Promotional inventory displays: an empirical analysis using IoT data

Jacob Z. Zeng
School of Business Administration, Gonzaga University
zijian.zeng@outlook.com

Ashish Agarwal, Ioannis Stamatopoulos
McCombs School of Business, The University of Texas at Austin
ashish.agarwal@mccombs.utexas.edu, yannis.stamos@mccombs.utexas.edu

Problem definition: Despite their widespread use, promotional inventory displays' campaign execution is opaque. Brands (manufacturers) can only verify in-store display presence through manual, on-site audits, which are costly and limited in scope. This lack of visibility into execution makes it difficult for brands to quantify the sales impact of displays and properly evaluate campaign performance.

Methodology/results: We use internet-of-things (IoT)-generated data on the real-time location of about 15 thousand displays from 10 display campaigns in about 5 thousand stores of a Fortune 500 retail chain, paired with the stores' point-of-sale (POS) data between September 2017 and March 2018, to measure the operational execution and effectiveness of display campaigns. First, we find that campaigns are poorly executed: 29 percent of displays never made it to a store's floor, and those that made it only spent 62 percent of the campaign there. Second, we find that poor execution deprives brands of substantial sales, especially during campaign weeks: placing a display on the floor during an arbitrary week increases the targeted products' sales by 7.3 percent, and placing it on the floor during a campaign week boosts sales by another 2.3 percent.

Managerial implications: Our results enable brands to optimize investments in promotional display campaigns and execution. We project that brands would suffer a 3.1 percent dollar sales decrease during campaign weeks by discontinuing current campaigns. However, they could achieve up to a 6.9 percent dollar sales increase during campaign weeks through improved execution.

Online appendix

		Proportion of:		
		All displays	Displays that carry the store brand	Displays that do not carry the store brand
Moved on floor–off floor:	On time–on time	0.02	0.01	0.02
	Early–on time	0.03	0.02	0.04
	Late–on time	0.03	0.01	0.04
	On time–late	0.06	0.05	0.06
	On time–early	0.08	0.12	0.06
	Early–late	0.10	0.08	0.12
	Late–early	0.10	0.12	0.09
	Late–late	0.13	0.05	0.17
	Early–early	0.16	0.27	0.11
	Never–never	0.29	0.27	0.30
Camp. days on/Camp. days:	0.00	0.32	0.31	0.33
	(0.00, 0.20]	0.09	0.12	0.07
	(0.20, 0.40]	0.08	0.14	0.05
	(0.40, 0.60]	0.10	0.14	0.07
	(0.60, 0.80]	0.12	0.08	0.15
	(0.80, 1.00)	0.14	0.10	0.15
	1.00	0.15	0.11	0.18

Table A

This table presents the distribution of "Moved on floor–off floor" status and of "Campaign days on/Campaign days" across displays. Following industry standards, we define "on time" as within three days from the planned date; "early" as at least four days before the planned date; and "late" as at least four days after the planned date. "Campaign days on" is a count of the campaign days a display spent on the floor, and "Campaign days" is the number of days in a campaign. Note that 29 percent of the displays were never placed on the floor and that only 15 percent of the displays spent the entire campaign period on the floor. The average value of Campaign days on/Campaign days is 0.44 for all displays, 0.39 for displays that carry store-brand products, and 0.48 for displays that do not carry store-brand products. Conditional on being positive, the same average is 0.62 for all displays, 0.56 for displays that carry store-brand products, and 0.71 for displays that do not carry store-brand products.

