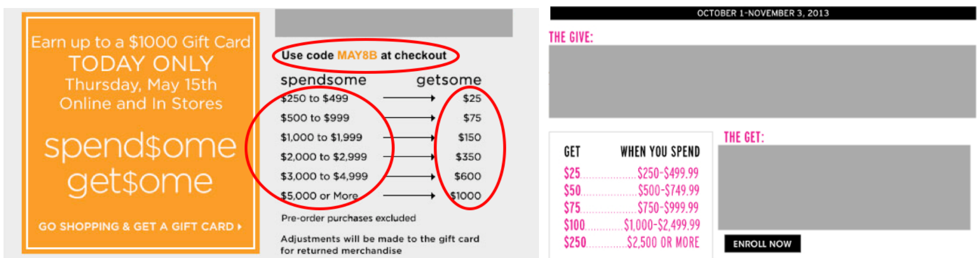
Electronic Companion to “Dual Value of Delayed Incentives: An Empirical Investigation of Gift Card Promotions”

EC.1. Examples of Gift Card Promotions

In Figure EC.1a, we provide examples of gift card promotions offered by luxury department stores. Some sections of these images have been shaded to conceal the identity of our partner retailer. Figures EC.1b and EC.1c provide examples of a gift card promotion run by a big-box retailer (Target) and a consumer electronic store (Best Buy), respectively. In these examples, a customer needs to spend more than specific expenditure levels during a predetermined time window to receive a gift card redeemable with the same retailer on a future purchase.



(a) Gift card promotions offered by luxury department stores.

(b) Gift card promotion offered by Target.

(c) Gift card promotion offered by Best Buy.

Figure EC.1 Examples of gift card promotions in luxury fashion and retail industries.

EC.2. Selection of Gift Card Promotions for RD Analysis

Customers in our dataset may be potentially treated in multiple gift card promotions. As a result, it is possible that the customer’s purchase recency with respect to a focal gift card promotion may be impacted by their exposure to a prior gift card promotion. In particular, participating/purchasing in the prior promotion lowers the customer’s purchase recency with respect to focal promotion and thus may determine which side of the focal promotion’s threshold the customer falls into and hence, their probability of being targeted in the focal promotion. Hence, the distribution of customer purchase recency (i.e., the Running_var) may be endogenous. Below we discuss the implications of multiple gift card promotion exposures on the identification of causal effects using RD design in our empirical setting.

We first note that the LATE estimation in the RD design for a particular purchase recency threshold uses customers only within the bandwidth of consideration around that threshold. We elaborate on how the bandwidths around each threshold are computed in Section 4.2.4. Therefore, we focus our attention on the impact of prior gift card promotion exposures on customers within the bandwidth. If customers within the bandwidth of the focal gift card promotion were treated uniformly in a prior promotion, then the discontinuity in the focal promotion's targeting policy still remains a valid exogenous shock that is amenable to causal analysis. However, if customers within the bandwidth were targeted discontinuously based on their purchase recency in the prior gift card promotion (i.e, there is a discontinuity in their treatment probability for the prior promotion), then the discontinuity in the focal gift card promotion is not conducive to causal analysis. In particular, the LATE estimated for the focal promotion conflates the effect of the two gift card promotions.

The retailer's promotion calendar revealed that four gift card promotions (corresponding to five thresholds) were subject to the identification issue discussed above, and so we excluded them from our analysis. In particular, the thresholds in the omitted gift card promotions overlapped with that of the threshold used in a preceding gift card promotion. As a result, the discontinuity in the targeting probability of the focal promotion does not identify the causal effect of the focal promotion; rather, it identifies the causal effect of *repeated* gift card promotion exposures, which is arguably diluted. Of the seven thresholds that we consider for RD analysis in the paper, three thresholds are not affected by this endogeneity issue by design—there were no gift card promotions in the preceding three months of these promotions. The remaining four thresholds do not suffer from the endogeneity issue because there are no other targeting discontinuities within the bandwidth of analysis arising from targeting policy in a prior promotion.

Empirically, the above endogeneity issue manifests in the form of a weaker instrument ($\mathbf{1}_{\{\mathsf{Running\_var}_{ij} \leq 0\}}$) in the two-stage IV regression (3)–(4). In particular, for promotions with an invalid RD design, we expect that the targeting policy in a prior promotion has a greater effect (compared to the instrument used in the focal promotion) on the probability of receipt of promotion email in the focal promotion (i.e., in the first-stage regression specified in Equation 3). To that end, we include the number of prior gift card promotion emails received in the past one, two, and three months, (each in a separate regression) as additional controls in the first-stage regression. In line with the above logic, we observe that the instruments of the four excluded promotions are weak. Furthermore, for the gift card promotions included in the RD analysis, we find that the coefficient of the instrument ($\mathbf{1}_{\{\mathsf{Running\_var}_{ij} \leq 0\}}$) and the predictive power of the first-stage regression remain unaffected (with and without the additional controls). These findings further validate the selection of the gift card promotions for the RD analysis. In Table EC.1 in Appendix EC.3, we present the estimates obtained in the first-stage of the 2SLS regression specified in Equation (3), including various performance measures. Further details of this analysis are available from the authors upon request.

EC.3. 2SLS First-Stage Regression Results

In Table EC.1, we present the estimates obtained in the first-stage of the 2SLS regression specified in Equation (3). The coefficient estimates of the binary (instrument) variable $\mathbf{1}_{\{\mathsf{Running\_var}_{ij} \leq 0\}}$, denoted in the table by inst, represents the magnitude of the discontinuity in targeting probability at corresponding thresholds; see also Figure 2.

Threshold	PD	inst	s.e.	N	R^2	RMSE	R^2-Adj	AIC
1	1	0.258	0.005	114,614	0.156	0.459	0.156	-178,386
	2	0.270	0.004	184,489	0.154	0.460	0.154	-286,897
2	1	0.216	0.008	53,202	0.245	0.422	0.244	-91,809
	2	0.227	0.006	101,230	0.255	0.420	0.255	-175,573
3	1	0.459	0.009	25,508	0.399	0.367	0.399	-51,065
	2	0.466	0.010	21,920	0.390	0.371	0.390	-43,488
4	1	0.753	0.004	127,516	0.567	0.320	0.567	-290,897
	2	0.757	0.004	176,238	0.574	0.316	0.574	-405,645
5	1	0.287	0.005	93,357	0.195	0.353	0.195	-194,355
	2	0.289	0.004	110,388	0.195	0.356	0.195	-227,706
6	1	0.490	0.005	85,328	0.416	0.381	0.416	-164,826
	2	0.514	0.004	178,670	0.406	0.385	0.406	-340,884
7	1	0.383	0.008	57,837	0.225	0.440	0.225	-94,904
	2	0.421	0.004	157,757	0.232	0.438	0.232	-260,682

Table EC.1 First-stage of the 2SLS regression specified in (3) with all the control variables included in the specification.

Notes: PD refers to degree of the polynomial functions (f, g) used in the RD estimation.

EC.4. Qualification Stage Data Preparation

We consider only customers whose emails are included in the retailer's master email list (see Section 3.3 for the details) and whose purchase recencies are around the thresholds in our dataset. To ensure validity and improve explanatory power of our empirical models, we excluded customers with abnormally high number and/or amount of purchases. Our conversations with the partner retailer executives revealed that such abnormally high transactions most likely indicate business accounts (a business account—as opposed to a personal (or individual) account—refers to a customer who shops on behalf of a company), which are unrepresentative of the general customer population.

