Incentives and Engagement: A Field Experiment to Evaluate the Effect of Survey Incentives on MIDS Alumni

Robert Hosbach, Thomas Gao, Devesh Khandelwal, Joe Mirza

DATASCI W241 | Summer 2021

Abstract

Studies have shown a variety of effects from different incentives to increase survey response rates. We conduct an experiment to test whether a direct financial incentive or a philanthropic incentive cause increased survey response rates among 1,105 alumni of the UC Berkeley I School's Master of Information and Data Science program. Surprisingly, we find that both incentives caused lower survey response rates, with the direct financial incentive causing significant intent-to-treat and complier average causal effects of approximately -6% and -10%, respectively. While these results likely do not generalize beyond the current population of MIDS alumni that comprised our sample frame, future research that evaluates different incentive levels and incentive types on survey response rates may prove fruitful for determining more general best practices for increasing survey response rates among graduate school alumni.

Introduction

A wealth of existing research shows that using monetary incentives can increase survey response rates (see Armstrong, 1975; Kanuk & Berenson, 1975; and Duncan, 1979). James & Bolstein (1992) found that mailing an incentive of only $1 with the survey significantly increased the response rate compared to people who received no incentive. More recently, Debell et al. (2020) found that including $5 of visible money in a mailed survey increased response rates from roughly 43% to 47%. Researchers have also studied response rates for surveys administered to college students. For instance, Szelényi et al. (2005) found that a $2 incentive led to increased response rates among college students compared to no incentive.

This body of research suggests that direct financial incentives can lead to higher survey response rates. But, while researchers have studied direct financial incentives in depth, the effect of indirect, philanthropic donations are comparatively less well understood. Gattellari & Ward (2001) found that offering donations to Australian surgeons' alma maters actually led to lower response rates when compared to offering no incentive at all. Nesrallah et al. (2014) and Warwick et al. (2019) had similar findings in medical-related settings. Outside of the medical field, Pedersen & Nielsen (2016) also found that incentives promising donations to a good cause led to decreased response rates relative to no incentives in online surveys.

We conducted a field experiment to test the effect of two different incentives on the survey response rates from alumni of UC Berkeley's Master of Information and Data Science (MIDS) program. One incentive (hereafter referred to as the "direct incentive") provides a chance to win a $25 Amazon gift card if the alumnus completes the survey. This direct financial incentive is akin to the direct financial incentives discussed above. The other incentive (hereafter referred to as the "philanthropic incentive") provides an opportunity to have us, the survey developers, donate funds to the Berkeley Student Food Collective (BSFC) if we achieve a sufficiently high survey response rate. While some studies found donation-based incentives counterproductive, our experiment takes a slightly different approach. Instead of promising a donation to BSFC for an individual's response, we will test the effect of survey response rates when we tell individuals that the donation will only occur if we meet a certain threshold of *group* response rate.

In this experiment, we evaluated the effects of the direct and philanthropic incentives to cause higher survey response rates when compared to a control group that received no incentive. We gave all three groups the same survey that we distributed via one of our personal email addresses; however, only alumni in the direct and philanthropic incentive groups received text in their emails regarding the incentive assigned to them. We hypothesized that both the direct and philanthropic incentives would cause higher survey response rates among the MIDS alumni. As mentioned, direct incentives have led to increased survey response rates in certain contexts, whereas donation-based incentives have shown far less promise. Nevertheless, we still expected the philanthropic incentive to garner higher survey response rates relative to the control group for the following reasons: 1) a group incentive, when compared to a personal incentive, may prove more inherently motivating for some people; 2) people may prefer the certainty of the donation to charity after reaching the response threshold over the uncertainty of receiving a personal monetary reward; and 3) trying to help a good cause can evoke a sense of meaning when compared to receiving money personally.

In the following section we describe our methodology for conducting this experiment, including the tools we used, our data collection process, how we dealt with attrition and non-compliance, and other details. Next, we report our experimental results and provide a discussion of their meaning and generalizability. We also discuss potential follow-on opportunities from this experiment. Finally, we conclude our paper and provide a series of appendices containing the email text we used, Python and R code we used for data collection and analysis, the actual survey, and a summary of the survey responses we received.

Methodology

In this section we explain the key components of our experimental design, including our sampling frame; data collection and blocking; power analysis; survey design and email sending strategy; identification of compliers, non-compliers, and attriters; and calculation of treatment effects.

Scraping alumni data

To obtain the official UC Berkeley School of Information (hereafter referred to as UCB I School) email addresses for all MIDS alumni, our team wrote a Python script to scrape selected pages from the UCB I School website's people directory ("People," 2021). The script relies on the Selenium package for Python, an open-source tool for browser automation (Muthukadan, 2021). We provide the script for reference in Appendix B.

Specifically, the team scraped information from all pages listing MIDS graduates, of which the website lists 1,105 as of the end of Spring 2021. These 1,105 alumni represent the sampling frame for this experiment. The web pages include not only the I School email address of each alumnus, but also the person's name and graduation year.[1]

Blocking

In order to improve the precision of our estimated causal effect, we randomly assigned the 1,105 alumni into the control and two treatment groups using blocks based on graduation year (2015-2021) and gender (male, female, or unknown). We expect blocking on graduation year to improve the precision of our estimated causal effect because of the likelihood that earlier MIDS graduates might have lost connection with the I School community relative to more recent graduates, leading to lower survey response rates, on average, for earlier graduates. Similarly, we blocked on gender because some studies have found that women may respond to surveys at higher rates than men in a variety of contexts (*e.g.*, Smith, 2008).

The UCB I School website's people directory does not provide alumni gender designations. Therefore, we inferred each alumnus' gender based on the person's name and photograph (from the people directory or LinkedIn, if available).[2] Out of the 1,105 MIDS alumni, we classified 789 (71.4%) as male and 299 (27.1%) as female. We classified the remaining 17 (1.5%) alumni as "unknown" because we did not have enough information to confidently classify them as male or female.

We used the blockTools library in R to block-randomize the 1,105 alumni into the control and two treatment groups by blocking on graduation year and inferred gender (Moore & Schnakenberg, 2016). Figure 1 shows the number of alumni assigned to each of the three groups by graduation year and gender. Each cell in Figure 1 contains three numbers, which correspond to the number of alumni assigned to each of three groups (control and two treatment) for that block. For a given block, we assign the same number of alumni to each group, when possible. In cases where the number of alumni in the block do not evenly divide into three groups, the number of alumni assigned to each group for a given block never differs by more than one.

[1] The scraped data did not contain two email addresses and one graduation year. We identified the email addresses using the UCB I School Slack workspace and validated the missing grladuation year using LinkedIn.

[2] We used a binary gender designation: male or female. Without more data, we could not make more specific gender designations. Additionally, blocking based on more disaggregated gender designations seems unlikely to improve the precision of our causal effect estimate relative to a binary gender designation.

Block-Randomization:
Alumni Assigned to each Group per Block

Inferred Gender	Graduation Year 2015	2016	2017	2018	2019	2020	2021
	2	6	13	19	20	29	10
Female	3	6	13	19	20	29	10
	2	6	14	19	21	29	9
	11	32	36	46	60	60	19
Male	10	32	36	46	60	61	18
	10	32	36	46	59	61	18
	0	0	2	1	1	1	1
Unknown	0	0	2	1	1	1	0
	1	1	1	1	2	0	0

Figure 1: Number of MIDS alumni block-randomized into each of three groups (control and two treatment) based on inferred gender and graduation year

Power analysis

We conducted a power analysis in advance of running our experiment to determine the likelihood that we would detect a significant effect if the effect were real. The studies referenced in the Introduction provide a range of survey incentive effects in different contexts. Debell et al. (2020) observed an increase in response rate from 43% to 47% (a 4% effect) by making a cash incentive plainly visible in a physically mailed envelope. In contrast, Pedersen & Nielsen (2016) found a variety of effects that ranged from -3.5% given a philanthropic incentive to 4.5% given an egotistic text appeal treatment.