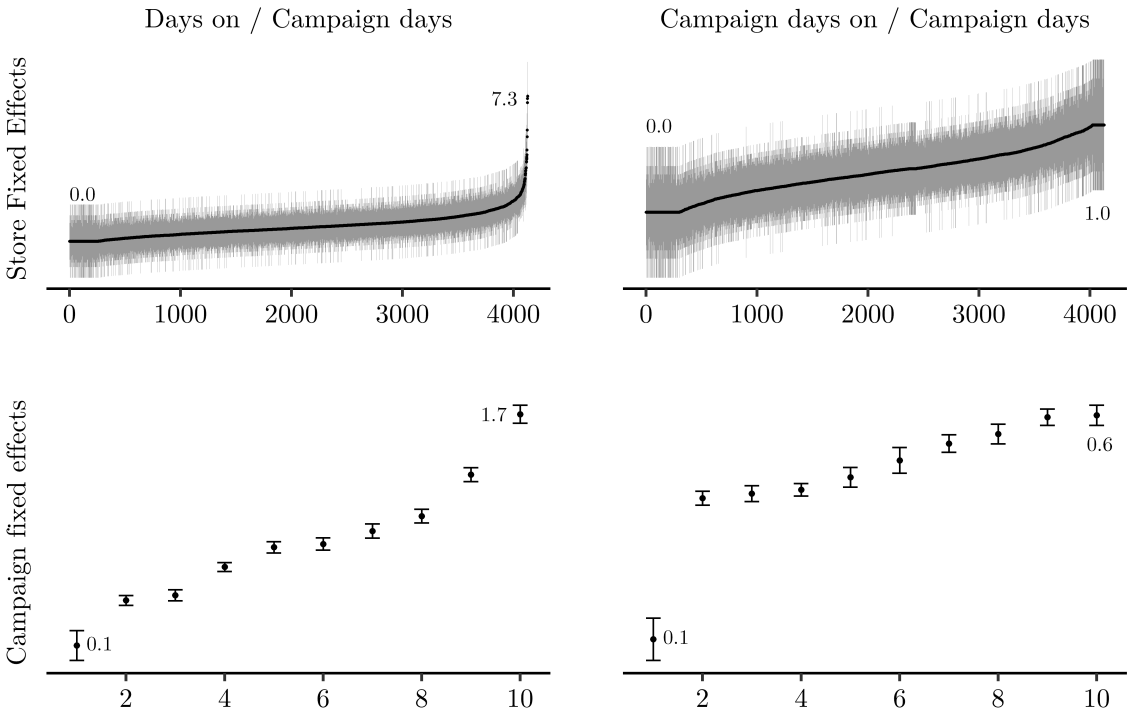

Figure A

The figures show regression coefficients from an OLS regression at the display level of Days on/Campaign days (left panel) and Campaign days on/Campaign days (right panel) on store fixed effects (top panel) and campaign fixed effects (bottom panel). The coefficients are sorted in ascending order and are accompanied by their 95 percent confidence intervals. The figures suggest that most stores achieve similar levels of compliance. In contrast, the different campaigns produced distinguishably different compliances.

	Top	Bottom	Difference
Number of products	27.47	25.99	1.48***
Tower display	0.83	0.83	0.01
Arrival date − Camp. start date	−1.39	−1.13	−0.27***
Arrived on the weekend	0.10	0.12	−0.02*
Number of other displays	0.51	0.57	−0.06*
Average product price	13.87	12.80	1.08***
Carries store-brand product	0.35	0.40	−0.05***
Number of past campaigns	3.65	4.38	−0.74***
Past unit sales	0.23	0.49	−0.26***
Had past promotion	0.24	0.34	−0.10***

Table B

The table reports the heterogeneity in the compliance drivers we examine in the main paper between top-complying stores and other stores. We define the top-complying stores as the stores corresponding to the highest 10 percent regression coefficients from the OLS regression at the display level of Days on/Campaign days on store fixed effects. Row definitions: "Number of products" is the number of different products carried by the display; "Tower display" is one for big (tower) displays; "Average product price" is the average price of the assortment carried by the display; "Past unit sales" is the total number of unit sales for all targeted products, divided by the number of targeted products, one week before the display arrived; "Carries store-brand product" is one if the display carried at least one store-brand product; "Arrival date − Campaign start date" is the number of days between the display's store arrival and the campaign start, divided by seven (so it is measured in weeks); "Arrived on the weekend" is one if the display arrived on the weekend; "Number of other displays" counts displays that meet four criteria: (i) they arrived at the store before the focal display, (ii) their campaign overlapped with the focal display's campaign, (iii) they made it to the floor, and (iv) they did not carry products in the same category as products carried by the focal display; "Number of past campaigns" is the number of past inventory display campaigns ran in the focal store by the time the display arrived; and "Had past price promotion" is one if at least one of the targeted products had a price promotion by the time the display arrived. All statistics are averages across stores. The p-values on the third column come from a simple t-test.

