Supplementary Online Appendix

Estimating Treatment Effects with Big Data When Take-up is Low: An Application to Financial Education

Gabriel Lara Ibarra, David McKenzie, and Claudia Ruiz-Ortega

Supplementary Online Appendix:

S1. Details of Simulations with Heterogeneous Treatment Effects Correlated with Take-Up Rates

To illustrate how statistical power changes with the take-up rate once the possibility of heterogeneous treatment effects is allowed for, a series of simulations are carried out. These are based on the case of the outcome "paying more than the minimum payment" and our coaching intervention. The study simulates a sample size of 5,000 units, 2,500 assigned to the treatment group, and 2,500 to the control group, and consider an outcome that has control mean 0.69 and standard deviation 0.46 (as is the case for paying more than the minimum in the coaching sample).

Let the data generating process be:

$$Y_i = 0.69 + \gamma_i s_i T_i + \varepsilon_i$$

Where the study draws ε_i from an i.i.d. normal distribution with mean 0 and standard deviation 0.46; T_i is 1 for the treatment group and 0 for the control group (none of whom receive the intervention). Let θ_i be individual i's underlying propensity to take up the intervention. The study then allows for this to be correlated with i's treatment effect by drawing γ_i and θ_i from a joint normal distribution:

$$\begin{pmatrix} \gamma_i \\ \theta_i \end{pmatrix} i.i.d.N\left(\begin{pmatrix} \mu_\gamma \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\gamma & \rho \\ \rho & 1 \end{pmatrix}\right)$$

Then let $\varphi_i = rank(-\theta_i)$, and consider a given take-up rate π_T. The study then sets the receipt of treatment (s_i) as follows:

$$s_i = 1 \quad \text{if} \quad T_i = 1 \quad \text{and} \quad \varphi_i < int\,(\pi_T * 2500)$$

And otherwise $s_i = 0$.

The study fixes $\mu_\gamma = 0.06$, which is the treatment effect found for those receiving coaching using the matched difference-in-differences approach. The study then estimates:

$$Y_i = a + bT_i + u_i$$

To get the ITT estimate $\hat{b}_{ITT}$, and then carry out the two-stage squares regression:

$$Y_i = a + bs_i + u_i$$

$$s_i = c + dT_i + v_i$$

Where s_i is instrumented with the random assignment to treatment (T_i) to get the IV estimate $\hat{b}_{IV}$.

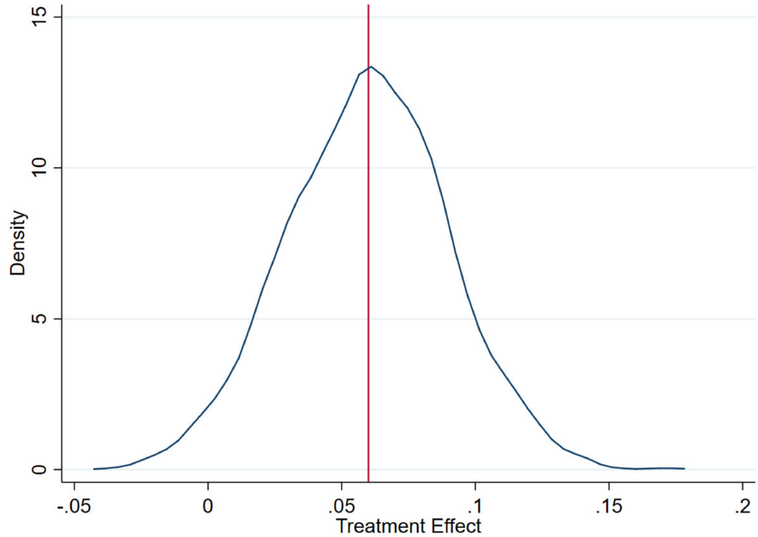
This process is replicated 1,000 times, and in each replication the study collects the estimated ITT and IV parameters, and tests the hypotheses $b_{ITT} = 0$ and $b_{IV} = 0$. Statistical power is then the percentage of replications in which the study rejects (at $\alpha = 0.05$) this null hypothesis, given the specified effect size distribution of $\gamma_i s_i$. The results on three types of simulations are presented, which are described below.

Simulation type 1. Holding σ_γ constant, how does power vary when we vary ρ?

The first set of simulations fixes $\sigma_\gamma = 0.03$, which is half the size of mean effect in this population. Figure S1.1 shows this distribution of individual treatment effects, in which a small number of individuals actually have negative effects (become less likely to pay the minimum payment after receiving coaching), and some have effects that could be more than double the average.

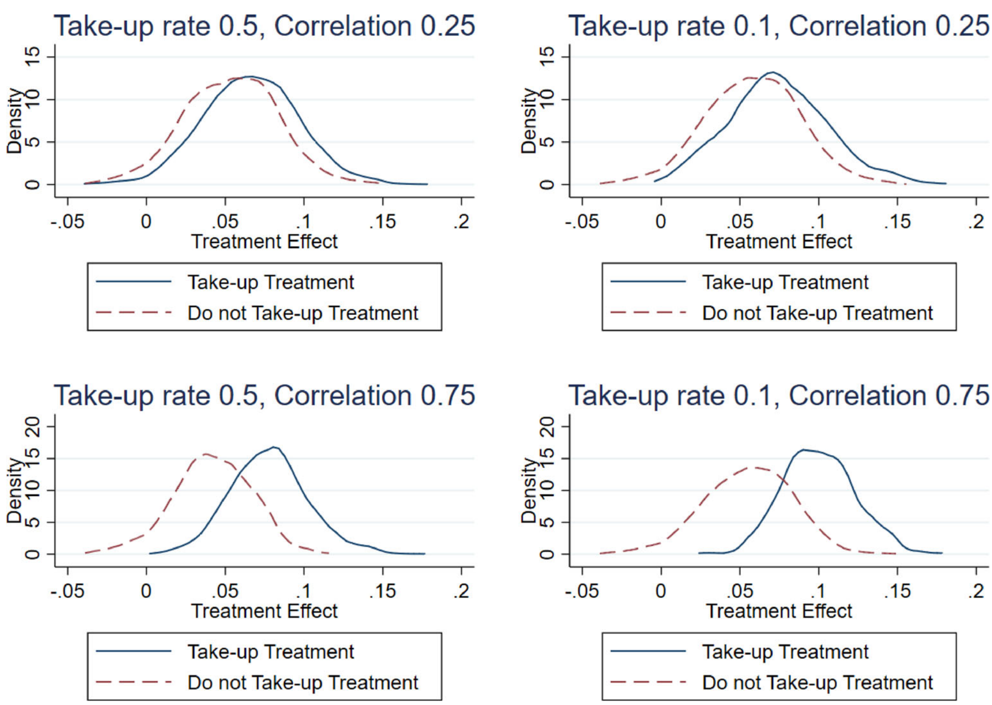
The study then explores how the treatment effects vary as ρ changes. When ρ is zero, those who take up the coaching intervention would be randomly drawn from this distribution, and so regardless of the take-up rate, the IV treatment effect for those who do take up coaching will have expected value equal to the mean of 0.06. Then as ρ becomes positive, individuals with larger treatment effects become more likely to take-up treatment. This is illustrated in fig. S1.2. When the correlation is positive, but modest, as