We used Alexander Coppock's Power Calculator application to perform this analysis (Coppock, 2021). The application requires certain inputs that we set as follows:

- **Binary dependent variable:** selected
- **Significance level:** 0.05
- **Proportion (DV = 1) in Control Group:** 0.1
- **Proportion (DV = 1) in Treatment Group:** 0.2
- **Power Target:** 0.8

The inputs indicate that we expect a control group response rate of 10%, and we hope (likely optimistically) to observe treatment group response rates of 20%. Given these assumptions, the application's output indicates that we need a sample size of at least 393 people to detect a difference between the control group and treatment group at the 5% significance level with 80% power. This breaks out to roughly 196 people per group, and we assigned approximately 370 alumni to each of our three groups. Therefore, our experiment can detect such a difference with more than 80% power. For reference, changing the assumed response rate in the treatment group from 20% to 15% while keeping the control group's response rate at 10% means we would need a sample size of 1,366 to achieve 80% power. This breaks out to 683 people per group, which is nearly double the allotment in our experiment. According to the application, with our group sizes of roughly 370 alumni, we can only detect this difference at a 5% significance level with 54% power. In sum, our experiment will lack sufficient power to detect an effect if we observe small differential response rates between the control and treatment groups.

Experimental design and survey delivery architecture

Our experiment followed a progression of randomization, treatment, and outcome measurement. Figure 2 depicts the ROXO representation of our experiment. We applied the same ROXO design to each of the 21 blocks we randomized within (denoted by the blue arrows). In Figure 2, "R" represents the block-randomization step, "X_1" and "X_2" represent the different treatment emails sent to the direct and philanthropic incentive groups, and "O" represents measurement of survey completions (our outcome).

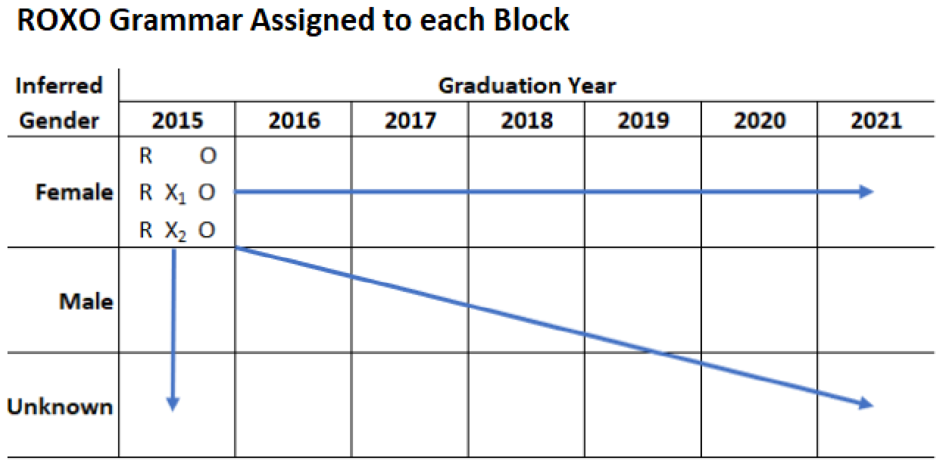

Figure 2: ROXO construct applied to each of the 21 randomization blocks

We delivered the control and treatment messages to the three groups of MIDS alumni using emails sent from Devesh Khandelwal's personal I School email address (deveshkhandelwal@berkeley.edu). (We found from testing that using a less personal email address resulted in more of our emails landing in recipients' "promotions" or "social" Gmail inboxes, rather than their "primary" inboxes.)

We used identical email subject lines for all three groups for the initial emails as well as the reminder emails (see Table 1). Moreover, we kept the body text of the emails identical for all three groups, except for the specific treatment text that we accentuated using bold typeface.

Table 1: Email subject lines for the initial and reminder emails

Email	Subject
Initial (July 2, 2021)	«Alumnus Name»-How Can MIDS Improve? Voice Your Opinion
July 9 and 16, 2021 Reminders	Reminder: «Alumnus Name»-How Can MIDS Improve? Voice Your Opinion
July 22, 2021 Reminder	ONE DAY LEFT : «Alumnus Name»-How Can MIDS Improve? Voice your opinion
July 23, 2021 Reminder	CLOSING TONIGHT: «Alumnus Name»-How Can MIDS Improve? Voice your opinion

For all three groups, the treatment text indicates that alumni who complete the survey will receive a summary of the survey results. In addition, we told the direct incentive group that we will enter alumni who complete the survey into a drawing to win a $25 Amazon gift card, and we told the philanthropic incentive group that if the overall survey response rate reached 60%, we would donate $250 to the Berkeley Student Food Collective. Table 2 shows the treatment text for all three groups, and Appendix A includes the original and reminder email templates sent to all three groups.

Table 2: Treatment email text for the control and two treatment groups

Group	Subject
Control	If you complete the survey, we will send you a summary of the results.
Direct incentive	If you complete the survey, we will send you a summary of the results and you will be entered to win an Amazon gift card for $25. Ten respondents will be selected at random to receive a gift card.
Philanthropic incentive	If you complete the survey, we will send you a summary of the results. Additionally, if we achieve a 60% response rate, we will donate $250 to the Berkeley Student Food Collective.

We measured survey completions as the outcome metric for our experiment. We designed a two-part survey using Qualtrics to ask about the perceived value of various aspects of the MIDS program (part one) as well as to understand the respondent's rationale for enrolling in MIDS (part two). We hypothesized that having a survey directly related to the MIDS experience and perceived benefits would invoke higher response rates among MIDS alumni. The first part asks the respondent to evaluate eight statements using a Likert scale. The second part asks the respondent to answer a single multiple choice question. We developed the survey so that respondents could complete it in 2-3 minutes, in order to maximize the likelihood of any given alumnus completing the survey. We provide the survey questions and a high-level results summary in Appendix C.

Sending emails and reminders

We sent initial emails to the control and two treatment groups on Friday, July 2, 2021 using Yet Another Mail Merge (YAMM). The YAMM service took approximately 12 minutes to send the emails to the 370 (approximately) alumni in each group. Table 3 indicates the start times for sending emails to each group.

Table 3: Start times for sending initial emails to the control and two treatment groups

Group	Email Batch Start Time on July 2, 2021 (PDT)
Direct incentive	9:26 am
Philanthropic incentive	10:14 am
Control	10:40 am

In addition to the initial email that we sent to all 1,105 alumni, we sent four follow-up reminder emails. We sent the first reminder email to all 1,105 alumni starting at 9:00 am (PDT) on July 9, 2021 (even to those who had already completed the survey). We sent the remaining three reminder emails starting at 9:15 am (PDT) on July 16, 8:00 am (PDT) on July 22, and 9:15 am (PDT) on July 23, 2021 only to alumni who had not already completed the survey by the evening prior to sending the emails. Figure 3 shows the cumulative number of completed surveys over time, with dashed lines signifying when we sent emails. The figure shows a clear pattern of a sharp increase in survey responses immediately after we sent the emails, with diminishing returns over time.