	Incremental F-statistics				
	On never	On early	On late	Days on/ Camp. days	Camp. days on/ Camp. days
	(1)	(2)	(3)	(4)	(5)
Number of products	1.67	22.22***	194.98***	334.06***	74.21***
Tower display	195.15***	12.45***	32.67***	9.47***	188.56***
Average product price	188.92***	102.53***	1.18	39.07***	27.77***
Past unit sales	430.67***	504.78***	145.46***	132.07***	148.97***
Carries store-brand product	64.07***	107.31***	135.98***	925.44***	135.31***
Arrival date - Camp. start date	299.63***	273.80***	121.33***	47.68***	89.28***
Arrived on the weekend	1.15	36.10***	7.04***	1.36	2.54
Number of other displays	384.37***	33.33***	9.58***	84.74***	130.88***
Number of past campaigns	21.01***	0.34	214.07***	274.50***	5.65**
Had past price promotion	253.42***	695.73***	729.75***	8.35***	72.01***
Store FEs	Yes	Yes	Yes	Yes	Yes
Observations	15,324	10,926	10,926	15,324	15,324

Table C

The table reports the F-statistics corresponding to the incremental R-squared associated with each driver for all performance metrics, and their respective p-values. We calculate the F-statistics as $((r_f - r_p)/(1 - r_f)) \times (n - k - 1)$, where r_f is the R-squared of the full model (with all the drivers), r_p is the R-squared of the partial model (without the focal driver), n is the number of observations, and k is the number of independent variables in the full model. We calculate the p-values with the F-distribution with $(1,\ n - k - 1)$ degrees of freedom. The rows correspond to potential compliance drivers, and the columns correspond to compliance measures. Row definitions: "Number of products" is the number of different products carried by the display; "Tower display" is one for big (tower) displays; "Average product price" is the average price of the assortment carried by the display; "Past unit sales" is the total number of unit sales for all targeted products, divided by the number of targeted products, one week before the display arrived; "Carries store-brand product" is one if the display carried at least one store-brand product; "Arrival date − Campaign start date" is the number of days between the display's store arrival and the campaign start, divided by seven (so it is measured in weeks); "Arrived on the weekend" is one if the display arrived on the weekend; "Number of other displays" counts displays that meet four criteria: (i) they arrived at the store before the focal display, (ii) their campaign overlapped with the focal display's campaign, (iii) they made it to the floor, and (iv) they did not carry products in the same category as products carried by the focal display; "Number of past campaigns" is the number of past inventory display campaigns ran in the focal store by the time the display arrived; and "Had past price promotion" is one if at least one of the targeted products had a price promotion by the time the display arrived. Column definitions: "On never" is one if the display never made it to the floor; "On early" is one if the display was moved to the floor at least four days before the campaign start date; "On late" is one if the display was moved to the floor at least four days after the campaign start date; "Days on/Campaign days" is the fraction of campaign length spent on the floor; "Campaign days on/Campaign days" is the fraction of the campaign window spent on the floor. Note that columns (2) and (3) focus on displays that made it to the floor and thus have fewer observations. *** p<0.01, ** p<0.05.

	On never	On early	On late	Days on/ Camp. days	Camp. days on/ Camp. days
	(1)	(2)	(3)	(4)	(5)
All drivers	0.430	0.681	0.681	0.400	0.395
All drivers, interacted with top	0.433	0.684	0.682	0.525	0.402
Incremental F-statistics	6.663***	5.388***	1.497***	292.344***	12.437***

Table D

The table reports the R-squared values and F-statistics corresponding to the incremental R-squared from interacting the drivers from Table E with "top store" dummies (where top stores are defined as in Footnote 10) for all performance metrics outcomes. We calculate the F-statistics with $((r_f - r_p)/(1 - r_f)) \times ((n - k - 1)/10)$, where r_f is the R-squared of the full model (with drivers plus interactions), r_p is the R-squared of the partial model (just the drivers, without the interactions), n is the number of observations, k is the number of independent variables in the full model, and 10 is the number of variables added to the partial model to get the full model. We calculate p-values with the F-distribution with $(10,\ n - k - 1)$ degrees of freedom. *** p<0.01.