To identify these *outliers*, we applied two criteria. First, we dropped any customer whose purchase frequency is within the highest 0.3 percentile of customers included in S^{4mo} and the highest 0.3 percentile of customers included in S^{13mo} and S^{16mo} combined together. Second, for each purchase recency (i.e., $\mathsf{Purchase\_rec}_{it}$) level around S^{13mo} and S^{16mo}, we computed the mean and standard deviation of the customers' qualification stage expenditures (i.e., $\mathsf{Expenditure}^{Q}_{it}$). Then we dropped customers whose $\mathsf{Expenditure}^{Q}_{it}$ are greater than three standard deviations above the mean. This criterion leads to the exclusion of 0.004% of the observations in our dataset (or 1.3% of observations who make a purchase during the qualification stage of a gift card promotion).

EC.5. Verifying Imprecise Control

The standard test to verify imprecise control assumption in RD designs is the McCrary test, proposed by McCrary (2008). To understand the intuition behind this test, recall that if the imprecise control assumption did not hold, i.e., if customers were able to precisely manipulate their purchase times to receive the gift card promotion email, then we would see significantly more customers on the targeted side of the threshold compared to the other side (within a small bandwidth from the threshold). Therefore, showing that the distribution of running variable, i.e., customers' purchase recency, continuously changes at the threshold provides a strong argument validating the identification assumption. Using this reasoning, McCrary (2008) provides a non-parametric test, based on local polynomial regression, to estimate the jump in the (logarithm of) density function of the running variable. The test provides an estimate for the difference in the logarithm of the density function below and above the threshold. In Table EC.2, we present the results of this test—performed separately for each threshold. The test results indicate that there is no statistically significant evidence to suggest that customers manipulate their purchasing times with the motivation of receiving the gift card promotion email. Specifically, all thresholds except $j = 5$ show no statistically significant jump in the density function of the running variable at the threshold. Furthermore, the sign of the (statistically significant) jump at threshold $j = 5$, if anything, provides evidence against manipulation of the running variable, i.e., customers' manipulating their purchase behavior to *not* receive the gift card promotion email. In particular, for threshold $j = 5$, the test estimates a positive jump from the left side of the threshold to the right side of it, even though the retailer targets the left side of the threshold for the gift card promotion; see Figure 2b. We provide a graphical illustration of the test in Figure EC.2.

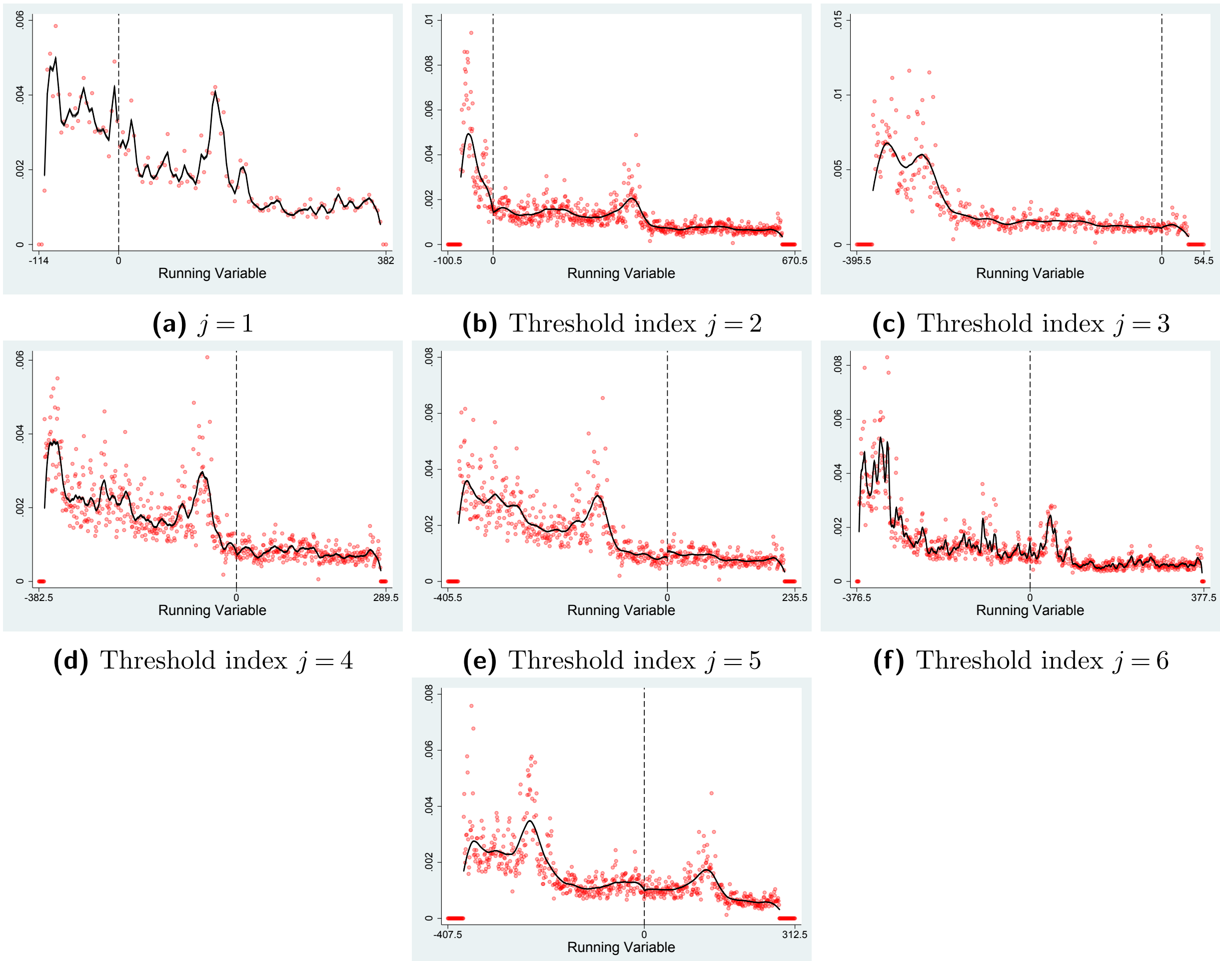
Index (j)	1	2	3	4	5	6	7
Est (diff in log)	-0.034	-0.009	0.041	-0.033	0.230***	-0.022	-0.018
	(0.022)	(0.017)	(0.026)	(0.042)	(0.027)	(0.036)	(0.022)

Standard errors are in parentheses. * $p < .1$, ** $p < 0.05$, *** $p < 0.01$

Table EC.2 Estimates of discontinuity in running variable obtained from the McCrary test.