Figure S1.1. Distribution of Simulated Treatment Effects for $\sigma_\gamma = 0.03$



Source: Illustration from simulated treatment effect where control mean is 0.69, control standard deviation is 0.46, sample size is 5,000 units divided equally between treatment and control, and individual treatment effects are drawn from a random normal distribution with mean 0.06 and standard deviation equal to 0.03.
Note: Vertical red line indicates mean of 0.06.

Figure S1.2. Varying ρ Changes the Selectivity of Who Takes Up Treatment



Source: Illustration from simulated treatment effect where control mean is 0.69, control standard deviation is 0.46, sample size is 5,000 units divided equally between treatment and control, and individual treatment effects are drawn from a random normal distribution with mean 0.06 and standard deviation equal to 0.03.
Note: Figures show the distribution of the treatment effect as the correlation between the order in which units take up treatment and their treatment effects moves from $\rho = 0.25$ (top figures) to $\rho = 0.75$ (bottom figures).

in the top row, there is considerable overlap in the distributions of treatment effects for those who take up treatment and those who do not. The bottom row shows the distributions when the correlation is high. When the correlation becomes a lot stronger, the two distributions diverge a lot more, and as the take-up rate falls, those who take up treatment almost all have larger than average treatment effects.

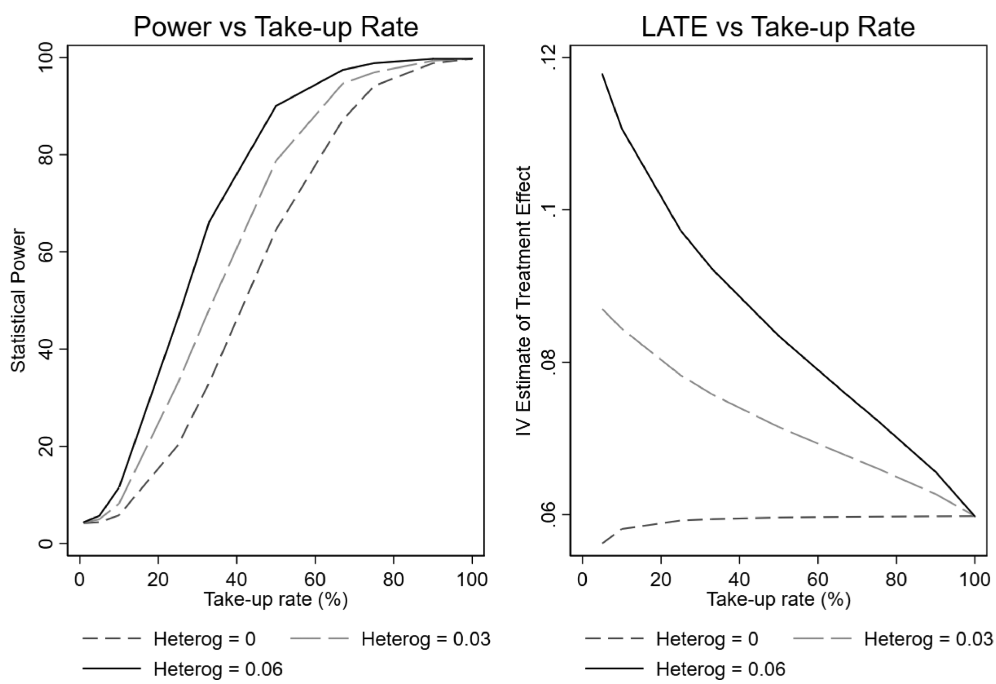
In the extreme case of $\rho = 1$, take-up perfectly sorts on the treatment effect, so that the $\pi_T * 2500$ individuals with the largest treatment effects take up treatment, and the remainder do not.

These simulations holding σ_γ at 0.03 and varying ρ are used to calculate the power and IV estimates presented in fig. S1.2.

Simulation type 2. Holding $\boldsymbol{\rho}$ constant, how does power vary when $\boldsymbol{\sigma_\gamma}$ is varied?

From equation (8), it is seen that σ_γ functions exactly like ρ in increasing the expected value of the ITT, and thereby increasing power as it increases. The study considers an intermediate case of $\rho = 0.5$, and in fig. S1.3 plots the resulting power and estimated LATE. It is seen that more heterogeneity across individuals' expected treatment effects increases power substantially when take-up rates are not near the tails. In contrast, when take-up rates get to 5 or 1 percent, more heterogeneity does not have much of an impact on power, even though it does result in much larger IV estimates for the few units that are treated.