	On never	On early	On late	Days on/ Camp. days	Camp. days on/ Camp. days
	(1)	(2)	(3)	(4)	(5)
Number of products	0.001	0.002***	−0.007***	0.018***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Tower display	0.223***	0.065***	0.103***	−0.107***	−0.201***
	(0.016)	(0.018)	(0.018)	(0.035)	(0.015)
Average product price	0.010***	0.009***	−0.001	−0.010***	−0.004***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Past unit sales	0.098***	−0.138***	0.072***	−0.118***	−0.053***
	(0.005)	(0.006)	(0.006)	(0.010)	(0.004)
Carries store-brand product	−0.081***	0.113***	−0.124***	−0.667***	−0.107***
	(0.010)	(0.011)	(0.011)	(0.022)	(0.009)
Arrival date - Camp. start date	−0.079***	−0.088***	0.057***	0.068***	0.039***
	(0.005)	(0.005)	(0.005)	(0.010)	(0.004)
Arrived on the weekend	0.015	−0.093***	0.040***	−0.036	−0.021
	(0.014)	(0.016)	(0.015)	(0.031)	(0.013)
Number of other displays	0.117***	−0.038***	0.020***	−0.119***	−0.063***
	(0.006)	(0.007)	(0.006)	(0.013)	(0.005)
Number of past campaigns	−0.011***	0.002	0.041***	−0.088***	−0.005**
	(0.002)	(0.003)	(0.003)	(0.005)	(0.002)
Had past price promotion	0.176***	−0.336***	0.336***	−0.069***	−0.086***
	(0.011)	(0.013)	(0.012)	(0.024)	(0.010)
Store FEs	Yes	Yes	Yes	Yes	Yes
Observations	15,324	10,926	10,926	15,324	15,324
R-squared	0.430	0.681	0.681	0.400	0.395

Table E

The table reports the regression outputs for different drivers. An observation is a display. The rows correspond to potential compliance drivers, and the columns correspond to compliance measures. Row definitions: "Number of products" is the number of different products carried by the display; "Tower display" is one for big (tower) displays; "Average product price" is the average price of the assortment carried by the display; "Past unit sales" is the total number of unit sales for all targeted products, divided by the number of targeted products, one week before the display arrived; "Carries store-brand product" is one if the display carried at least one store-brand product; "Arrival date − Campaign start date" is the number of days between the display's store arrival and the campaign start, divided by seven (so it is measured in weeks); "Arrived on the weekend" is one if the display arrived on the weekend; "Number of other displays" counts displays that meet four criteria: (i) they arrived at the store before the focal display, (ii) their campaign overlapped with the focal display's campaign, (iii) they made it to the floor, and (iv) they did not carry products in the same category as products carried by the focal display; "Number of past campaigns" is the number of past inventory display campaigns ran in the focal store by the time the display arrived; and "Had past price promotion" is one if at least one of the targeted products had a price promotion by the time the display arrived. Column definitions: "On never" is one if the display never made it to the floor; "On early" is one if the display was moved to the floor at least four days before the campaign start date; "On late" is one if the display was moved to the floor at least four days after the campaign start date; "Days on/Campaign days" is the fraction of campaign length spent on the floor; "Campaign days on/Campaign days" is the fraction of the campaign window spent on the floor. Note that columns (2) and (3) focus on displays that made it to the floor and thus have fewer observations. Robust standard errors are in parentheses, clustered at the store level; *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

	On never	On early	On late	Days on/ Camp. days	Camp. days on/ Camp. days
	(1)	(2)	(3)	(4)	(5)
Number of products	0.009	0.027	−0.105	0.210	0.071
Tower display	0.086	0.019	0.040	−0.032	−0.103
Average product price	0.085	0.056	−0.008	−0.065	−0.039
Past unit sales	0.107	−0.113	0.079	−0.099	−0.076
Carries store-brand product	−0.044	0.047	−0.068	−0.285	−0.078
Arrival date − Camp. start date	−0.124	−0.103	0.090	0.083	0.082
Arrived on the weekend	0.006	−0.028	0.016	−0.011	−0.011
Number of other displays	0.109	−0.027	0.019	−0.086	−0.078
Number of past campaigns	−0.033	0.004	0.120	−0.198	−0.021
Had past price promotion	0.100	−0.144	0.192	−0.031	−0.065