EC.6. Continuity of Control Variables

We first graphically check how the control variables vary as a function of the running variable around each threshold. In Figures EC.3–EC.5, we present scatter plots of each of the control variables (log(Avg_exp), Purchase_freq, Promo_email_freq, Web_freq) against the running variable for the all thresholds. Each point in these plots is the average value (across customers) of the control variables at each value of running variable. These scatter plots suggest that there are no noticeable discontinuities in the control variables around the threshold. Next, we formally test for the presence of a discontinuity in each of the four control variables at each of the seven thresholds using a local RD analysis (which is based on Section 3.2.5 of Cattaneo et al. 2023). Specifically, we use a local polynomial (degree one and two) regression with uniform kernel function and corresponding data-driven bandwidths to estimate the intent-to-treat (ITT) effect at the threshold by treating each control variable (log(Avg_exp), Purchase_freq, Promo_email_freq, Web_freq) as the outcome variable in the RD analysis. The ITT effect measures the effect of customers being on either side of the threshold, i.e., the effect of the binary threshold indicator $\mathbf{1}_{\{\mathsf{Running\_var}_{ij} \geq 0\}}$ on the outcome variable. A statistically significant coefficient estimate of the threshold indicator would point to the presence of a discontinuity in the control variable at the threshold. We present the coefficient estimates of the threshold indicator in Table EC.3. The estimates presented in Table EC.3 suggest that there are no robust, statistically significant discontinuities in the control variables at each of the seven thresholds (except for Web_freq for threshold index 6), thus validating the RD design. We further investigate the discontinuity in Web_freq for threshold index 6 by considering observations closer to the threshold. We find that the (mean, std. err.; nbr. of obs.) within a bandwidth of 3, 2, 1 are (-1.97, 1.27; 7,962), ($-.847$, 1.45; 6,282), and (2.26, 2.04; 3,896), respectively. None of these effects sizes are statistically significant (at *p-value* < 0.1). These volatile results point to outlier values around the threshold potentially explaining the observed discontinuity in Web_freq.



(a) $j = 1$ **(b)** Threshold index $j = 2$ **(c)** Threshold index $j = 3$

(d) Threshold index $j = 4$ **(e)** Threshold index $j = 5$ **(f)** Threshold index $j = 6$

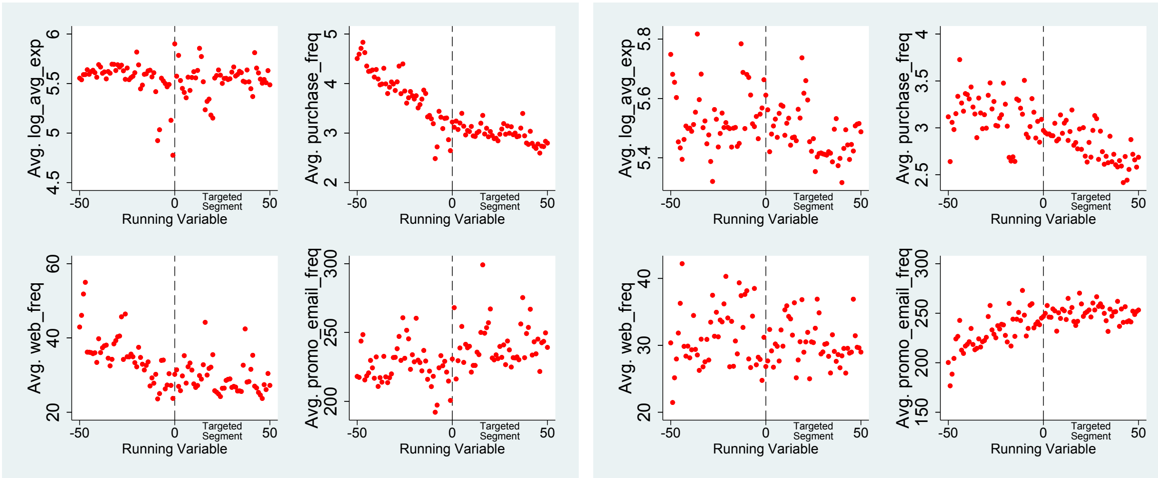
(g) Threshold index $j = 7$

Figure EC.2 Density of running variable smoothened using local linear regression separately on both sides of the threshold (McCrary 2008).

Furthermore, evidence from other validation checks (e.g., imprecise control condition is satisfied for threshold index 6 and Web_freq is continuous at all of the other thresholds), also suggests that the discontinuity in Web_freq for threshold index 6 alone is not a cause for concern.

EC.7. Exclusion Restriction

We empirically test if the instrument variable ($\mathbf{1}_{\{\text{Running\_var} \geq 0\}}$) in our fuzzy RD analysis satisfies the exclusion restriction by conducting the following local RD analysis suggested in Section 3.2.2 of Cattaneo et al. (2023). Consider customers who are in the neighborhood of the threshold (which is normalized to 0). Conditional on a customer receiving the treatment (gift card promotion email), exclusion restriction implies that the instrument variable—whether the customer is to the left or right of the threshold—should *not* have an effect on the outcome (because the threshold indicator should only affect the outcome through the treatment status which is now fixed). We conduct this local RD analysis around the threshold. In Table EC.4, we present the coefficient estimates of the threshold indicator $\mathbf{1}_{\{\text{Running\_var} \geq 0\}}$ (denoted by inst) on the outcome variable of interest (Expenditure^{Q}) conditional on customers not receiving the treatment (columns 3 and 4) and receiving



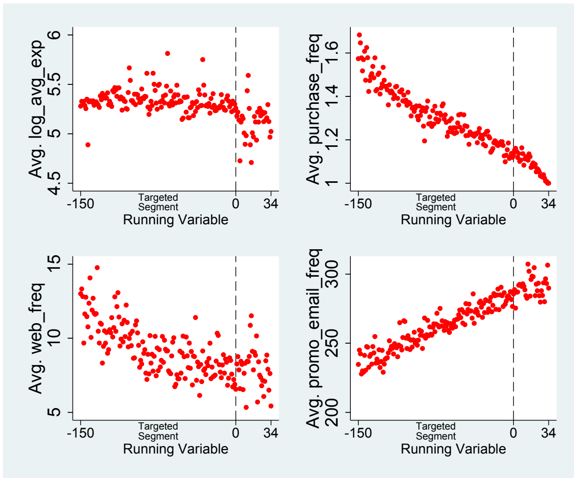
(a) Threshold index $j = 1$. **(b)** Threshold index $j = 2$.

Figure EC.3 Scatter plots for controls (vs. running variable) at thresholds in set S^{4mo}.