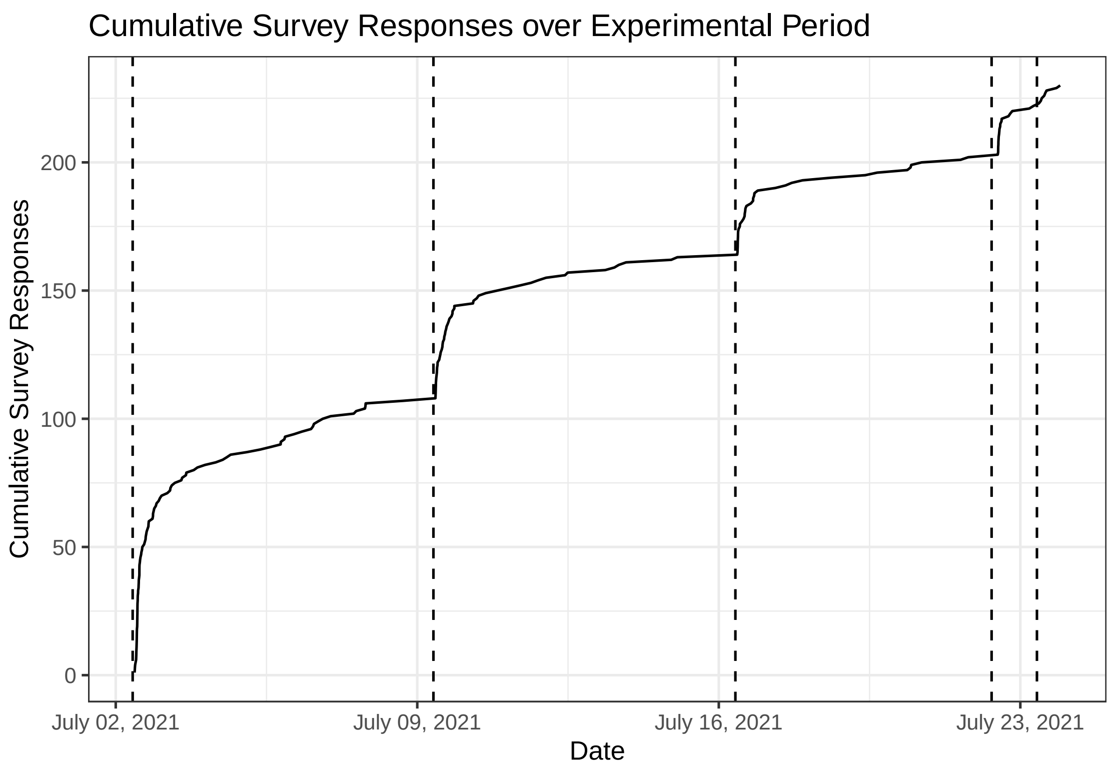

Figure 3: Cumulative survey completions over the experimental period with initial and reminder email send dates indicated by dashed vertical lines

Identifying compliers, non-compliers, and attriters

The YAMM service provides a near-real time merge status indicator for each email that it sends. YAMM labels each email as sent, opened, clicked, responded, bounced, or unsubscribed; however, YAMM does not guarantee 100% accuracy in applying these labels ("How Accurate Is Our Tracking Tool?" n.d.). Between the different email batch sends, we noticed inconsistencies in the labeling. For instance, we observed some alumni with bounced emails after sending the first reminder email who did not have their initial email bounce. We also observed the opposite, where alumni with emails that initially bounced were later relabeled as opened. We investigated these inconsistencies and applied the following assumptions to our data in order to identify attriters, compliers, and non-compliers:

- We considered any alumni with initial emails that were labeled as bounced right after sending to have inactive email accounts, and therefore we assumed all subsequent emails bounced as well (regardless of the actual labels YAMM applied).
- We considered any alumni with all emails labeled as sent (not opened, clicked, or responded) to have received, but not opened, any of their emails.
- We assumed the remaining alumni received and opened at least one of their emails, thereby exposing themselves to the treatment text contained in the email.

We placed the treatment text specific to the control, direct incentive, and philanthropic incentive groups in the body of the emails, rather than in the email subject lines. This allowed us to define a complier as someone who opened the email, given the assumption that if the person opened the email they read the treatment text. (Alternatively, we could have defined compliance as reception of the email; however, that criterion seemed too lenient for assuming the recipient actually received the treatment.) We therefore identified compliers as those alumni for whom YAMM reported that they opened, clicked, or responded to at least one of the emails we sent them. Conversely, we define non-compliers as those who receive at least one of the emails, but never open any emails they receive. As a result, these alumni never receive a dosage of the treatment corresponding to the group to which we assigned them. We identified non-compliers in our study as those for whom the email did not bounce, and YAMM reported that the person never opened, clicked, or responded to any of the emails we sent them. Finally, we define attriters as those alumni who never receive any of the emails we sent them. We deemed any alumni for whom YAMM reported a bounced email from our initial email batch an attriter. The flow diagram in Figure 4 shows the number of alumni in each group after accounting for attriters and non-compliers. In total, 229 MIDS alumni completed the survey, providing an overall response rate of approximately 21%.

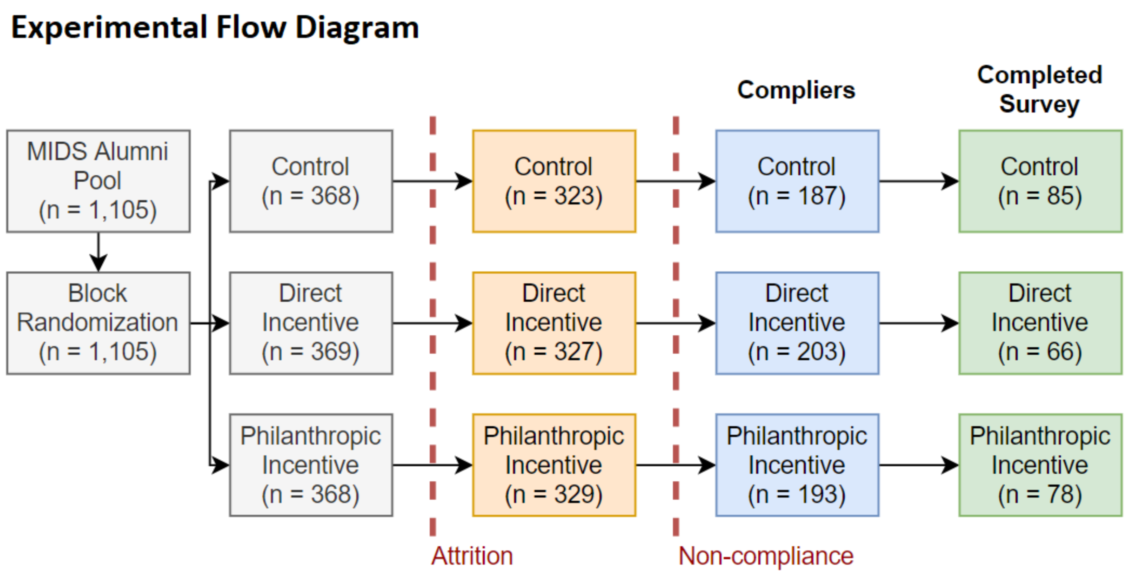

Figure 4: Flow diagram indicating the number of MIDS alumni by group after accounting for attrition and non-compliance

Data flow and aggregation

As highlighted in the previous sections, our data come from three main sources:

- Alumni scrape: consists of alumni details that we obtain from web scraping.
- YAMM: data from the mail-merge tool that helps track the status of emails we sent.
- Qualtrics: the survey platform we used that provides information on survey respondents.

Figure 5 represents the high-level data flow and aggregation of our data sources to arrive at our final data source for analysis.