Table F

The table reports the power of each driver in explaining between-store variation in compliance. Specifically, we construct this table by multiplying each coefficient in Table E by the between-store standard deviation of its corresponding compliance driver, and then dividing it by the between-store interquartile range (IQR) of its corresponding compliance metric. For example, the coefficient of "Number of products" for "Days on/Camp. days" in Table E is 0.018, the standard deviation of "Number of products" is 7.56, and the IQR of "Days on/Camp. days" is 0.646. Therefore, everything else equal, adding 7.56 more products to every campaign would cover $0.018 \times 7.56/0.646 = 0.210$ or 21 percent of the distance between the 25-th top-complying store and the 75-th top-complying store, when performance is measured according to "Days on/Camp. days."

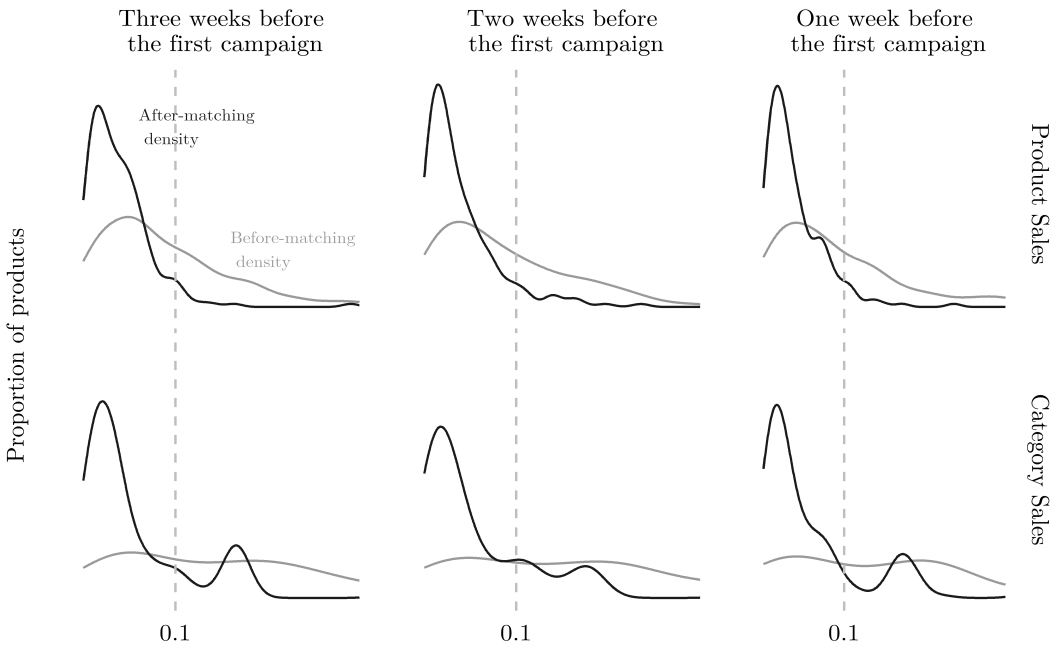

Figure B

This figure depicts the similarity in the sales patterns of targeted and non-targeted stores before and after matching, product-by-product. Specifically, for each of the three weeks leading to the first campaign that targets a product, we calculate the absolute standardized difference in product sales (first row) and category sales (second row) between targeted and non-targeted stores, before and after matching. The black curves correspond to the after-matching densities of these standardized differences (across products), and the gray curves correspond to the before-matching densities. The gray dashed line marks the absolute standardized difference of 0.1, the literature standard for valid matching. Note that almost all of the after-matching densities are to the left of the gray dotted lines, indicating that the matching procedure was successful.