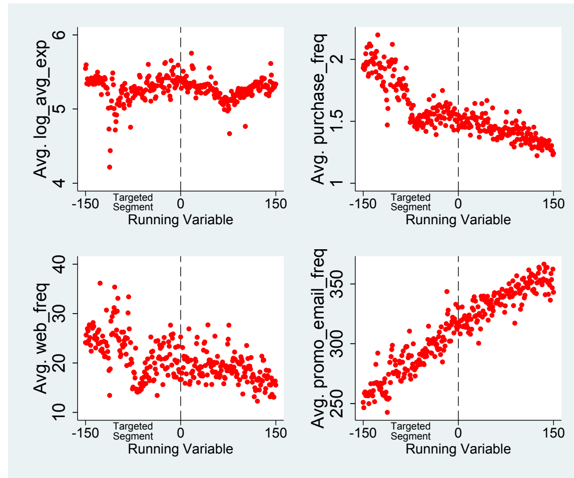
Threshold index (j)	log(Avg_exp) PD = 1	log(Avg_exp) PD = 2	Purchase_freq PD = 1	Purchase_freq PD = 2	Web_freq PD = 1	Web_freq PD = 2	Promo_email_freq PD = 1	Promo_email_freq PD = 2
1	-0.004	-0.003	-0.014	-0.029	-1.004	-0.728	-1.581**	0.948
	(0.003)	(0.002)	(0.022)	(0.020)	(0.644)	(0.543)	(0.741)	(0.619)
bw^*	8	17	16	21	9	14	14	20
Nbr. Obs	48,042	88,711	84,667	110,232	55,627	73,705	73,705	106,312
2	0.003	0.003	-0.024	-0.019	-0.527	-0.706	3.074	-0.028
	(0.005)	(0.002)	(0.035)	(0.024)	(1.004)	(0.666)	(2.909)	(2.473)
bw^*	5	25	12	26	8	20	15	21
Nbr. Obs	12,932	78,993	30,719	82,689	21,026	59,289	41,699	63,501
3	0.001	0.001	-0.004	0.001	0.289	-0.352	0.05	-2.764
	(0.001)	(0.001)	(0.003)	(0.002)	(0.351)	(0.291)	(2.362)	(2.027)
bw^*	13	23	20	32	16	24	19	27
Nbr. Obs	19,895	36,314	32,723	50,443	25,508	37,664	30,824	42,433
4	4e-5	0.001	-0.001	-0.007	-0.840	0.298	4.821	0.954
	(0.002)	(0.001)	(0.009)	(0.006)	(0.812)	(0.658)	(3.070)	(2.718)
bw^*	11	49	24	53	20	30	16	21
Nbr. Obs	13,394	64,368	30,049	73,720	25,766	37,109	20,163	26,545
5	0.001	0.001	-0.004	-0.003	0.058	-0.294	2.323	-0.948
	(0.001)	(0.001)	(0.007)	(0.006)	(0.613)	(0.423)	(2.181)	(1.900)
bw^*	21	42	29	50	24	48	31	41
Nbr. Obs	27,633	54,086	36,953	66,114	31,212	63,695	39,443	53,108
6	0.001	0.001	0.002	0.021**	-3.697***	-3.765***	4.787*	0.742
	(0.002)	(0.001)	(0.016)	(0.010)	(0.841)	(0.566)	(2.501)	(1.601)
bw^*	12	25	7	20	8	18	25	64
Nbr. Obs	25,397	53,064	14,606	39,931	16,707	35,984	53,064	156,114
7	-2e-5	-9e-4	-0.007	-0.004	-0.863	-0.635	-4.037	0.104
	(0.002)	(0.001)	(0.009)	(0.008)	(0.819)	(0.435)	(2.725)	(2.160)
bw^*	11	18	24	30	15	51	38	58
Nbr. Obs	16,736	29,216	38,451	47,977	23,218	80,173	59,623	91,529

Table EC.3 Intent-to-treat (ITT) effect on control variables at each threshold

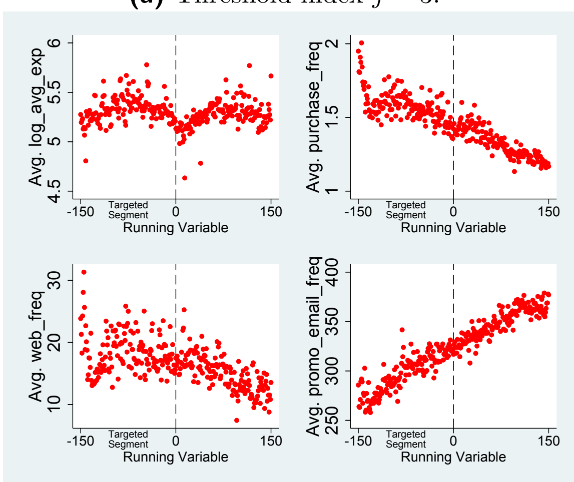
Notes: Standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. For each combination of outcome variable, threshold index and polynomial degree (one and two), the optimal bandwidth was computed using uniform kernel function based on the approaches in Calonico et al. (2014). Additional control variables are included in the regressions.



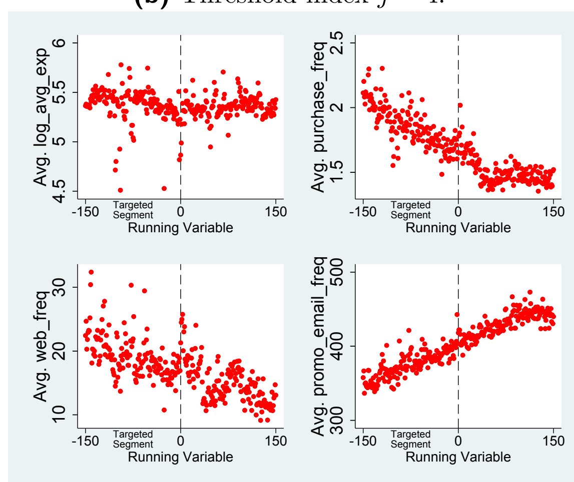
(a) Threshold index $j = 3$.



(b) Threshold index $j = 4$.



(c) Threshold index $j = 5$.



(d) Threshold index $j = 6$.

Figure EC.4 Scatter plots for controls (vs. running variable) at thresholds in set $S^{13\text{mo}}$.

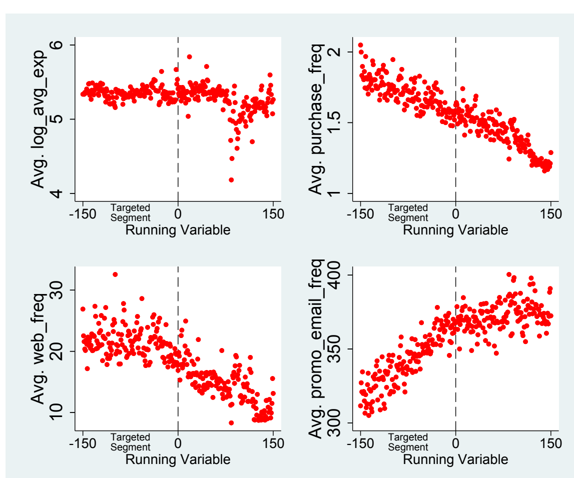


Figure EC.5 Scatter plot for controls (vs. running variable) for threshold index $j = 7$ in set $S^{16\text{mo}}$.

the treatment (columns 5 and 6). We do not find a robust, statistically significant effect of inst. Thus, results from this RD analysis further suggest that the instrument variable satisfies the exclusion restriction.

		Not treated		Treated	
Index (j)	PD	bw^*	inst	bw^*	inst
1	1	28	1.67 (1.11)	18	0.22 (0.92)
	2	48	2.07** (0.86)	27	0.24 (1.12)
2	1	41	0.53 (1.22)	21	2.15 (1.56)
	2	53	-0.03 (1.26)	41	2.4** (1.19)
3	1	29	0.82 (0.53)	28	-1.12 (0.75)
	2	31	0.68 (0.57)	30	-1.1 (1.01)
4	1	67	-0.03 (0.62)	40	2.35* (1.36)
	2	120	-0.3 (0.54)	96	1.78 (1.08)
5	1	41	-0.09 (0.56)	58	-1.82 (1.94)
	2	75	0.41 (0.49)	76	-3.14 (2.24)
6	1	61	0.58 (0.52)	39	0.21 (0.70)
	2	85	0.58 (0.55)	66	0.74 (0.66)
7	1	53	-0.24 (0.32)	45	0.51 (0.55)
	2	85	-0.43 (0.30)	78	-0.02 (0.44)

Table EC.4 Effect of instrument (inst) on Expenditure^Q conditional on customers' treatment (GC_email) status.

Notes: Standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. For each combination of threshold index and polynomial degree, we computed the optimal bandwidth based on the approaches provided in Calonico et al. (2014). Control variables are included in the regressions.