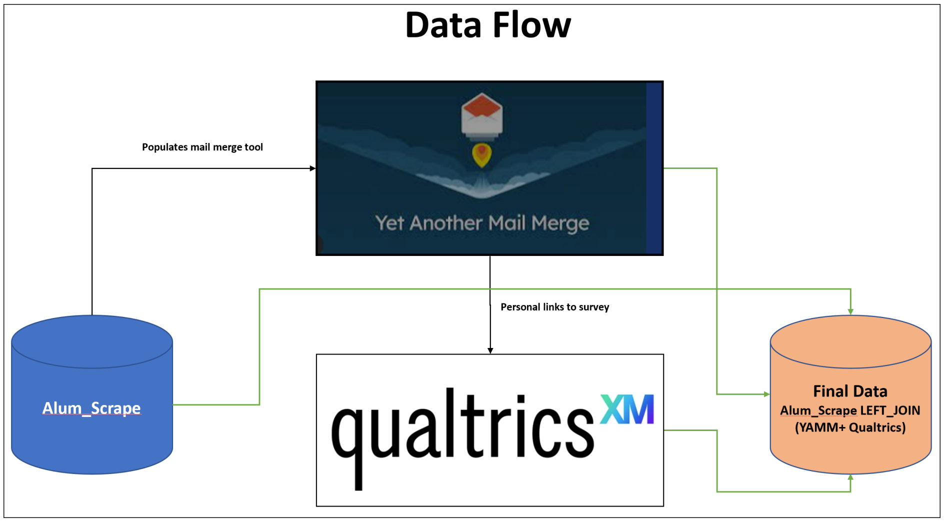

Figure 5: Summary diagram of data flows and aggregation

Calculating treatment effects

For this experiment, we calculated the following treatment effects: intent-to-treat (ITT) effect, complier-average causal effect (CACE), and heterogeneous treatment effects (HTEs) using gender and graduation year subgroups. When calculating these treatment effects, we omitted attriters from the analysis. We omitted attriters for two reasons: 1) We do not have a measured outcome for these alumni, and 2) Omitting attriters should not bias our estimated treatment effects because the group assignment has no influence on whether or person's email account is active or not (which determines whether an email bounces).

The ITT effect calculation provides an estimated treatment effect based only on the group assignment. Given our concerns regarding the accuracy of our compliance data, we chose to use the ITT as our primary treatment effect. The CACE, in contrast, does require knowledge of the number of compliers and non-compliers in each group. Due to the aforementioned issues with YAMM, obtaining these estimates required us to make assumptions about who actually opened at least one of the emails they received. As a result, we do not have as much trust in our CACE estimates as we do for our ITT effect estimates. Finally, we analyzed HTEs for gender and graduation year subgroups based only on group assignment. This allowed us to avoid making compliance assumptions, and it means that one should interpret the estimates we obtained as heterogeneous ITT effects.

Analysis assumptions

To obtain unbiased causal effect estimates, we satisfied a series of key assumptions:

- **Exclusion restriction:** We assume that potential outcomes in our experiment respond only to the treatment a subject receives, and not the group assignment we gave the subject or any other causal pathway. While we cannot verify this assumption empirically, we have no reason to believe that the exclusion restriction is violated in our experiment.
- **Non-interference:** We assume that no subjects are affected by treatment of other subjects in our experiment. We sent personal emails containing specific treatment text according to each subject's group assignment, which prevents interference. While it remains possible that some MIDS alumni discussed the survey and emails we sent them, we assume no spillover effects due to the small likelihood of that occurring.
- **Monotonicity:** We assume that none of our subjects are defiers (*i.e.*, people who receive treatment when assigned to control, or vice-versa). We can safely make this assumption due to the personal emails we sent each MIDS alumnus.

Results and Discussion

In this section we will discuss our covariate balance checks to validate our randomization; provide estimated ITT, CACE, and HTE effects, and discuss the generalizability of our results.

Checks and balances

We checked for covariate balance to validate our randomization, which ensures a proper "apples-to-apples" comparison. Our block randomization ensures balance in inferred gender and graduation year among our three groups; so, here we used the YAMM data to analyze the balance of compliers, attriters and non-compliers between groups. Figure 6 plots the number of attriters, non-compliers, and compilers for each group. By visual inspection, we see balance between the groups.

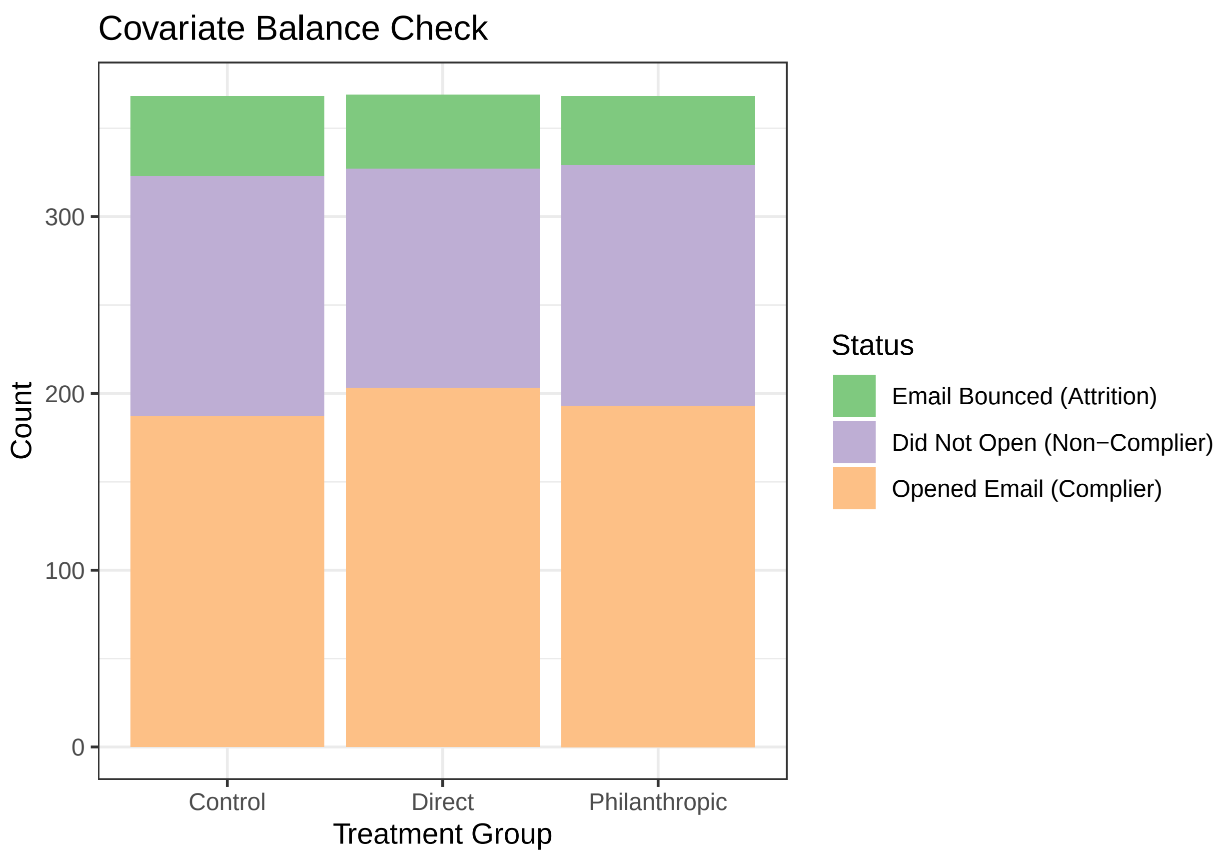

Figure 6: Stacked bar chart showing number of attriters, compliers, and non-compliers in each group

More formally, Table 4 provides the results of a regression analysis of attrition, non-compliance, and compliance on the group assignment.[3] As shown by the lack of statistically significant coefficients, we did not find statistically significantly different results between the control and two treatment groups in this analysis.