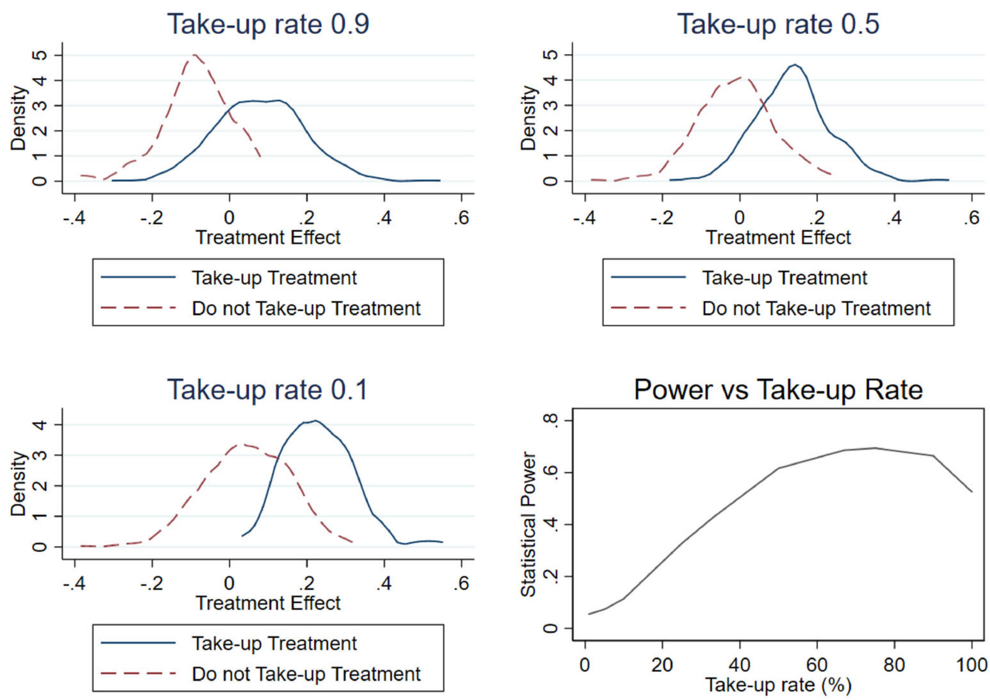
Figure S1.3. With Take-up Positively Correlated with Treatment Effects, Power Falls Less Steeply with Take-up the Larger the Heterogeneity in Individual Treatment Effects



Source: Illustration from simulated treatment effect where control mean is 0.69, control standard deviation is 0.46, sample size is 5,000 units divided equally between treatment and control, and individual treatment effects are drawn from a random normal distribution with mean 0.06 and standard deviation equal to 0, 0.03 or 0.06, and a correlation between the order in which units take up treatment and their treatment effect of 0.5.

Note: Figures plot the statistical power and estimated LATE as heterogeneity and take-up rates increase.

Figure S1.4. In the Extreme Case, Moving from 100 Percent Take-up to Lower Take-up Rates Can Increase Power, so Long as Take-up is Not Too Low



Source: Illustration from simulated treatment effect where control mean is 0.69, control standard deviation is 0.46, sample size is 1,000 units divided equally between treatment and control, and individual treatment effects are drawn from a random normal distribution with mean 0.06 and standard deviation equal to 0.12, and a correlation between the order in which units take up treatment and their treatment effect of 0.75.

Note: Top two figures and bottom left show kernel densities of the distribution of treatment effects for those who take up and do not take up treatment as the take-up rate is varied. Bottom right figure shows the resulting statistical power.

Simulation type 3. A case where lower take-up can increase power

If power is not as high to begin with, and there are many units with large negative treatment effects, it is possible that a drop in the take-up rate might actually increase power. This arises because individuals for whom the treatment would have a negative effect do not take the treatment up, and the replacement of these negative treatment effects with zero effect therefore increases the ITT. To illustrate this, the study chooses a relatively high correlation ($\rho = 0.75$), reduces the sample size to 500 treated and 500 control, and increases the heterogeneity to 0.12 (twice the mean). Figure S1.4 then shows that at a take-up rate of 90 percent, those who do not take up treatment are heavily drawn from those with large negative treatment effects. As a result, moving from 100 percent take-up to 75 percent take-up in this setting increases power from 52.5 percent to 69.4 percent. But even in this extreme case, at very low take-up rates, power is still low as many of the individuals who do not take up treatment would still have had positive treatment effects.

S2. Robustness Checks and Additional Statistics

Table S2.1. Contents of Credit Card Financial Literacy Course

Session 1. Credit cards (duration: 1 hour)

This session explains that debt can be useful if you know how to use it correctly. The session also covers how to apply for a loan and the different types of loans there are.

Objectives:

1. Participants know what debt is
2. Participants know what a credit card is
3. Participants know good habits with their credit cards

Content:

- Types of loans available
- Advantages and disadvantages of each type of loan
- What a credit card is
- Credit cards' elements
- The personal identification number (PIN)
- The bank statement and how to read it
- Credit cards APR (annual percentage rate)
- Healthy use of credit cards

Exercises:

- Case study: Identifying what kind of credit is best for each situation
- Interactive exercise: Identifying the parts of a bank statement

Session 2. Healthy credit rating (duration: 1 hour)

This session focuses on understanding what a credit rating is and how to keep a good credit score.

Objectives:

1. Participants learn what a credit rating is
2. Participants learn how to obtain their own credit rating
3. Participants learn what they can do if they have credit problems

Content:

- Credit ratings and the importance of having a good credit rating
- The credit bureau
- How and where to get a credit report
- What my credit report means
- Self-diagnose credit health
- Advise to maintain or improve your credit rating

Exercise:

- Case study: Helping someone to solve their financial problems

Source: BBVA Bancomer Adelante con tu futuro ©.

Note: The table describes the topics, objectives, content, duration, and exercises covered throughout the workshops.