EC.8. Placebo Thresholds

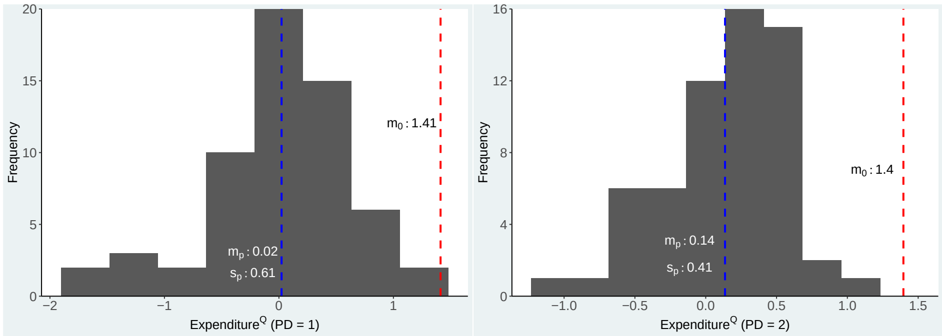
We follow Cattaneo et al. (2020) to implement the placebo analysis to estimate the intent-to-treat (ITT) effect corresponding to the two primary outcome variables in our analysis: total effect of gift card promotion on Expenditure^Q and Purchase^Q, respectively. To estimate the placebo effects at the 4- and 13-month threshold sets, we rely on local (linear and quadratic) polynomial regressions estimated within a bandwidth around the placebo threshold. Importantly, for this analysis we use only observations above (below) the true threshold for estimating the effect at placebo thresholds which are above (below) the true threshold. We do so to avoid "contamination" that results from the treatment effect at the true threshold (Imbens and Lemieux 2008, Cattaneo et al. 2020). Specifically, when the bandwidth of analysis around a placebo threshold includes a true discontinuity in treatment probability, the placebo effect may be contaminated due to the treatment effect at the *true* threshold.

We choose 60 placebo thresholds -50, -49, ...,-21, 21, ..., 49, 50 for this analysis. Recall that the true thresholds in the seven RD designs are normalized to zero (see Section 4.1.1). We exclude placebo thresholds $-20, -19, \ldots, -1, 1, \ldots, 19, 20$ to avoid potentially underpowering the placebo analysis. Specifically, excluding 20 placebo thresholds on either side of the true threshold prevents the bandwidth of analysis around each placebo threshold be sharply truncated (on one side) by the true threshold at zero. For each combination of placebo threshold (60 of them), polynomial degree one and two, and the two outcome variables, we conducted the local analysis to estimate the intent-to-treat (ITT) effect at each of the placebo thresholds. We included control variables in all specifications. In the following, we present results from the placebo analysis corresponding to the (unconditional) expenditure and purchase likelihood during qualification stage of the gift card promotion.

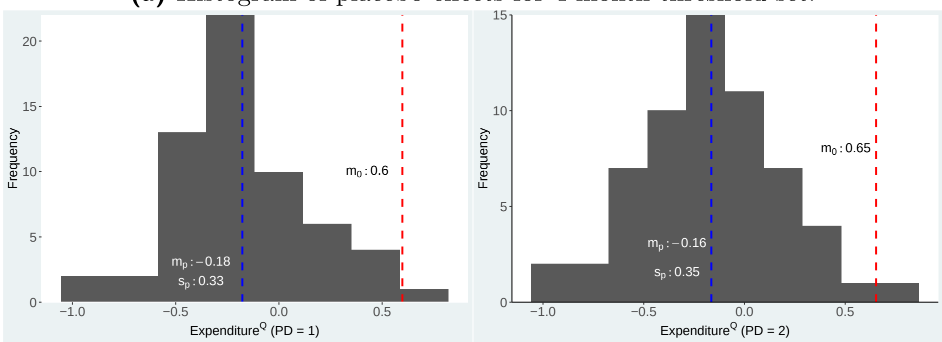
In Figures EC.6a and EC.6b, we present the histograms of the placebo effects corresponding to qualification stage expenditure for the 4-month and 13-month threshold sets, respectively. In Figures EC.7a and EC.7b, we present the histograms of the placebo effects corresponding to qualification stage purchase likelihood for the 4-month and 13-month threshold sets, respectively. The mean and standard deviation of the placebo distribution—denoted by $\mathrm{m_p}$ and $\mathrm{s_p}$, respectively—are provided in the figures. The mean of the placebo distribution is also indicated by the blue dotted line in the figures. The effect at the true threshold (i.e., at threshold 0)—denoted by $\mathrm{m_0}$—is indicated by the red dotted line in the figures. It is clear from these figures that the effect (corresponding to qualification stage expenditure and purchase likelihood) at the true threshold is substantially greater (more than one standard deviation) than the mean of the estimates obtained from placebo thresholds. These results, therefore, suggest that the effects observed at the true threshold for 4-month and 13-month threshold sets are not spurious.

EC.9. Compliance in Fuzzy RD Design

In a fuzzy RD setting, customers may be classified into four categories based on how their treatment status changes depending on which side of the threshold they belong to (Angrist et al. 1996): *Compliers, always-takers, never-takers,* and *defiers.* Compliers are customers who receive the gift card promotion email (treatment) if and only if they belong to the targeted side of the threshold. Always-takers (resp., never-takers) receive (resp., do not receive) the gift card promotion email regardless of which side of the threshold they belong to. Defiers receive the gift card promotion email if and only if they belong to the control side of the threshold. It is reasonable to assume that the retailer's targeting probability of gift card promotion email changes monotonically in customers' purchase recency (at least around the threshold value). Therefore, we assume that, changing the purchase recency threshold value should not increase the probability of receipt of gift card promotion email for some customers while at the same time reduce the probability of receipt of gift card promotion email for others. Under this monotonicity assumption, we note that defiers do not exist. Therefore, in our RD designs, there are compliers, never-takers, and always-takers, on both sides of the threshold. We refer to always-takers and never-takers collectively as *noncompliers.* In our fuzzy RD design, noncompliers exist primarily because of the retailer's targeting policy which is based on both observable and unobservable (to us) factors. For example, customers who unsubscribe from the retailer's marketing program



(a) Histogram of placebo effects for 4-month threshold set.



(b) Histogram of placebo effects for 13-month threshold set.

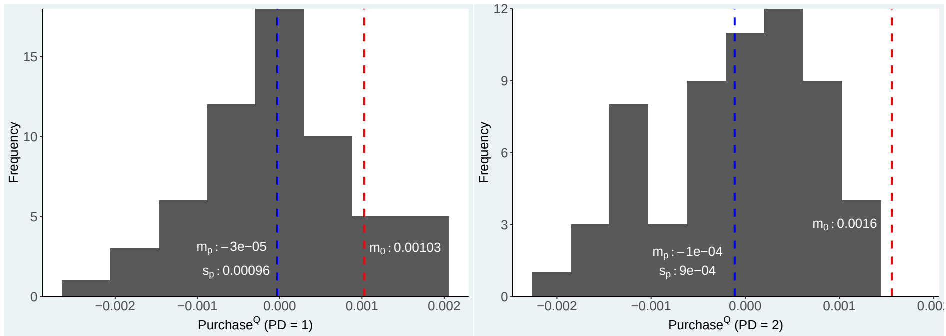
Figure EC.6 **Histogram of placebo effects corresponding to qualification stage expenditure for 4-month and 13-month threshold sets. On the left (right) panel of each figure is the histogram with polynomial degree one (two). The mean of the placebo distribution (m_p) is indicated by blue dotted line and the effect at the true threshold (m_0) is indicated by red dotted line. Control variables are included in both specifications.**

(fewer than 0.50% of observations) are never-takers. In this case, the customer's subscription status is unobservable to us, but is part of the retailer's targeting policy because retailers are legally mandated to honor customers' unsubscription requests. In addition, noncompliance may also result from customer actions. For example, some customers may seek out gift card promotion information even if they are not targeted by the retailer. These customers are classified as always-takers because they receive the treatment regardless of which of the threshold they belong to.