[3]Robust standard errors are provided in this regression table as well as the forthcoming regression tables.

Table 4: Covariate balance regression analysis results

Covariate Balance Check

	Dependent variable:		
	Attrition	Non-Compliance	Compliance
	(1)	(2)	(3)
Group: Direct	-0.008	-0.034	0.042
	(0.024)	(0.035)	(0.037)
Group: Philanthropic	-0.016	-0.000	0.016
	(0.023)	(0.036)	(0.037)
Constant	0.122***	0.370***	0.508***
	(0.017)	(0.025)	(0.026)
Observations	1,105	1,105	1,105
R2	0.0004	0.001	0.001
F Statistic (df = 2; 1102)	0.242	0.600	0.661
Note:	*p<0.1; **p<0.05; ***p<0.01		

Non-compliance

As with any survey campaign, we had non-compliance issues. Our experiment is only impacted by *one-sided non-compliance*, whereby some participants assigned to a particular treatment did not receive that treatment (*i.e.*, the participant received but did not open the email), but no participants assigned to control received either of the treatments. This is because we sent personalized emails to each participant with email text specific to the group to which we assigned the participant.

With this one-sided non-compliance in mind, we calculate the following treatment effects in the subsequent sections:

- Intent-to-treat effects
- Complier-average causal effects
- Heterogeneous treatment effects

Intent-to-treat effects

The goal of our experiment is to determine the average treatment effect (ATE); however, due to practical challenges such as non-compliance, it is not always possible to obtain an accurate measure of the ATE. We start by estimating the ITT, which is the effect of treatment *assignment* on outcome, ignoring non-compliance. Mathematically, we define ITT as:

$$ITT = E[Y_i(z = 1)] - E[Y_i(z = 0)]$$

where z denotes assignment to one of the experimental groups. After removing the attriters, we perform two regressions to estimate the ITT effect. First, the survey completion variable is regressed on the group assignment variable, using the control group as the reference. We also perform a second regression, which additionally incorporates inferred gender and graduation year covariates. Table 5 summarizes the regression results.

Table 5: Intent-to-treat regression analysis results

ITT Estimate with Attriters Removed

	Dependent variable:	
	Survey Completion	
	(1)	(2)
Group: Direct	-0.061*	-0.062*
	(0.033)	(0.033)
Group: Philanthropic	-0.026	-0.026
	(0.034)	(0.034)
Constant	0.263***	0.248***
	(0.025)	(0.085)
Gender Fixed Effects?	No	Yes
Cohort Fixed Effects?	No	Yes
Observations	979	979
R2	0.004	0.007
Note:	*p<0.1; **p<0.05; ***p<0.01	

Below, we provide a list of our key findings from the ITT analysis (specifically, for the second model that includes cohort and gender fixed effects):

- Direct incentive reduced the likelihood to complete the survey by 6.2% (relative to control), with a 95% confidence interval of [-12.7%, 0.3%].
- Philanthropic incentive reduced the likelihood of completion by 2.6% (relative to control), with a 95% confidence interval of [-9.3%, 4.1%].
- The direct incentive ITT effect is statistically significant at the 10% significance level, whereas the philanthropic incentive ITT effect is not statistically significant.
- The coefficients and standard errors are approximately the same for the two models, which provides validation that block randomization worked as we expected.

Complier-average causal effects

The presence of non-compliers in our experiment mean that our causal estimates from the ITT analysis are diluted relative to the causal estimates we would obtain if only considering compliers. Ultimately, we want to estimate the causal effect of our treatments on those alumni who did comply with our experiment. To do this, we calculate the CACE:

$$CACE = E[Y_i(d=1) - Y_i(d=0) | d_i(1) = 1]$$

Based on our non-interference and exclusion restriction assumptions, we obtain an unbiased estimate of the CACE using the following equations:

$$ITT_D = E[d_i(z=1) - d_i(z=0)]$$

$$CACE = ITT/ITT_D$$

We calculate the CACE by first calculating the take-up rate, ITT_D, which is the proportion of compliers in each group. We then scale the estimated ITT effects by the take-up rate. As shown in Table 6, we observed take-up rates of approximately 60% in both treatment groups, which scales our estimated ITT effects by a factor of roughly 1.7.

Table 6: Complier-average causal effect calculation

CACE = ITT / ITTd

Treatment	ITT	ITTd	CACE
Direct	-0.062	0.621	-0.099
Philanthropic	-0.026	0.587	-0.045

In addition to calculating the CACE using the ratio of ITT to the take-up rate, we also calculate the CACE using instrumental variables (IV) regression. We developed separate IV regression models for each treatment group, where the outcome variable (*i.e.*, survey completion) is regressed on a binary indicator of treatment, using the group assignment as the instrument. Table 7 shows the results of these IV models.

Table 7: Complier-average causal effect calculation using instrumental variables regression

CACE (IVreg) Estimate with Attriters Removed

	Dependent variable:	
	Survey Completion	
	Direct	Philanthropic
	(1)	(2)
Treated	-0.099*	-0.044
	(0.055)	(0.059)
Constant	0.263***	0.263***
	(0.025)	(0.025)
Observations	650	652
Note:	*p<0.1; **p<0.05; ***p<0.01	

Below, we provide a list of our key findings from the CACE analysis:

- Direct incentive reduced the likelihood to complete the survey by 9.9% (relative to control), with a 95% confidence interval of [-20.6%, 0.8%].
- Philanthropic incentive reduced the likelihood of completion by 4.4% (relative to control), with a 95% confidence interval of [-16%, 7.1%].

- The direct incentive CACE is statistically significant at the 10% significance level, whereas the philanthropic incentive CACE is not statistically significant.

Heterogeneous treatment effects

We calculated heterogeneous treatment effects (HTEs) to determine whether response rates differed between specific subgroups. Due to the non-random nature of subgroup analyses, one should not interpret the effect estimates presented in this section as strictly causal effects. Nevertheless, this analysis could point to subgroups that appear more likely to respond to our treatments which may inform follow-up research.

Table 8 shows the results of our HTE analysis for inferred gender subgroups, where we excluded those with an "unknown" gender designation for simplicity.

Table 8: Heterogeneous treatment effect results for inferred gender subgroups

Heterogeneous Treatment Effect

	Dependent variable:	
	Survey Completion	
	(1)	(2)
Male	0.072	0.074
	(0.054)	(0.054)
Group: Direct	-0.025	-0.025
	(0.061)	(0.061)
Group: Philanthropic	0.023	0.024
	(0.063)	(0.064)
Male + Direct	-0.056	-0.057
	(0.073)	(0.074)
Male + Philanthropic	-0.074	-0.074
	(0.075)	(0.076)
Constant	0.216***	0.222**
	(0.044)	(0.093)
Cohort Fixed Effects?	No	Yes
Excl. Gender Unknown?	Yes	Yes
Observations	962	962
R2	0.006	0.008
Note:	*p<0.1; **p<0.05; ***p<0.01	

The difference between the two models shown in Table 8 is that the second model includes graduation year fixed effects. Importantly, none of the interaction terms in either model are statistically significant, indicating that the HTEs we observe on gender are statistically indistinguishable from random chance.

To facilitate interpretation, we provide an explanation of the first, simpler model. Based on Table 8, the resulting model equation is:

$$Y = 0.216 + (0.072 * Male) - (0.025 * Direct) + (0.023 * Philanthropic) - (0.056 * Male * Direct) - (0.074 * Male * Philanthropic) + \epsilon$$

where:

- $Male$ = a dummy variable for gender,
- $Direct$ = a dummy variable for the direct incentive group,
- $Philanthropic$ = a dummy variable for the philanthropic incentive group, and
- ϵ = the error term.