Table S2.2. Golden Rules from BBVA Bancomer's Financial Literacy Course

15 GOLDEN RULES

Know them to keep or improve your financial health.
Consider them when using your credit card. Having them in mind will help take maximum advantage of your card:

1. Write down everything that you **earn** and **spend**. This will help you plan all your purchases and save.
2. When recording **expenses**, divide them into fixed (such as electricity) and variable (such as the coffee that you regularly buy)
3. Purchases with your credit card are a loan that you must pay back. Consider them as part of your **expenses** and **avoid delays**.
4. Know and keep in mind the **fees** and **interest rates of your credit cards.** This will help you manage it better and know its total cost.
5. Always consider your **closing and due dates**, to know when to make purchases and avoid delays.
6. Remember that when you buy with your credit card the **right after** the **closing date, you have up to 50 days to pay with no interests.**
7. Failing to cover at least the **minimum payment** will put you at risk of not being able to use your credit card and **paying additional fees**.
8. Failing to cover at least the **minimum payment** by the **due date** will result in **late fees and interests**.
9. Remember that credit cards purchases at **'months without interest' are not considered in your minimum payment**.
10. When you pay **more than the minimum**, your debt will decline and you will pay sooner. **The more you pay, the more your debt drops**.
11. Know and use available means of payment (such as automatic payments) to **never delay in the payment of your credit card**.
12. Only use credit cards that your **income and expenditures** level can allow you to pay and manage.
13. Avoid loan payments **higher than 30% your monthly income**; this will help take care of your credit health.
14. If you have an **extra income** throughout the month, try to use part of it to **pay** or **reduce** the debt in your credit card.
15. You are not alone. If you have problems with your credits, you can approach your **account executive or seek for help in your bank.**

Source: BBVA Bancomer Adelante con tu futuro ©.
Note: The figure shows a translated flyer of the 15 "golden rules" received by participants to the workshops.

Table S2.3. Description of Coaching Sessions

Sessions	Objectives	Topics to be covered	Tools/materials used
Session 1: Diagnostic	Identify financial concerns based on credit situation	Introduce coaching sessions Invite client to participate in the sessions Mention the 5-session program Identify client's issues/concerns Introduce solutions to issues/concerns identified	- Client's information - Diagnostic questionnaire from the Credit Health workshop
Session 2: Budget	Prepare savings plan and discuss barriers that prevent client from keeping up with payments	Analyze client's expenditures Classify expenditures into fixed and variable Self-evaluation Request to prepare the coaching worksheet for the next session	- Coaching format for creating a budget
Session 3: Credit	Analyze key aspects of a credit card and credit card statement Develop action plan to improve credit card payment behavior	Review topics from previous session Explain main parts of the bank statement Highlight key elements to improve overall use of the credit card Request that coaching worksheet is ready for the next session	- Client's bank statements - Credit card
Session 4: Healthy credit	Suggest alternatives that help client pay outstanding loans	Use client's bank statement to check credit card behavior Evaluate if recommendations are improving client's credit health	- Coaching worksheet - Client information
Session 5: Goals met?	Measure effectiveness of coaching sessions	Discuss advantages of proper use of credit card and good credit history	- Bank statements - Client information

Source: BBVA Bancomer Adelante con tu futuro ©.
Note: The table summarizes the objectives and topics covered throughout the coaching intervention, as well as the tools and materials used by the coach to monitor progress.

Table S2.4. Current Credit Card Offerings by BBVA Bancomer Mexico

Name of card	Minimum payment (%)	Minimum income to be demonstrated	Late fee	Annual fee	Annual percentage rate (APR)
Platinum	20	50,000	377	2177	34.9%
Visa infinite	20	150,000	–	5275	18.6%
Oro	20	20,000	377	972	75.3%
Afinidad UNAM	20	12,000	377	972	88.0%
e	25	5,000	348	580	ND
Azul	20	6,000	377	631	90.2%
IPN	20	6,000	377	631	91.6%
Congelada	20	4,000	235	290	115.6%
Educacion	20	6,000	377	631	68.2%
Mi Primera Tarjeta	20	6,000	–	–	79.9%
Rayados	20	6,000	377	631	83.0%

Source: Comisión Nacional para la Protección y Defensa de los usuarios de Servicios Financieros (accessed October 13, 2017).
Note: APR does not include taxes.

Table S2.5. Comparison of Pre-Intervention Means for Workshop Intervention

	Treatment	Full control sample		In common support		NN all vars		NN lasso		NN outcome	
	Received workshop	Normalized difference	*p*-value	Normalized difference	*p*-value	Normalized difference	*p*-value	Normalized difference	*p*-value	Normalized difference	*p*-value
Female	0.511	−0.016	0.705	−0.015	0.753	−0.021	0.743	−0.150	0.013		
Age	45.995	−0.040	0.345	−0.065	0.166	−0.174	0.008	−0.074	0.218		
Years as client	15.189	−0.096	0.072	−0.104	0.074	−0.260	0.000	−0.173	0.004		
Mean min pay	0.916	−0.286	0.000	−0.147	0.004	0.038	0.559	−0.121	0.045	0.020	0.755
Mean delay in paying	0.003	0.153	0.002	0.003	0.950	0.020	0.763	0.175	0.004	−0.010	0.863
Mean log spending	6.732	−0.441	0.000	−0.346	0.000	−0.011	0.873	−0.108	0.075	−0.007	0.910
Mean deposit account	0.751	−0.145	0.001	−0.147	0.002	0.035	0.592	−0.202	0.001	−0.003	0.961
Mean profitable client	0.770	0.181	0.000	0.200	0.000	−0.032	0.626	0.252	0.000	0.000	1.000
Sample size	583	36946		26811		465		547		469	

Source: Authors' analysis based on the study implementation data provided by BBVA Bancomer.
Note: The control group varies with outcome for the last approach (NN outcome), and so normalized differences and *p*-values are shown using the outcome specific control group.