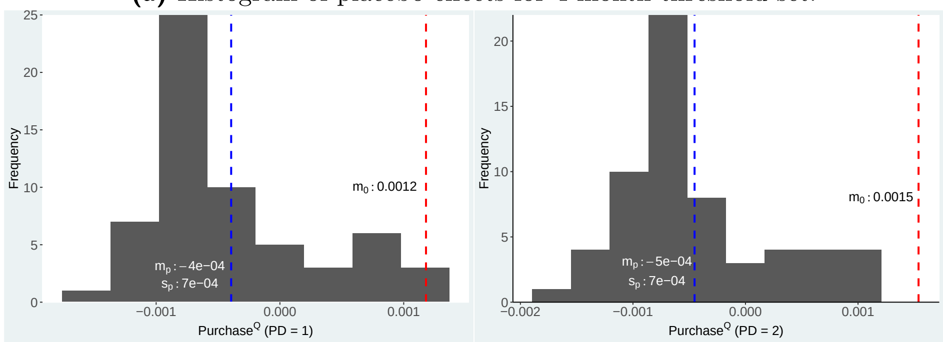
Estimation of LATE in fuzzy RD designs accounts for noncompliers as follows. As discussed in Section 4.2, the denominator in Equation (2) accounts for the non-compliance in the RD design and attributes the LATE to compliers. The presence of non-compliers, however, limits external validity of the LATE estimates, i.e., the segment of population over which the LATE can be applied beyond complier-customers. In Section 4.3.1, we profile compliers to estimate the size of the complier population at each threshold using the approach proposed by Marbach and Hangartner (2020).

EC.10. Scatter Plots of Average Customer Expenditure

In Figure EC.8, we provide scatter plots of average expenditure (aggregated by running variable) against the running variable for all thresholds. The fitted lines in the scatter plots are second-degree polynomial



(a) Histogram of placebo effects for 4-month threshold set.



(b) Histogram of placebo effects for 13-month threshold set.

Figure EC.7 Histogram of placebo effects corresponding to qualification stage purchase likelihood for 4-month and 13-month threshold sets. On the left (right) panel of each figure is the histogram with polynomial degree one (two). The mean of the placebo distribution (m_p) is indicated by blue dotted line and the effect at the true threshold (m_0) is indicated by red dotted line. Control variables are included in both specifications.

functions fitted based on the ordinary least squares criterion. We observe that a lower degree polynomial function captures the non-linearity of customer expenditure reasonably well around the threshold.

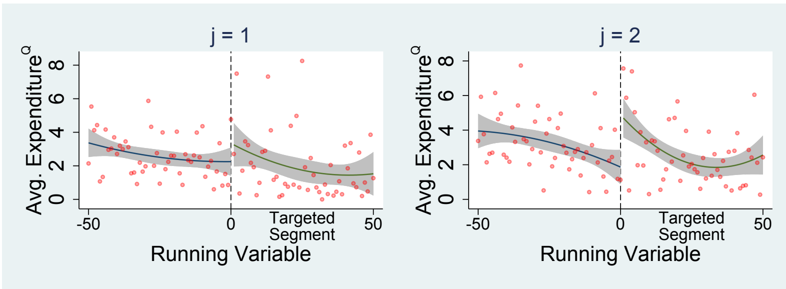
EC.11. Optimal Bandwidth and Summary Statistics

In Table EC.5, we present the magnitude of the optimal bandwidth, number of observations within the bandwidth (i.e., size of the estimation sample), and averages of the Expenditure^Q and Purchase^Q variables in the estimation sample for each of the seven thresholds.

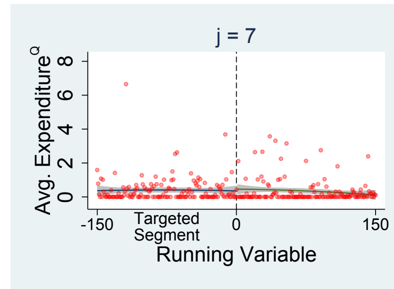
Index (j)		1	2	3	4	5	6	7
Threshold		106	120	376	381	387	388	481
	PD							
bw	1	22	18	16	74	71	38	37
(days)	2	39	32	14	100	82	75	100
Nbr. Obs.	1	114,614	53,202	25,508	127,516	93,357	85,328	57,837
	2	184,489	101,230	21,920	176,238	110,388	178,670	157,757
average	1	2.19	2.62	0.73	1.29	1.26	1.54	0.37
Expenditure^Q	2	2.23	2.79	0.68	1.36	1.31	1.58	0.40
average	1	0.40	0.53	0.19	0.30	0.32	0.53	0.07
Purchase^Q (%)	2	0.41	0.53	0.19	0.31	0.32	0.49	0.08

Table EC.5 Optimal bandwidth, sample size, and average Expenditure^Q and Purchase^Q in the estimation sample.

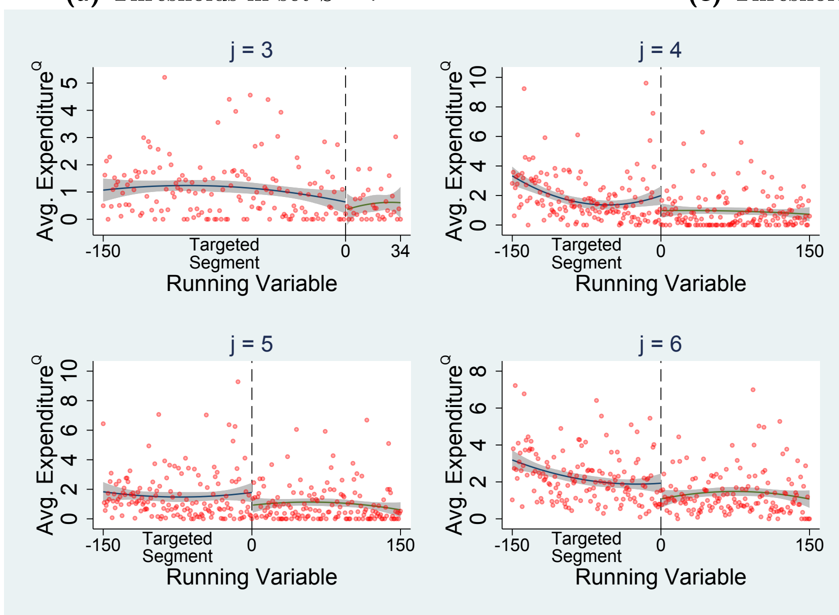
Notes: PD refers to degree of the polynomial functions (f, g) used in the RD estimation.



(a) Thresholds in set S^{4mo}.



(c) Threshold in set S^{16mo}.



(b) Thresholds in set S^{13mo}.

Figure EC.8 Scatter plot of average Expenditure^Q **against the running variable.**

Notes: The lines on either sides of the threshold are second degree polynomial functions (of running variable) fitted to the observed customer expenditure using ordinary least squares (OLS) method. The shaded region represents the 90% confidence interval of the predicted values.

EC.12. Robustness Checks

Here we establish robustness of LATE estimates to the choices of tuning parameters in the local estimation approach. The inputs to the non-parametric estimation of LATE include polynomial degree, kernel function, and bandwidth of analysis (see Section 4.2). Next, we discuss the robustness of our results with respect each of these input choices.