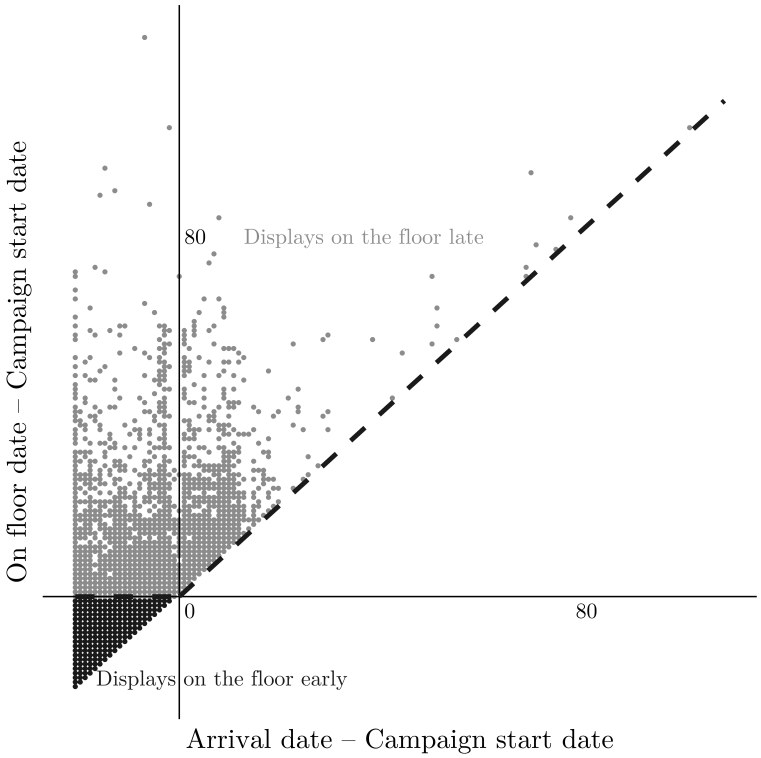

Figure C

This figure demonstrates the relationship between a display's arrival at a store and its placement on the floor. The horizontal axis marks the difference between a display's store arrival date and its corresponding campaign start date (negative values indicate early store arrival), and the vertical axis marks the difference between a display's on-floor date and its corresponding campaign start date (negative values indicate early on-floor placement). Black dots are displays that were placed on the floor early, gray dots are displays that were placed on the floor late, and the dashed line represents perfect compliance given the display's arrival date. Note that, perhaps surprisingly, several of the displays that arrived at a store early were placed on the floor early. Anecdotal evidence suggests that this occurs because the displays take up space in the stores' backrooms.

	Display on floor	
	(1)	(2)
During campaign	0.269***	0.388***
	(0.003)	(0.004)
Arrival date − Campaign start date	-0.011***	-0.003*
	(0.001)	(0.002)
During campaign × (Arrival date − Campaign start date)	0.026***	0.044***
	(0.001)	(0.001)
Product–store FEs	Yes	Yes
Product–week FEs	Yes	No
Product–month FEs	No	Yes
R-squared	0.570	0.639
First-stage F-Statistic	3,018	2,233
Observations	3,099,476	900,230

Table G

This table shows the output of the first-stage instrumental variable (IV) regressions of both the weekly-level and the monthly-level analyses. Column (1): The unit of observation is a product–store–week. The sample consists of product–store–weeks such that there is a campaign that targets this product at this store at most three weeks away from this week. We code the display dummy, "Display on floor," as one if at least one display that carried this product spent at least one day of that week on that store's floor; and we code the campaign dummy, "During campaign," as one if at least one display was planned to be on that store's floor for at least one day of that week. Alternative specifications of the display and campaign dummy variables (e.g., demanding that the display spent at least half the week on the store's floor) yield qualitatively similar results. To construct the instrument, "Arrival date − Campaign start date," for each product–store–week, we identify the corresponding campaign and displays, calculate the number of days between each display's store arrival and the campaign start, divide it by seven (so it is measured in weeks), and then take the median value across all corresponding displays. The instrument is predictive of treatment (display on floor). Robust standard errors are in parentheses, clustered at the product–store level; *** $p<0.01$. Column (2): The unit of observation is a product–store–month. The instrument is predictive of treatment (display on floor) in this specification too. The sample and variable definitions are analogous to those of Column (1). As in Column (1), robust standard errors are in parentheses, clustered at the product–store level; *** $p<0.01$, * $p<0.10$.

	log(Sales in units of product)			log(Sales in USDs)		
	During campaign	Display on floor	Interaction term	During campaign	Display on floor	Interaction term
Minimum	0.004	0.023	0.003	0.006	0.070	0.018
First Quartile	0.004	0.025	0.004	0.007	0.072	0.021
Median	0.004	0.025	0.005	0.007	0.074	0.022
Mean	0.004	0.025	0.005	0.007	0.073	0.022
Third Quartile	0.005	0.026	0.005	0.008	0.074	0.023
Maximum	0.005	0.028	0.007	0.009	0.079	0.027

Table H

This table summarizes the distribution of regression coefficients we get by applying our main specification (i.e., the regression specification behind Table 1's columns (2) and (6)) on 30 product-by-product, random store-orderings.