Some conclusions from this model include:

- Females in the control group had a 21.6% response rate, whereas men in the control group had a 28.8% response rate.
- Females in the direct incentive group had a slightly lower response rate than females in the control group, but females in the philanthropic incentive group had a slightly higher response rate.
- Compared to men in the control group, men in both incentive groups had lower response rates.

We also performed an HTE analysis on graduation year subgroups and evaluated both models using F-tests to determine if any of the interaction models are better at explaining the variance of the response rates compared to models without interaction terms (*i.e.*, models that do not include HTEs). Table 9 and Table 10 show the results of these F-tests.

Table 9: F-test on inferred gender interaction terms

```
Analysis of Variance Table

Model 1: completion ~ Gender + Block
Model 2: completion ~ Gender * Block
  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1    958 172.59
2    956 172.40  2   0.19372 0.5371 0.5846
```

Table 10: F-test on graduation year interaction terms

```
Analysis of Variance Table

Model 1: completion ~ factor(Year_Graduation) + Block
Model 2: completion ~ factor(Year_Graduation) * Block
  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1    953 172.45
2    941 169.29 12    3.1601 1.4638 0.1319
```

The p-values for both F-tests are greater than 0.05, indicating that including interaction terms does not statistically significantly improve the models at a 5% significance level, when compared to analogous models that omit the interaction terms.

Please refer to Appendix D for the detailed results of the HTE analysis on graduation year.

Summary of results

Our ITT and CACE estimates, while not statistically significant at the 5% significance level, indicated negative treatment effects for the direct and philanthropic incentive groups. This is contrary to our hypothesis that the incentives would motivate MIDS alumni to respond to the surveys at a higher rate when compared to the control group. This result could be due to a number of reasons, including:

- MIDS is a graduate degree program with a vast majority of people working full-time. Therefore, $25 might not be enough of an incentive to encourage higher response rates.
- Similarly, MIDS alumni may have perceived a donation of only $250 to the BSFC as inadequate to really make a difference.

- The email body text was longer for the direct and philanthropic groups than for the control group. Due to busy schedules and the fact that many MIDS alumni may have read their emails on a hand-held device (which is more arduous to read a longer email on than a computer monitor), having a longer body of text may have effectively discouraged alumni in the treatment groups from taking the time to take on another task and complete the survey.

MIDS is a professional degree program through the School of Information at UC Berkeley. Our experiment was targeted exclusively at MIDS alumni and the "call to action" was to provide feedback to improve the overall MIDS program. With that in mind, we do not expect our results to generalize well. While our results may generalize to near-term future MIDS cohorts, we do not expect the results to generalize to other universities, or even UC Berkeley and other I School programs (due to the specific nature of the MIDS program). The results might also not even generalize to other professional degree programs. This is because we expect that the response rate is influenced by overall experience of the students. We recommend follow-up studies with a mix of students in part-time and full-time degree programs, as well as investigations into the right degree of incentives to encourage higher survey response rates.

Conclusion

In higher education, surveying is an important technique used to gather important feedback and make improvements to the programs. Studies have shown that monetary incentives can improve survey response rates. In our field experiment, we set out to understand the effects of two incentives—a direct financial incentive and a philanthropic incentive—on survey response rates among UC Berkeley Master of Information and Data Science (MIDS) alumni.

Our experiment was not able to find statistically significant positive effects on response rate. On the contrary, we found negative effects from our incentives on survey response rates. The negative effect of direct incentive is significant at the 10% significance level, while the effect of philanthropic incentive is not statistically significant. We also observed some heterogeneous treatment effects, such as female alumni responding more favorably to the philanthropic incentive. But, none of the heterogeneous treatment effects were statistically significant.

The findings from this experiment suggests that monetary incentives are overall not effective at increasing survey response rates among MIDS alumni. We do not expect the results of this experiment to generalize well to other educational programs. However, future work could investigate the generalizability of our results and also analyze the reasons behind monetary incentives being less effective at encouraging higher survey response rates. The findings can contribute to the survey practices in graduate schools and higher education.

References

Armstrong, J. S. (1975). Monetary Incentives in Mail Surveys. *The Public Opinion Quarterly*, *39*(1), 111–116. Retrieved from https://www.jstor.org/stable/2748074

Coppock, A. (2021). *Power Calculator*. Retrieved from https://egap.shinyapps.io/power-app/

Debell, M., Maisel, N., Edwards, B., Amsbary, M., & Meldener, V. (2020). Improving Survey Response Rates with Visible Money. *Journal of Survey Statistics and Methodology*, *8*(5), 821–831. https://doi.org/10.1093/jssam/smz038

Duncan, W. J. (1979). Mail Questionnaires in Survey Research: A Review of Response Inducement Techniques. *Journal of Management*, *5*(1), 39–55. https://doi.org/10.1177/014920637900500103

Gattellari, M., & Ward, J. E. (2001). Will donations to their learned college increase surgeons' participation in surveys?: A randomized trial. *Journal of Clinical Epidemiology*, *54*(6), 645–649. https://doi.org/10.1016/S0895-4356(00)00350-4

How accurate is our tracking tool? (n.d.). Retrieved from https://support.yet-another-mail-merge.com/hc/en-us/articles/212022085-How-accurate-is-our-tracking-tool-

James, J. M., & Bolstein, R. (1992). Large Monetary Incentives and Their Effect on Mail Survey Response Rates. *The Public Opinion Quarterly*, *56*(4), 442–453. Retrieved from https://www.jstor.org/stable/2749201

Kanuk, L., & Berenson, C. (1975). Mail Surveys and Response Rates: A Literature Review. *Journal of Marketing Research*, *12*(4), 440–453. https://doi.org/10.1177/002224377501200408

Moore, R. T., & Schnakenberg, K. (2016). *Package 'blockTools'*. CRAN. Retrieved from https://cran.r-project.org/web/packages/blockTools/blockTools.pdf

Muthukadan, B. (2021). *Selenium with Python*. Retrieved from https://selenium-python.readthedocs.io/

Nesrallah, G., Barnieh, L., Manns, B., Clase, C., Mendelssohn, D., & Guyatt, G. (2014). A charitable donation incentive did not increase physician survey response rates in a randomized trial. *Journal of Clinical Epidemiology*, *67*(4), 482–483. https://doi.org/10.1016/j.jclinepi.2013.11.005

Pedersen, M. J., & Nielsen, C. V. (2016). Improving Survey Response Rates in Online Panels: Effects of Low-Cost Incentives and Cost-Free Text Appeal Interventions. *Social Science Computer Review*, *34*(2), 229–243. https://doi.org/10.1177/0894439314563916

People. (2021). Retrieved from https://www.ischool.berkeley.edu/people

Smith, W. (2008). Does Gender Influence Online Survey Participation? A Record-Linkage Analysis of University Faculty Online Survey Response Behavior. *Online Submission*.

Szelényi, K., Bryant, A. N., & Lindholm, J. A. (2005). What Money Can Buy: Examining the effects of prepaid monetary incentives on survey response rates among college students. *Educational Research and Evaluation*, *11*(4), 385–404. https://doi.org/10.1080/13803610500110174

Warwick, H., Hutyra, C., Politzer, C., Francis, A., Risoli, T. J., Green, C., … Mather, R. C. I. (2019). Small Social Incentives Did Not Improve the Survey Response Rate of Patients Who Underwent Orthopaedic Surgery: A Randomized Trial. *Clinical Orthopaedics and Related Research®*, *477*(7), 1648–1656. https://doi.org/10.1097/CORR.0000000000000732

Appendix A: Emails

This appendix contains the initial email templates that were sent out to the control and two treatment groups of MIDS alumni on July 2, 2021, as well as the reminder email templates that were sent in subsequent weeks. Areas in the templates demarcated by "«. . . »" indicate dynamic fields that are filled in by YAMM as the email gets sent.