Table S2.6. Comparison of Pre-Intervention Means for Coaching Intervention

	Treatment	Full control sample		In common support		NN all vars		NN lasso		NN outcome	
	Received coaching	Normalized difference	*p*-value	Normalized difference	*p*-value	Normalized difference	*p*-value	Normalized difference	*p*-value	Normalized difference	*p*-value
Female	0.354	0.251	0.000	0.182	0.019	−0.076	0.461	0.320	0.001		
Age	45.618	0.030	0.661	0.039	0.618	0.155	0.131	0.066	0.492		
Years as client	14.557	−0.074	0.248	−0.028	0.709	0.035	0.733	0.039	0.682		
Mean min pay	0.737	−0.561	0.000	−0.346	0.000	0.032	0.753	−0.150	0.117	0.024	0.813
Mean delay in paying	0.011	0.280	0.001	0.018	0.811	0.045	0.661	0.094	0.325	−0.029	0.747
Mean log spending	5.283	−0.423	0.000	−0.316	0.000	−0.027	0.790	0.087	0.361	0.020	0.830
Mean deposit account	0.862	0.068	0.292	0.030	0.689	−0.047	0.645	0.041	0.665	0.002	0.981
Mean profitable client	0.846	0.086	0.190	0.093	0.203	0.047	0.648	0.087	0.360	0.000	1.000
Sample size	246	2504		1563		190		219		192	

Source: Authors' analysis based on the study implementation data provided by BBVA Bancomer.
Note: The control group varies with outcome for the last approach (NN outcome), and so normalized differences and *p*-values are shown using the outcome specific control group.

Table S2.7. Post-Treatment Mean Outcomes in the Control Group Sample by Risk Classification

	High risk	Medium risk	Low risk	All
Share of debt paid	0.39	0.56	0.55	0.53
	1.06	1.29	1.25	1.24
	57,944	131,748	111,295	300,987
Delay in payment (>1 day)	0.08	0.05	0.04	0.05
	0.27	0.23	0.19	0.22
	61,572	142,757	121,386	325,715
Payment above minimum required	0.73	0.81	0.84	0.81
	0.45	0.39	0.36	0.39
	57,818	131,448	111,040	300,306
Owns basic deposit account with bank	0.77	0.7	0.67	0.7
	0.42	0.46	0.47	0.46
	61,666	141,995	119,916	323,577
Profitability to the bank	0.81	0.76	0.8	0.79
	0.39	0.43	0.4	0.41
	48,808	112,497	96,130	257,435
Monthly balance in MXN pesos	9.68	9.1	9.28	9.29
	2.69	2.8	2.82	2.79
	60,477	139,003	117,769	317,249
Monthly credit card purchases in MXN pesos	4.82	5.31	5.72	5.4
	4.23	4.11	3.97	4.09
	61,572	142,757	121,386	325,715

Source: Authors' analysis based on the study implementation data provided by BBVA Bancomer.

Note: The table presents the mean, standard deviation, and number of observations of the sample of 36,946 clients assigned to the control group for the post-treatment period of June 2016 to February 2017. Each column presents a risk group category, which was assigned by BBVA Bancomer.

Table S2.8. Estimated Heterogeneous Treatment Effects for Those Who Received Workshops

			Nearest-neighbor matching		
	Full control sample	Common support	All vars	Lasso	Outcome
Panel A: Pay more than the minimum payment					
Receive Workshop*Post-Intervention	0.038***	0.047***	0.048***	0.044***	0.106***
	(0.010)	(0.011)	(0.014)	(0.014)	(0.016)
Receive Workshop*Post-Intervention*Med Risk	0.016	0.003	−0.030*	0.011	−0.006
	(0.017)	(0.019)	(0.018)	(0.018)	(0.019)
Receive Workshop*Post-Intervention*High Risk	0.036	0.019	0.034	0.029	0.017
	(0.026)	(0.027)	(0.028)	(0.027)	(0.027)
Panel B: Delay in payment					
Receive Workshop*Post-Intervention	−0.035***	−0.035***	−0.020***	−0.037***	−0.037***
	(0.004)	(0.004)	(0.007)	(0.008)	(0.008)
Receive Workshop*Post-Intervention*Med Risk	−0.006	−0.001	0.003	−0.005	0.002
	(0.006)	(0.007)	(0.007)	(0.007)	(0.006)
Receive Workshop*Post-Intervention*High Risk	−0.001	−0.004	−0.010	0.005	0.010
	(0.012)	(0.010)	(0.011)	(0.013)	(0.014)
Panel C: Log monthly spending on card					
Receive Workshop*Post-Intervention	0.333**	0.396**	0.556***	0.460***	0.568***
	(0.146)	(0.155)	(0.181)	(0.166)	(0.164)
Receive Workshop*Post-Intervention*Med Risk	0.172	−0.090	−0.467*	−0.053	0.204
	(0.249)	(0.272)	(0.260)	(0.247)	(0.239)
Receive Workshop*Post-Intervention*High Risk	0.353	0.262	0.196	0.083	−0.051
	(0.361)	(0.388)	(0.389)	(0.366)	(0.347)
Panel D: Has deposit account					
Receive Workshop*Post-Intervention	0.039**	0.039**	0.042*	0.058**	0.018
	(0.020)	(0.020)	(0.022)	(0.023)	(0.022)
Receive Workshop*Post-Intervention*Med Risk	−0.048	−0.048	−0.046	−0.040	0.008
	(0.037)	(0.037)	(0.042)	(0.037)	(0.035)
Receive Workshop*Post-Intervention*High Risk	0.043	0.043	0.025	0.006	0.044
	(0.036)	(0.036)	(0.041)	(0.037)	(0.038)
Panel E: Profitable client for the bank					
Receive Workshop*Post-Intervention	0.031**	0.024	−0.010	0.022	0.033*
	(0.015)	(0.016)	(0.020)	(0.018)	(0.020)
Receive Workshop*Post-Intervention*Med Risk	−0.021	−0.010	0.014	−0.012	−0.035
	(0.025)	(0.027)	(0.028)	(0.025)	(0.026)
Receive Workshop*Post-Intervention*High Risk	0.009	0.015	0.001	0.006	0.011
	(0.028)	(0.029)	(0.032)	(0.029)	(0.033)

Source: Authors' analysis based on the study implementation data provided by BBVA Bancomer.