EC.12.1. Non-Parametric Estimation Using Triangular Kernel Function

One of the inputs to the non-parametric estimation of LATE is the kernel function (see Section 4.2.3), which is used to weigh the observations within the bandwidth around the threshold. The results presented in Section 4.3 use uniform kernel function in the LATE estimation. Here, we present LATE estimates with triangular kernel function using the `rdrobust` package in STATA. For each threshold set (4mo, 13mo, 16mo), we estimate LATE by varying the polynomial degree (one, two), with/without controls, and two choices of bandwidth (`mserd, msetwo`) resulting in eight combinations of the input parameters. Both `mserd` and `msetwo` are MSE-optimal bandwidth choices, but `msetwo` uses asymmetric bandwidths on either side of the threshold. We present results with bandwidth computed using `mserd`. The results with bandwidth computed using `msetwo` are similar to those with `mserd`, and hence omitted. For each of these eight combinations of polynomial degree-bandwidth choice-controls, three LATE estimates denoted by $\text{LATE}_{\texttt{cnv}}$, $\text{LATE}_{\texttt{bias\_crr}}$, and $\text{LATE}_{\texttt{rb}}$, along with the MSE-optimal bandwidths (bw), are presented in Table EC.6. $\text{LATE}_{\texttt{cnv}}$ is the conventional LATE estimator based on the discussion above. $\text{LATE}_{\texttt{bias\_crr}}$ debiases $\text{LATE}_{\texttt{cnv}}$ estimate when using the MSE-optimal bandwidth. $\text{LATE}_{\texttt{rb}}$ has the same point estimate as $\text{LATE}_{\texttt{bias\_crr}}$ but uses a theoretically justified approach to compute the variance of the estimator. We find that the LATE estimates obtained using the triangular kernel are consistent in terms of the mean and the standard errors compared to the LATE estimates obtained with uniform kernel in Table 2. Thus, our findings are robust to the choice of kernel function.

Threshold set	PD	Controls	bw^*	$\text{LATE}_{\texttt{cnv}}$	$\text{LATE}_{\texttt{bias\_crr}}$	$\text{LATE}_{\texttt{rb}}$	*Nbr.* Obs.
4mo	1	N	23	8.84^{***}	9.77^{***}	9.77^{***}	190,531
				(3.30)	(3.30)	(3.87)	
4mo	1	Y	23	8.84^{***}	9.77^{***}	9.77^{***}	190,531
				(3.30)	(3.30)	(3.87)	
4mo	2	N	55	8.72^{***}	9.51^{***}	9.51^{***}	440,185
				(3.22)	(3.22)	(3.62)	
4mo	2	Y	55	8.72^{***}	9.51^{***}	9.51^{***}	440,185
				(3.22)	(3.22)	(3.62)	
13mo	1	N	62	1.24^{***}	1.23^{***}	1.23^{**}	403,474
				(0.46)	(0.46)	(0.54)	
13mo	1	Y	62	1.24^{***}	1.23^{***}	1.23^{**}	403,474
				(0.46)	(0.46)	(0.54)	
13mo	2	N	82	1.08^{*}	1.00^{*}	1.00	533,815
				(0.59)	(0.59)	(0.65)	
13mo	2	Y	82	1.08^{*}	1.00^{*}	1.00	533,815
				(0.59)	(0.59)	(0.65)	
16mo	1	N	64	0.68	0.85	0.85	100,289
				(0.63)	(0.63)	(0.73)	
16mo	1	Y	64	0.68	0.85	0.85	100,289
				(0.63)	(0.63)	(0.73)	
16mo	2	N	101	0.70	0.85	0.85	159,483
				(0.76)	(0.76)	(0.83)	
16mo	2	Y	101	0.70	0.85	0.85	159,483
				(0.76)	(0.76)	(0.83)	

Table EC.6 LATE estimates using triangular kernel function in RD estimation.

Notes: Cluster-robust standard errors are in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

EC.12.2. Robustness to Bandwidth Selection

To check the robustness of the LATE estimates to bandwidth size, we re-estimate the total effect of gift card promotion on customer expenditure (Section 4.3, Table 2) using bandwidth values that are scaled multiples of the MSE-optimal bandwidth proposed by Calonico et al. (2014). Specifically, we consider bandwidth sizes that are 0.8, 0.9, 1.1, and 1.2 times the MSE-optimal bandwidth. We present the LATE estimates obtained from using these alternative bandwidths in Table EC.7. These results suggest that our LATE estimates are robust to the size of bandwidth.

Threshold set		4-month		13-month		16-month	
$0.8 \times bw^*$	PD = 1	8.99**	6.23*	1.11*	1.00	1.30	1.30
		(3.71)	(3.68)	(0.67)	(0.68)	(0.82)	(0.81)
	PD = 2	6.79**	5.33**	1.29**	1.43***	-0.13	-0.09
		(2.68)	(2.71)	(0.51)	(0.52)	(0.49)	(0.49)
$0.9 \times bw^*$	PD = 1	9.92***	7.47**	1.39**	1.32**	1.03	1.03
		(3.52)	(3.50)	(0.61)	(0.63)	(0.72)	(0.71)
	PD = 2	6.48***	5.23**	1.13**	1.23**	-0.16	-0.09
		(2.47)	(2.49)	(0.49)	(0.49)	(0.46)	(0.46)
$1.1 \times bw^*$	PD = 1	7.39**	5.39*	1.22**	1.22**	0.81	0.85
		(3.12)	(3.13)	(0.56)	(0.57)	(0.71)	(0.71)
	PD = 2	6.42***	5.26**	1.22***	1.27***	-0.07	-0.02
		(2.22)	(2.23)	(0.44)	(0.44)	(0.42)	(0.42)
$1.2 \times bw^*$	PD = 1	6.51**	4.49	1.39**	1.46***	0.52	0.54
		(2.96)	(3.02)	(0.54)	(0.56)	(0.71)	(0.71)
	PD = 2	6.05***	4.82**	1.06**	1.04**	-0.27	-0.20
		(2.10)	(2.10)	(0.42)	(0.43)	(0.43)	(0.42)
Controls		N	Y	N	Y	N	Y

Table EC.7 Robustness of LATE^k estimations for each threshold set and polynomial degree (PD).

Notes: Standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The first column indicates the size of the bandwidth used in the RD estimation. bw^* denotes the MSE-optimal bandwidth proposed by Calonico et al. (2014).

EC.13. Effect of Gift Card Promotion's Tiered Expenditure Structure

Here, we elaborate on the nested analysis used to tease out the tier-specific effect of participation on expenditure (Section 4.3.4). Note that there are six expenditure levels used by the retailer: \$500, \$750, \$1000, \$2500, \$5000 and \$10000. First, we set expenditures (strictly) less than the \$500 expenditure level to zero, and estimate the total effect, the advertisement effect, and the participation effect (as in Sections 4.3.1 and 4.3.2), denoted by $\mathsf{LATE}(500)$, $\mathsf{LATE}_{\mathsf{adv}}(500)$, and $\Delta(500)$, respectively, such that $\Delta(500) = \mathsf{LATE}(500) - \mathsf{LATE}_{\mathsf{adv}}(500)$. $\Delta(500)$ quantifies the effect of participation on expenditures greater than \$500. Next, we set expenditures less than \$750 to zero and estimate $\mathsf{LATE}(750)$, $\mathsf{LATE}_{\mathsf{adv}}(750)$, and $\Delta(750)$. We repeat this analysis until the last expenditure level (\$10,000). Note that the estimates presented in Tables 2 and 5 correspond to $\mathsf{LATE}(0)$ and $\mathsf{LATE}_{\mathsf{adv}}(0)$, respectively. We obtain the tier-specific effect, i.e, the (net) effect of participation on expenditures within a tier $[\underline{\ell}, \bar{\ell})$ as $\delta(\underline{\ell}) := \Delta(\underline{\ell}) - \Delta(\bar{\ell})$. For example, $\delta(500) = \Delta(500) - \Delta(750)$ is the effect of participation on expenditures falling into the tier $[500, 750)$. The effect of participation on expenditures in tier $[0, 500)$ is given by $\delta(0) = \Delta(0) - \Delta(500)$, and as discussed in Section 4.3.4, we expect $\delta(0) \le 0$. We present the estimates of $\mathsf{LATE}(\cdot)$, $\mathsf{LATE}_{\mathsf{adv}}(\cdot)$, and $\Delta(\cdot)$ obtained from the nested analysis (with controls included in specification) in Tables EC.8 and EC.9 for the 4-month and 13-month threshold sets, respectively. Estimates without controls in the specification are qualitatively similar to those with controls included in the specification, and are available from the authors upon request.