Initial emails sent on July 2, 2021

Control group

Hi, «Alumnus Name»:

We are a group of MIDS students conducting research to gather insights on the experience of MIDS alumni. Please take a 2-3 minute survey to provide your opinion of the program. The insights from this survey will help us understand the value of the program as perceived by you, the alumni.

If you complete the survey, we will send you a summary of the results.

The deadline to submit your survey response is Friday, July 23, 2021 by midnight (PDT).

Click here to take the survey and ensure your voice is heard.

Thanks for your time, Devesh, Robert, Thomas, and Joe

Direct incentive group

Hi, «Alumnus Name»:

We are a group of MIDS students conducting research to gather insights on the experience of MIDS alumni. Please take a 2-3 minute survey to provide your opinion of the program. The insights from this survey will help us understand the value of the program as perceived by you, the alumni.

If you complete the survey, we will send you a summary of the results and you will be entered to win an Amazon gift card for $25. Ten respondents will be selected at random to receive a gift card.

The deadline to submit your survey response is Friday, July 23, 2021, by midnight (PDT).

Click here to take the survey and ensure your voice is heard.

Thanks for your time, Devesh, Robert, Thomas, and Joe

Philanthropic incentive group

Hi, «Alumnus Name»:

We are a group of MIDS students conducting research to gather insights on the experience of MIDS alumni. Please take a 2-3 minute survey to provide your opinion of the program. The insights from this survey will help us understand the value of the program as perceived by you, the alumni.

If you complete the survey, we will send you a summary of the results. Additionally, if we achieve a 60% response rate, we will donate $250 to the Berkeley Student Food Collective.

The deadline to submit your survey response is Friday, July 23, 2021 by midnight (PDT).

Click here to take the survey and ensure your voice is heard.

Thanks for your time, Devesh, Robert, Thomas, and Joe

Reminder emails sent on July 9 and July 16, 2021

Control group

Hi, «Alumnus Name»:

We are a group of MIDS students conducting research to gather insights on the experience of MIDS alumni. On July 2nd we sent you a link to a 2-3 minute survey that allows you to provide your opinion of the MIDS program.

If you have not yet completed the alumni survey, this is a reminder that you have «14 or 7» days left to respond.

If you complete the survey, we will send you a summary of the results.

Click here to take the survey and ensure your voice is heard.

Thanks for your time, Devesh, Robert, Thomas, and Joe

Direct incentive group

Hi, «Alumnus Name»:

We are a group of MIDS students conducting research to gather insights on the experience of MIDS alumni. On July 2nd we sent you a link to a 2-3 minute survey that allows you to provide your opinion of the MIDS program.

If you have not yet completed the alumni survey, this is a reminder that you have «14 or 7» days left to respond.

If you complete the survey, we will send you a summary of the results and you will be entered to win an Amazon gift card for $25. Ten respondents will be selected at random to receive a gift card.

Click here to take the survey and ensure your voice is heard.

Thanks for your time, Devesh, Robert, Thomas, and Joe

Philanthropic incentive group

Hi, «Alumnus Name»:

We are a group of MIDS students conducting research to gather insights on the experience of MIDS alumni. On July 2nd we sent you a link to a 2-3 minute survey that allows you to provide your opinion of the MIDS program.

If you have not yet completed the alumni survey, this is a reminder that you have «14 or 7» days left to respond.

If you complete the survey, we will send you a summary of the results. Additionally, if we achieve a 60% response rate, we will donate $250 to the Berkeley Student Food Collective.

Click here to take the survey and ensure your voice is heard.

Thanks for your time, Devesh, Robert, Thomas, and Joe

Reminder emails sent on July 22, 2021

Control group

Hi, «Alumnus Name»:

We are a group of MIDS students conducting research to gather insights on the experience of MIDS alumni. On July 2nd we sent you a link to a 2-3 minute survey that allows you to provide your opinion of the MIDS program.

If you have not yet completed the alumni survey, this is a reminder that you have 1 more day left to respond.

If you complete the survey, we will send you a summary of the results.

Click here to take the survey and ensure your voice is heard.

Thanks for your time, Devesh, Robert, Thomas, and Joe

Direct incentive group

Hi, «Alumnus Name»:

We are a group of MIDS students conducting research to gather insights on the experience of MIDS alumni. On July 2nd we sent you a link to a 2-3 minute survey that allows you to provide your opinion of the MIDS program.

If you have not yet completed the alumni survey, this is a reminder that you have 1 more day left to respond.

If you complete the survey, we will send you a summary of the results and you will be entered to win an Amazon gift card for $25. Ten respondents will be selected at random to receive a gift card.

Click here to take the survey and ensure your voice is heard.

Thanks for your time, Devesh, Robert, Thomas, and Joe

Philanthropic incentive group

Hi, «Alumnus Name»:

We are a group of MIDS students conducting research to gather insights on the experience of MIDS alumni. On July 2nd we sent you a link to a 2-3 minute survey that allows you to provide your opinion of the MIDS program.

If you have not yet completed the alumni survey, this is a reminder that you have 1 more day left to respond.

If you complete the survey, we will send you a summary of the results. Additionally, if we achieve a 60% response rate, we will donate $250 to the Berkeley Student Food Collective.

Click here to take the survey and ensure your voice is heard.

Thanks for your time, Devesh, Robert, Thomas, and Joe

Final reminder emails sent on July 23, 2021

Control group

Hi, «Alumnus Name»:

We are a group of MIDS students conducting research to gather insights on the experience of MIDS alumni. On July 2nd we sent you a link to a 2-3 minute survey that allows you to provide your opinion of the MIDS program.

If you have not yet completed the alumni survey, this is a reminder that you have a few hours left to respond.

If you complete the survey, we will send you a summary of the results.

Click here to take the survey and ensure your voice is heard.

Thanks for your time, Devesh, Robert, Thomas, and Joe

Direct incentive group

Hi, «Alumnus Name»:

We are a group of MIDS students conducting research to gather insights on the experience of MIDS alumni. On July 2nd we sent you a link to a 2-3 minute survey that allows you to provide your opinion of the MIDS program.

If you have not yet completed the alumni survey, this is a reminder that you have a few hours left to respond.

If you complete the survey, we will send you a summary of the results and you will be entered to win an Amazon gift card for $25. Ten respondents will be selected at random to receive a gift card.

Click here to take the survey and ensure your voice is heard.

Thanks for your time, Devesh, Robert, Thomas, and Joe

Philanthropic incentive group

Hi, «Alumnus Name»:

We are a group of MIDS students conducting research to gather insights on the experience of MIDS alumni. On July 2nd we sent you a link to a 2-3 minute survey that allows you to provide your opinion of the MIDS program.

If you have not yet completed the alumni survey, this is a reminder that you have a few hours left to respond.

If you complete the survey, we will send you a summary of the results. Additionally, if we achieve a 60% response rate, we will donate $250 to the Berkeley Student Food Collective.

Click here to take the survey and ensure your voice is heard.