Note: Robust standard errors in parentheses, clustered at the client level. *, **, *** denote significance at the 10, 5, and 1 percent levels only. The five columns show estimated treatment impacts of take-up in coaching treatment, using as control groups: all clients randomly assigned to the control (column 1); those within the common support when matching on all pre-intervention variables (column 2); using single nearest-neighbor matching within this common support (column 3); using single nearest-neighbor matching with the common support when using lasso to select variables for propensity score (column 4), and using nearest neighbor matching within the common support (column 5).

S3. Additional Outcomes

Additional outcomes are investigated to shed some light on the drivers behind the results. These variables help explain where the increase in profits of the bank is coming from, and whether financial education helps individuals improve their creditworthiness and better manage their credit card debt loads, by for instance, avoiding unnecessary interests and fees.

Table S3.1. Estimated Treatment Effects on Other Outcomes for Those Who Received Workshops

			Nearest-neighbor matching		
	Full control sample	**In common support**	**on all variables**	**using lasso**	**on outcome**
Panel A: >90 days delay in payment					
Receive Workshop*Post-Intervention	−0.023***	−0.019***	−0.017	−0.014***	−0.013***
	(0.002)	(0.007)	(0.015)	(0.005)	(0.005)
Sample size	967,442	127,429	5,275	28,888	30,037
Mean	0.029	0.029	0.027	0.019	0.019
p-values for test common linear pre-trend	0.287	.	.	0.340	1.000
p-values for test common non-linear pre-trend	.	.	.	0.391	0.849
Panel B: Log monthly balance on card					
Receive Workshop*Post-Intervention	0.218***	0.133	0.006	0.147*	0.250***
	(0.057)	(0.113)	(0.145)	(0.081)	(0.082)
Sample size	942,989	127,198	5,261	28,710	26,587
Mean	9.296	10.43	10.96	10.10	9.957
p-values for test common linear pre-trend	0.058	0.563	0.837	0.873	0.281
p-values for test common non-linear pre-trend	0.293	0.291	0.797	0.000	0.993
Panel C: Log net income from interest and fees					
Receive Workshop*Post-Intervention	0.323**	−0.227	0.264	0.544**	0.154
	(0.149)	(0.380)	(0.538)	(0.214)	(0.219)
Sample size	449,122	58,942	2,436	13,360	13,948
Mean	4.269	5.982	5.759	5.600	3.845
p-values for test common linear pre-trend	0.099	0.722	0.316	0.028	0.547
p-values for test common non-linear pre-trend	0.161	0.757	0.677	0.000	0.926
Panel D: Number of BBVA products					
Receive Workshop*Post-Intervention	0.146***	0.200**	0.262*	0.549***	0.195***
	(0.042)	(0.079)	(0.134)	(0.058)	(0.062)
Sample Size	1,002,206	132,570	5,481	30,024	31,235
Mean	4.646	6.553	7.069	6.398	5.089
p-values for test common linear pre-trend	0.211	0.604	0.197	0.000	0.851
p-values for test common non-linear pre-trend	0.230	0.956	0.984	0.000	0.916
Panel E: Monthly purchases/Credit limit					
Receive Workshop*Post-Intervention	0.011	0.011	0.006	0.028	0.023
	(0.014)	(0.015)	(0.024)	(0.018)	(0.016)
Sample size	270,114	119,007	4,933	17,604	5,582
Mean	0.114	0.086	0.101	0.219	0.073
p-values for test common linear pre-trend	0.323	0.423	0.837	0.000	0.677
p-values for test common non-linear pre-trend	0.175	0.709	0.435	0.000	0.937

Source: Authors' analysis based on the study implementation data provided by BBVA Bancomer.

Note: Robust standard errors in parentheses, clustered at the client level. *, **, *** denote significance at the 10, 5, and 1 percent levels only. The five columns show estimated treatment impacts of taking part in the coaching treatment, using different control groups. Column (1) uses all clients randomly assigned to the control; column (2) uses those within the common support when matching on all pre-intervention variables; column (3) uses single nearest-neighbor matching within this common support; column (4) uses single nearest-neighbor matching with the common support when using lasso to select variables for propensity score, and then nearest-neighbor matching within the common support.

Table S3.2. Estimated Treatment Effects on Other Outcomes for Those Who Received Workshops

	Full control sample	In common support	Nearest neighbor-matching on all variables	Nearest neighbor-matching using lasso	Nearest neighbor-matching on outcome
Panel F: Share of balance paid					
Receive Workshop*Post-Intervention	0.008	−0.006	0.080	0.118***	0.058*
	(0.029)	(0.069)	(0.088)	(0.039)	(0.035)
Sample size	861,249	121,955	5,051	27,330	22,852
Mean	0.526	0.312	0.320	0.547	0.451
p-values for test common linear pre-trend	0.729	0.784	0.799	0.858	0.650
p-values for test common non-linear pre-trend	0.741	0.456	0.855	0.000	0.380
Panel G: Full payment of balance					
Receive Workshop*Post-Intervention	0.003	0.010	0.018	0.063***	0.020**
	(0.007)	(0.015)	(0.021)	(0.010)	(0.010)
Sample size	861,249	121,955	5,051	27,330	22,895
Mean	0.116	0.059	0.066	0.125	0.093
p-values for test common linear pre-trend	0.741	0.624	0.890	0.572	0.251
p-values for test common non-linear pre-trend	0.607	0.561	0.861	0.000	0.995

Source: Authors' analysis based on the study implementation data provided by BBVA Bancomer.
Note: Robust standard errors in parentheses, clustered at the client level. *, **, *** denote significance at the 10, 5, and 1 percent levels only. The five columns show estimated treatment impacts of taking part in the coaching treatment, using different control groups. Column (1) uses all clients randomly assigned to the control; column (2) uses those within the common support when matching on all pre-intervention variables; column (3) uses single nearest-neighbor matching within this common support; column (4) uses single nearest neighbor matching with the common support when using lasso to select variables for propensity score, and then nearest neighbor matching within the common support.