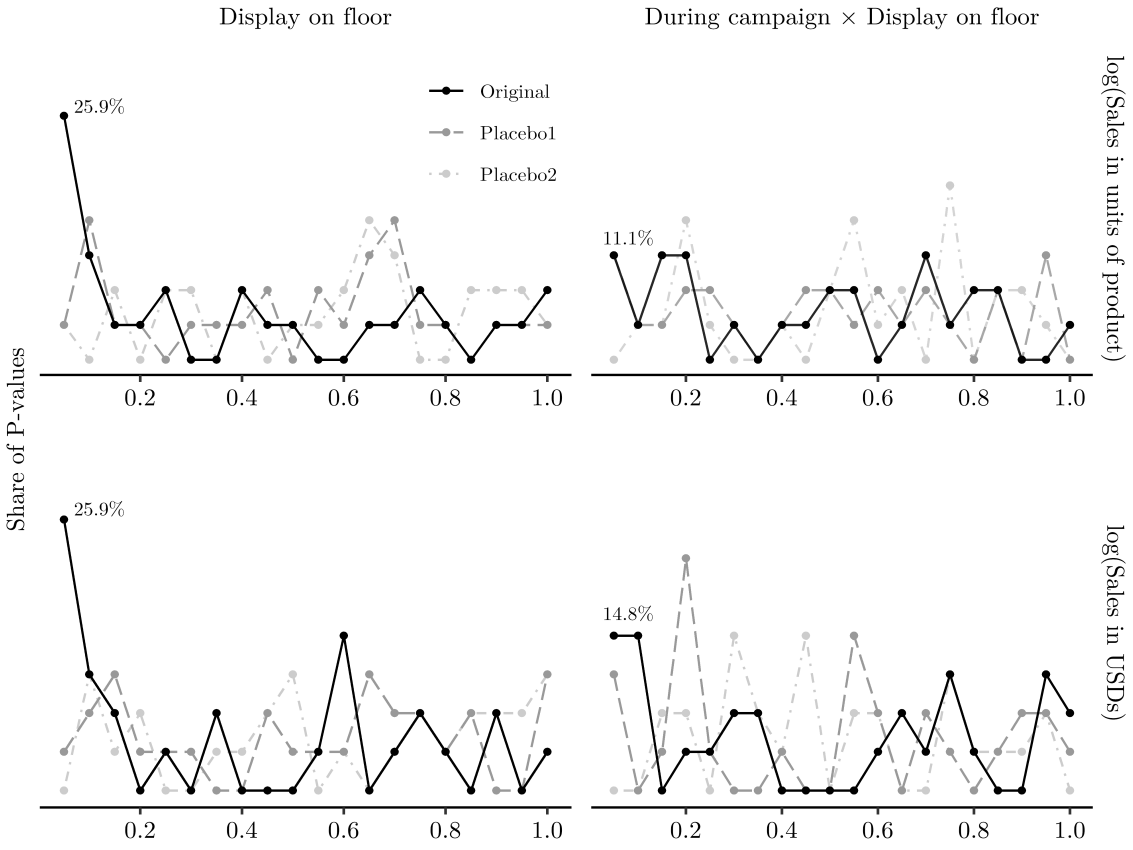

Figure D

This figure depicts p-curves (i.e., distributions of p-values) corresponding to the original dataset and the two placebo datasets. To construct these p-curves, we repeat our identification strategy (matching of stores, IV regression, fixed-effects regression) for each product that was affected by at least two campaigns during our period of study (27 of 186 products), and we collect the second-stage regression p-values of the display dummy and the interaction of the display dummy with the campaign dummy. We execute this process both for the original dataset and for the two placebo datasets and then plot the distribution of the corresponding p-values. To create the first placebo dataset, we first randomly select 30 percent of the matched product–store pairs, where each pair consists of a "treated" product–store (i.e., one targeted by a campaign) and a "control" product store (i.e., one not targeted by a campaign). Second, for each control product–store we copy its corresponding treated product–store's campaign and display movement information. Third, we drop all treated product–stores from the data (100 percent of them, not just the 30 percent selected in the first step). To create the second placebo dataset, we first select 50 percent of the control product–stores and randomly assign each to a treated product–store. This assignment can be one-to-many. Second, for each of these selected control product–stores, we copy their corresponding treated product–store's campaign and display movement information. Third, we drop all treated product–stores from the data. We test the null that a p-curve is generated from a uniform distribution in $[0, 1]$ using Fisher's method (see Simonsohn et al. 2014). The Fisher-method p-values for this null are (original, placebo 1, placebo 2): top left (0.000, 0.335, 0.856); top right (0.049, 0.153, 0.836); bottom left (0.000, 0.433, 0.848); bottom right (0.049, 0.150, 0.751).

	log(Sales in units of product)					
	(1)	(2)	(3)	(4)	(5)	(6)
During campaign	0.004***	0.005***	0.004***	0.005***	0.004***	0.003*
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Display on floor	0.025***	0.023***	0.025***	0.024***	0.025***	0.031***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
During camp. × Display on floor	0.005*	0.005*	0.005*	0.005*	0.005*	0.004*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
R-squared	0.531	0.531	0.531	0.531	0.531	0.531
	log(Sales in USDs)					
	(7)	(8)	(9)	(10)	(11)	(12)
During campaign	0.007*	0.008**	0.007*	0.007*	0.007*	0.004
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Display on floor	0.073***	0.070***	0.073***	0.073***	0.074***	0.085***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