EC.14. Details of the Propensity Score Matching Analyses

Here we first describe the details of how GC_value variable was derived and then describe the details of the propensity score matching (PSM) analysis, whose results are presented in Section 5. We derive the GC_value variable as follows. In our data, 105 redeemers (out of the 425 redeemers in total) qualified for multiple gift cards with redemption stages overlapping with the focal redemption stage. Similarly, 46 non-redeemers (out of the 378 non-redeemers in total) qualified for multiple gift cards with redemption stages overlapping with

Expenditure^{Q}	PD = 1			PD = 2		
	$\text{LATE}^{4\text{mo}}(\cdot)$	$\text{LATE}^{4\text{mo}}_{\text{adv}}(\cdot)$	Participation Effect (Δ)	$\text{LATE}^{4\text{mo}}(\cdot)$	$\text{LATE}^{4\text{mo}}_{\text{adv}}(\cdot)$	Participation Effect (Δ)
≥ 0	5.99^{*}	3.34	2.65	5.64^{**}	4.42^{**}	1.22
	(3.21)	(2.63)		(2.32)	(1.91)	
≥ 500	5.30^{*}	2.65	2.65	4.83^{**}	3.61^{*}	1.22
	(3.17)	(2.57)		(2.28)	(1.86)	
≥ 1000	5.74^{*}	3.06	2.68	4.59^{**}	3.26^{*}	1.33
	(3.06)	(2.44)		(2.19)	(1.76)	
≥ 2500	4.16	1.59	2.57	3.85^{**}	2.06	1.79
	(2.67)	(1.95)		(1.93)	(1.44)	
≥ 5000	4.28^{*}	1.71	2.57	3.03^{*}	1.23	1.80
	(2.37)	(1.52)		(1.65)	(1.04)	
≥ 10000	1.50	1.50	0.00	0.97	0.97	0.00
	(1.50)	(1.50)		(0.99)	(0.99)	

Table EC.8 Tier-specific participation effect within 4-month threshold set (with controls).

Notes: Cluster-robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Expenditure^{Q}	PD = 1			PD = 2		
	$\text{LATE}^{13\text{mo}}(\cdot)$	$\text{LATE}^{13\text{mo}}_{\text{adv}}(\cdot)$	Participation Effect (Δ)	$\text{LATE}^{13\text{mo}}(\cdot)$	$\text{LATE}^{13\text{mo}}_{\text{adv}}(\cdot)$	Participation Effect (Δ)
≥ 0	1.08^{*}	0.91	0.17	1.37^{***}	1.31^{***}	0.06
	(0.59)	(0.58)		(0.47)	(0.46)	
≥ 500	0.85	0.67	0.18	0.85^{**}	0.79^{*}	0.06
	(0.52)	(0.51)		(0.42)	(0.41)	
≥ 750	0.92^{*}	0.77	0.15	0.8^{**}	0.74^{*}	0.06
	(0.49)	(0.48)		(0.40)	(0.39)	
≥ 1000	0.7	0.56	0.14	0.58	0.5	0.08
	(0.44)	(0.43)		(0.36)	(0.35)	
≥ 2500	0.18	0.18	0.00	0.05	0.05	0.00
	(0.19)	(0.19)		(0.15)	(0.15)	
≥ 5000	-0.03	-0.03	0.00	-0.06	-0.06	0.00
	(0.03)	(0.03)		(0.06)	(0.06)	
≥ 10000	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)		(0.00)	(0.00)	

Table EC.9 Tier-specific participation effect within 13-month threshold set (with controls).

Notes: Cluster-robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

the focal redemption stage. These customers did not redeem any of their gift cards in the focal redemption stage. Our datasets do not indicate the specific promotion in which the redeemed gift card was earned. We conservatively assume that whenever such multiple gift cards' redemption stages are overlapping with the focal redemption stage by at least one month, the redeemer redeems all their earned gift cards in the focal redemption stage. Therefore, we derived GC_value variable by combining the face values of such multiple gift cards whose redemption stages are overlapping with the focal redemption stage by at least one month.

The PSM method is a statistical approach used to estimate the effect of a treatment on outcome variable using observational data. This method accounts for the bias due to self-selection, based on observable factors. In the PSM analysis of each of the three promotions, for selecting customers who receive a gift card promotion email *similar* to those who do not receive this email, we first compute the propensity to receive a gift card promotion email by running a logistic regression. We used Purchase_rec, Purchase_freq, Avg_exp, Web_freq, and Promo_email_freq variables as the explanatory variables of the logistic regression. We derived these explanatory variables starting from the beginning of the focal promotions' qualification stages and going back for 13 months. For every customer who receives a gift card promotion email (i.e., treated customers), at least one customer with similar predicted propensity but who do not receive a gift card promotion email (i.e., control customers) is matched (with replacement). The difference in the matched (control and treatment) customers' (average) net total redemption stage expenditures (i.e., $\text{NetExpenditure}^{R}$) is averaged to compute the effect of receiving a gift card promotion email on the net total redemption stage expenditure (i.e., *average treatment effect on the treated* or ATET; note that estimating the ATET only requires finding matches for the treated customers as opposed to finding matches for both the treated and control customers). To ensure that similar customers are matched, we restrict the difference between the propensity scores of matched customers to be less than 0.01. This restriction did not exclude any customer for whom a suitable match cannot be found. We used default (i.e., independent and identically distributed) Abadie–Imbens standard errors.

EC.15. Supporting Tables

Threshold set	4-month
Number of promotions in threshold set	2
Total RegularPriceExpenditure within the bandwidth	$460,852
Number of customers who received promotion email on the treated side	81,577
Fraction of compliers (based on Marbach and Hangartner 2020)	26.70%
Estimated LATE on RegularPriceExpenditure (from Table 6)	4.28**
Percentage of incremental RegularPriceExpenditure due to gift card promotion	25.36%

Table EC.10 Effect of the gift card promotion email on RegularPriceExpenditure.

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. We consider the bandwidths corresponding to specification with PD-two and controls included. Calculations of the values in the last row are similar to the ones in Table 4.

Category	High-end Brands	Mid-range Brands	Low-end Brands
Dresses (*Best-selling*)	Adam Lippes, Akris	10 Crosby Derek Lam, 12th Street By Cynthia Vincent	4.Collective, 525 America
Jackets (*Mid-selling*)	Akris, Alexander Mcqueen	A.l.c., Akris Punto	7 For All Mankind, Addison
Scarves (*Niche*)	Donna Karan, Giorgio Armani	Alexander Mcqueen, Armani Collezioni	Anna Coroneo, Diane Von Furstenberg

Table EC.11 Examples of brand names belonging to different brand type-category classes.