These additional outcomes are 1) likelihood of charge-offs, defined as an indicator variable equal to 1 if a payment is late for ninety or more days and 0 otherwise; 2) monthly revenue collected by the bank from interest and fee payments minus operational expenses (in logs); 3) number of BBVA products that a client has at each month (as proxy for product cross-selling); 4) ratio of credit card purchases to credit limit (as proxy for credit line utilization); 5) log of revolving balances; 6) share of total credit card debt that was paid at each month; 7) likelihood of fully paying outstanding credit card balance. These outcomes are included when matching the compliers in the treatment group to similar individuals in the control group via propensity score matching. As with this study's baseline approach, the control groups are selected relying on five different matching methods. The study then uses difference-in-differences on these matched samples to estimate the impact of attending the workshops or receiving coaching.

The results, summarized in tables S3.1 to S3.4, provide some explanations for the increase in profits experienced by the bank. The reduced likelihood of delinquency and charge-offs (of 3.4 and 1.3 pp for workshops and 2.6 and 3.6 pp for coaching) can help BBVA Bancomer reduce collection costs. Participants of the interventions also increase credit card monthly purchases, which should in turn translate into higher interchange fees for BBVA Bancomer. While interchange fees are not observed in the data, it is found that bank revenues from fees and interest appear to increase. Importantly, the fact that participants are more likely to pay on time and pay more than the minimum payment, suggests that the increase in revenue is not coming from extra fees from late or missed payments. Participants of the interventions, particularly workshops, increase their number of BBVA products. However, the economic magnitude of this effect seems low, suggesting that cross-selling is likely not a major driver of the increased profits of the bank. Compared to clients in the control group that have on average 5.09 BBVA products, clients receiving workshops increase their number of products to 5.28.

The results further suggest that both coaching and workshops help clients improve their creditworthiness by improving their payment timeliness (measured by the reduction in delinquency rates) and reducing their probability of default (proxied by the drop in charge-off rates).

Table S3.3. Estimated Treatment Effects on Other Outcomes for Those Who Received Coaching

	Full control sample	In common support	Nearest-neighbor matching on all variables	Nearest-neighbor matching using lasso	Nearest-neighbor matching on outcome
Panel A: >90 days delay in payment					
Receive Coaching*Post-Intervention	−0.041***	−0.040***	−0.023*	−0.045***	−0.036***
	(0.007)	(0.009)	(0.012)	(0.014)	(0.013)
Sample size	70,498	14,307	4,601	12,041	12,628
Mean	0.062	0.045	0.027	0.061	0.051
p-values for test common linear pre-trend	0.740	0.582	0.994	0.532	0.622
p-values for test common non-linear pre-trend	0.000	0.582	0.994	0.262	0.722
Panel B: Log monthly balance on card					
Receive Coaching*Post-Intervention	0.239**	−0.035	−0.101	0.464***	−0.021
	(0.120)	(0.161)	(0.187)	(0.150)	(0.157)
Sample size	69,407	14,288	4,592	11,960	11,416
Mean	9.47	10.83	10.86	10.08	10.31
p-values for test common linear pre-trend	0.003	0.192	0.527	0.000	0.707
p-values for test common non-linear pre-trend	0.015	0.931	0.936	0.000	0.929
Panel C: Log net income from interest and fees					
Receive Coaching*Post-Intervention	0.954***	1.178***	0.063	1.304***	1.111***
	(0.220)	(0.342)	(0.424)	(0.331)	(0.345)
Sample size	32,987	6,624	2,124	5,568	5,880
Mean	4.857	6.062	6.722	5.117	4.605
p-values for test common linear pre-trend	0.442	0.911	0.887	0.738	0.636
p-values for test common non-linear pre-trend	0.155	0.810	0.945	0.044	0.992
Panel D: Number of BBVA products					
Receive Coaching*Post-Intervention	0.131*	0.118	0.110	0.333***	0.176*
	(0.070)	(0.088)	(0.126)	(0.100)	(0.095)
Sample size	73,706	14,900	4,777	12,513	13,190
Mean	5.617	6.872	6.942	6.008	5.949
p-values for test common linear pre-trend	0.754	0.587	0.636	0.502	0.412
p-values for test common non-linear pre-trend	0.101	0.999	0.866	0.087	0.972
Panel E: Monthly purchases/credit limit					
Receive Coaching*Post-Intervention	0.003	0.011	0.017	0.004	0.023
	(0.011)	(0.016)	(0.024)	(0.015)	(0.016)
Sample size	43,855	13,467	4,344	9,250	6,116
Mean	0.064	0.074	0.092	0.057	0.078
p-values for test common linear pre-trend	0.890	0.982	0.988	0.648	0.599
p-values for test common non-linear pre-trend	0.786	0.940	0.999	0.135	0.985

Source: Authors' analysis based on the study implementation data provided by BBVA Bancomer.
Note: Robust standard errors in parentheses, clustered at the client level. *, **, *** denote significance at the 10, 5 and 1 percent levels only. The five columns show estimated treatment impacts of taking part in the coaching treatment, using different control groups. Column (1) uses all clients randomly assigned to the control; column (2) uses those within the common support when matching on all pre-intervention variables; column (3) uses single nearest-neighbor matching within this common support; column (4) uses single nearest-neighbor matching with the common support when using lasso to select variables for propensity score, and then nearest neighbor matching within the common support.

Altogether, the evidence suggests that the financial education interventions help clients to manage their credit cards better. Clients make substantially more purchases with their credit cards. This increase in purchases in turn increases revolving balances by 25 percent for the workshop group, with no change for clients that took coaching. This change in purchases is not accompanied by an increase in payment delays, or by paying a lower fraction of the outstanding balance. What is more, the share of debt paid of clients in the treated arms is similar to that of the control group, and the fraction of clients paying their balance in full remains unchanged even though treated clients use their credit cards more.