During camp. × Display on floor	0.023***	0.023***	0.022***	0.022***	0.022***	0.022***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
R-squared	0.412	0.412	0.412	0.412	0.412	0.412
Product-store FEs	Yes	Yes	Yes	Yes	Yes	Yes
Product-week FEs	Yes	Yes	Yes	Yes	Yes	Yes
Arrival date - Camp. start date IV	Yes	Yes	Yes	Yes	Yes	No
Squared IV, version 1	No	Yes	No	No	No	No
Squared IV, version 2	No	No	Yes	No	No	No
Squared IV, version 3	No	No	No	Yes	No	No
Arrival of display on a weekend IV	No	No	No	No	Yes	Yes
First-stage F-Statistics	3,018	2,035	2,715	1,829	1,532	117
First-stage robust F-Statistics	498	984	1367	1,400	507	29
Observations	6,198,952	6,198,952	6,198,952	6,198,952	6,198,952	6,198,952

Table I

The table reports the main results using four different alternative first-stage regressions. Columns (1) and (7) are our original results for unit sales and dollar sales respectively. Columns (2)—(4) and Columns (8)–(10) are the results obtained by incorporating both the original IV and its second-order terms. We use three different definitions for the second-order term of our original IV. In the first version, we define the second-order term as the square of the original IV value (the difference between the arrival date and the campaign start date). (See results in Columns (2) and (8).) Although this approach is straightforward, it yields a squared term that is always positive. To address this, we also used a second version, where the second-order term was defined as the product of the original IV with its absolute value. (See results in Columns (3) and (9).) This approach ensures that the second-order term has the same sign as the original IV value. We also employed a third version, in which we separated the second-order term into its positive part and its negative part. The positive part is the square of the original IV value times an indicator that the original IV is greater than 0. The negative part is the negative square of the original IV value times an indicator that the original IV is less than or equal to 0. (See results in Columns (4) and (10).) Columns (5) and (11) report results obtained by incorporating both the original IV and the weekend display arrival—defined as one if a display that carries that product arrived at that store during a weekend, and zero otherwise—as an additional instrument. Columns (6) and (12) report results obtained by incorporating the weekend display arrival as an instrument only.

References

Simonsohn, Uri, Leif D Nelson, Joseph P Simmons. 2014. P-curve: a key to the file-drawer. *Journal of experimental psychology: General* 534. 12