Table S3.4. Estimated Treatment Effects for Those Who Received Coaching

			Nearest-neighbor matching		
	Full control sample	In common support	on all variables	using lasso	on outcome
Panel F: Share of balance paid					
Receive Coaching*Post-Intervention	−0.012	−0.064	−0.114	0.042	−0.030
	(0.037)	(0.049)	(0.096)	(0.046)	(0.056)
Sample size	61,452	13,706	4,409	11,296	9,191
Mean	0.271	0.261	0.373	0.237	0.287
p-values for test common linear pre-trend	0.290	0.221	0.899	0.001	0.826
p-values for test common non-linear pre-trend	0.042	0.940	0.695	0.000	0.978
Panel G: Full payment of balance					
Receive Coaching*Post-Intervention	0.001	−0.020	−0.011	0.015	−0.003
	(0.009)	(0.013)	(0.019)	(0.011)	(0.011)
Sample Size	61,452	13,706	4,409	11,296	9,235
Mean	0.049	0.044	0.046	0.038	0.040
p-values for test common linear pre-trend	0.362	0.297	0.815	0.021	0.648
p-values for test common non-linear pre-trend	0.120	0.944	0.495	0.001	0.913

Source: Authors' analysis based on the study implementation data provided by BBVA Bancomer.

Note: Robust standard errors in parentheses, clustered at the client level. *, **, *** denote significance at the 10, 5, and 1 percent levels only. The five columns show estimated treatment impacts of taking part in the coaching treatment, using different control groups. Column (1) uses all clients randomly assigned to the control; column (2) uses those within the common support when matching on all pre-intervention variables; column (3) uses single nearest-neighbor matching within this common support; column (4) uses single nearest-neighbor matching with the common support when using lasso to select variables for propensity score, and then nearest-neighbor matching within the common support.

S4. Heterogeneous Treatment Effects

For each intervention, the study estimates heterogeneous treatment effects (HTEs) on predicted take-up to investigate whether the ITT effects of workshops and coaching vary depending on predicted take-up. To do this, the sample is limited to clients in the treatment group, and the study fits a probit model relating the probability of take-up in the future with the variables used to obtain the common support sample in the pre-training months. These 73 variables include gender and the monthly levels of the five benchmark outcomes (likelihood of paying more than the minimum payment, likelihood of delay in payment, log of monthly spending, likelihood of having a deposit account, and likelihood of being a profitable client). From this regression, the predicted take-up rates are obtained of all individuals in our sample.

The study then runs the treatment effects of the regression summarized in equation (10), allowing for heterogeneous effects based on the fitted probability of take-up. That is, the study includes the new variable of predicted take-up as well as its interaction with the *TreatmentOffered* variable. The study is then interested in the coefficient of this interaction, which identifies if the treatment effect is higher for those who are more likely to take up the intervention. The results from this regression, summarized in table S4.1, find little in the way of statistically differential effects for individuals with higher take-up rates. However, this reflects a combination of both large standard errors, and the difficulty in predicting precisely which individuals will be compliers to the treatment. As noted in section 5, with the matched difference-in-differences approach, it is possible instead to identify the compliers in the treatment group precisely, and then it is only necessary to find individuals in the control group whose time trends on specific outcomes match those of the compliers in the treatment group, which allows for more precision.

Table S4.1. Heterogeneous Treatment Effects on Predicted Take-up

	Share of debt paid by due date	Client classified as not in good standing	Delay in payment	Pays more than minimum	Has basic deposit account with bank	Profitable client (bai)	Profitable client (nibt)	Log of monthly balance	Log of monthly spending
Panel A: Impact of workshops									
Offered workshops	−0.008	−0.005	0.001	0.005	0.002	0.002	0.004	0.049**	0.045
	(0.007)	(0.007)	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.020)	(0.035)
Predicted takeup	4.260***	−3.363***	−4.098***	4.878***	3.735***	2.324***	2.668***	38.828***	52.326***
	(0.676)	(0.534)	(0.230)	(0.291)	(0.278)	(0.275)	(0.309)	(1.618)	(3.110)
Offered workshops*Predicted takeup	1.102	0.639	0.083	−0.445	−0.251	0.001	−0.047	−3.798**	−4.589
	(0.806)	(0.653)	(0.290)	(0.352)	(0.333)	(0.330)	(0.371)	(1.910)	(3.346)
Sample size	649,527	197,376	662,888	647,085	653,357	501,182	501,182	656,850	662,888
Panel B: Impact of Coaching									
Offered coaching	−0.022	0.006	0.012	0.004	−0.017	−0.010	−0.008	−0.007	−0.039
	(0.016)	(0.017)	(0.013)	(0.013)	(0.012)	(0.012)	(0.012)	(0.058)	(0.118)
Predicted takeup	−0.045	−0.250**	−0.287***	0.483***	−0.062	0.250***	0.302***	2.200***	4.551***
	(0.109)	(0.127)	(0.083)	(0.108)	(0.093)	(0.088)	(0.091)	(0.373)	(1.044)
Offered coaching*Predicted takeup	0.293*	−0.005	−0.033	0.012	0.194	0.072	0.070	−0.132	0.336
	(0.159)	(0.163)	(0.119)	(0.134)	(0.123)	(0.113)	(0.117)	(0.509)	(1.272)
Sample size	34,335	21,793	34,911	34,186	34,682	26,503	26,503	34,750	34,911

Source: Authors' analysis based on the study implementation data provided by BBVA Bancomer.

Note: Robust standard errors in parentheses, clustered at the client level. *, **, *** denote significance at the 10, 5, and 1 percent levels respectively. Estimation is by Ancova, and includes mean of outcome over baseline periods, time period fixed effects, and strata fixed effects.