Thanks for your time, Devesh, Robert, Thomas, and Joe

Appendix B: Scraping MIDS Alumni Data

This appendix contains the Python script we used to scrape the name, email, and graduation date for each MIDS alumnus from the UC Berkeley School of Information website's people directory.

```
from collections import namedtuple # Facilitate collect/convert into Pandas
import time # To pause between page loads

from bs4 import BeautifulSoup # To parse the HTML pages
import pandas as pd
# Selenium allows us to open a web browser and act on web pages programmatically
from selenium import webdriver
from selenium.webdriver.common.keys import Keys


# Open a Chrome browser
driver = webdriver.Chrome()

# This file contains a person's username and password
# (each on its own line) for logging into I-School
with open("creds.txt", "r") as file:
    creds = file.readlines()

# Navigate to login page and authenticate
driver.get("https://www.ischool.berkeley.edu/user/login?destination=home")
username = driver.find_element_by_id("edit-name") # Find the username entry
password = driver.find_element_by_id("edit-pass") # Find the password entry
username.send_keys(creds[0].strip()) # Fill in the username
password.send_keys(creds[1].strip()) # Fill in the password
driver.find_element_by_id("edit-submit").click() # Click the login button

# Navigate to first page of alumni, then pause for 5 seconds for loading
start_url = (
  "https://www.ischool.berkeley.edu/people?name=&role=126&degr=MIDS" +
  "&year%5Bvalue%5D%5Byear%5D=&spec=All&emp=&faculty_type=All"
)
driver.get(start_url)
time.sleep(5)

# Initialize namedtuple for data collection
Alumni = namedtuple('Alumni', ["name", "cohort", "email"])
# This list will store `Alumni` namedtuples for each person
all_data = []

# Scrape all data we can for each page
flag = True
page = 1
while flag:
    print(f"Starting page {page}.")
    # Get the HTML from the current browser page
    soup = BeautifulSoup(driver.page_source, 'html5lib')

    # This is a way to get a list of "chunks" of the HTML that
    # correspond to a given alumnus
```

```
    data = [x.parent.text.strip() for x in soup.find_all(
                "div",
                class_='views-field views-field-field-profile-fullname')
    ]

    for person in data:
        name = str(person.split("\n")[0].strip()) # 1st item is always the name
        try:
            # If an entry contains "20," that's the cohort
            cohort = str([x.strip() for x in person.split("\n") if "20" in x][0])
        except IndexError:
            cohort = ""
        try:
            # If an entry contains "@," that's the email
            email = str([x.strip() for x in person.split("\n") if "@" in x][0])
        except IndexError:
            email = ""
        # Append an `Alumni` namedtuple for the person to `all_data`
        all_data.append(Alumni(name, cohort, email))

    try:
        # If there is another page of alumni, click the "next" button
        driver.find_element_by_class_name("pager__item--next").click()
        page += 1
        time.sleep(5) # Pause for not overloading server
    except:
        print("Finished with the last page!")
        flag = False

results_df = pd.DataFrame(all_data)
```

Appendix C: Qualtrics Survey and Results Summary

This appendix contains the actual Qualtrics survey that was sent to each of the MIDS alumni, as well as a graphical summary of the completed survey responses. Figure 7 is the screenshot of the Qualtrics survey that was sent to each of the MIDS alumni. Figure 8 and Figure 9 are summary visualizations of the 229 responses we received from the MIDS alumni.

Berkeley
UNIVERSITY OF CALIFORNIA

Please indicate whether you agree with the following statements. If you are unable to answer the question, please choose "Not applicable".

	Strongly disagree	Somewhat disagree	Neutral	Somewhat Agree	Strongly Agree	Not Applicable
I would recommend MIDS to others interested in pursuing a career in data science.	○	○	○	○	○	○
MIDS offers good value-for-money relative to alternatives, including self-education, other programs, etc.	○	○	○	○	○	○
MIDS has helped me achieve my career objectives.	○	○	○	○	○	○
MIDS provided unique benefits relative to other data science degrees.	○	○	○	○	○	○
I have found MIDS for Life to be valuable post-graduation.	○	○	○	○	○	○
The MIDS network of students, alumni and instructors have helped me achieve my career objectives post-graduation.	○	○	○	○	○	○
The instruction for the program as a whole was valuable, including live sessions, assigned readings, problem sets, projects, Slack, etc.	○	○	○	○	○	○
MIDS career services has helped me achieve my career objectives.	○	○	○	○	○	○

What was your rationale for enrolling in MIDS?

- To build on an existing data science career you were already employed in.
- To enable a career change into data science.
- To add data science to your resume or skill set, for work in a non-data science role.
- Other

Figure 7: Screenshot of Qualtrics survey

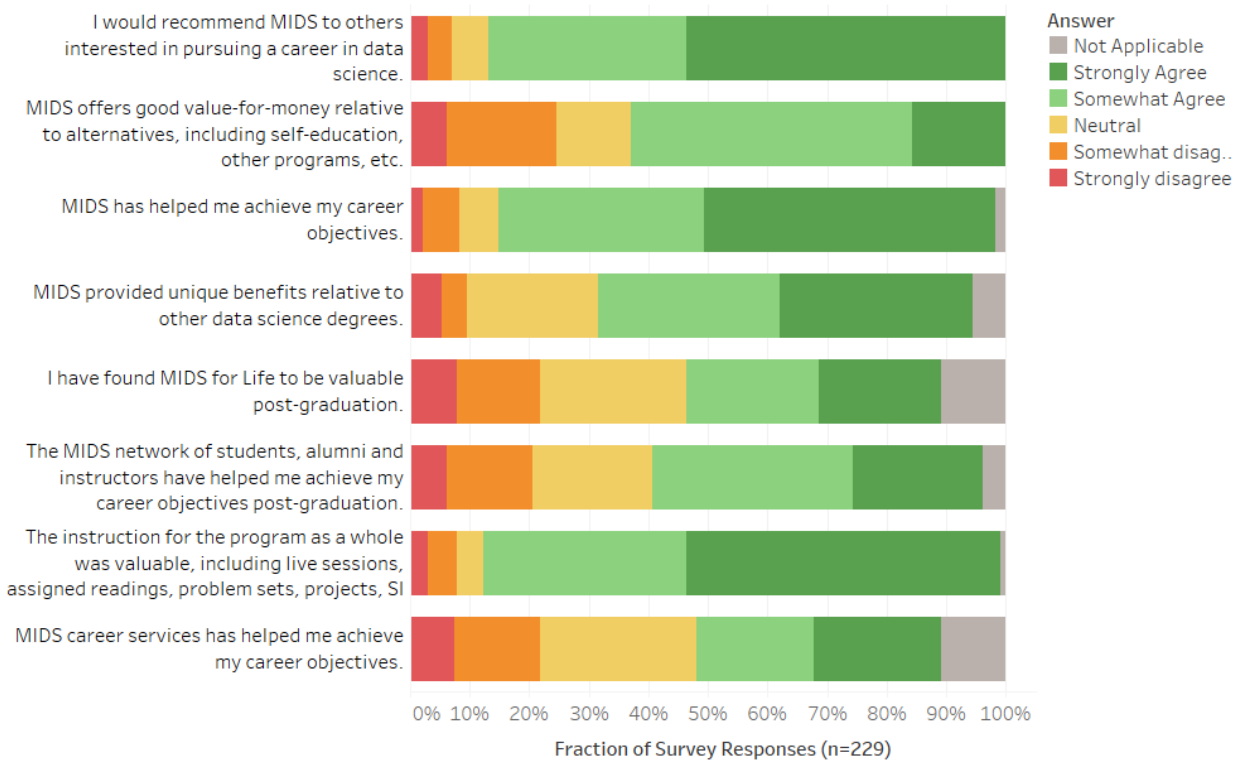

Figure 8: Summary of responses to the first set of statements from the Qualtrics survey

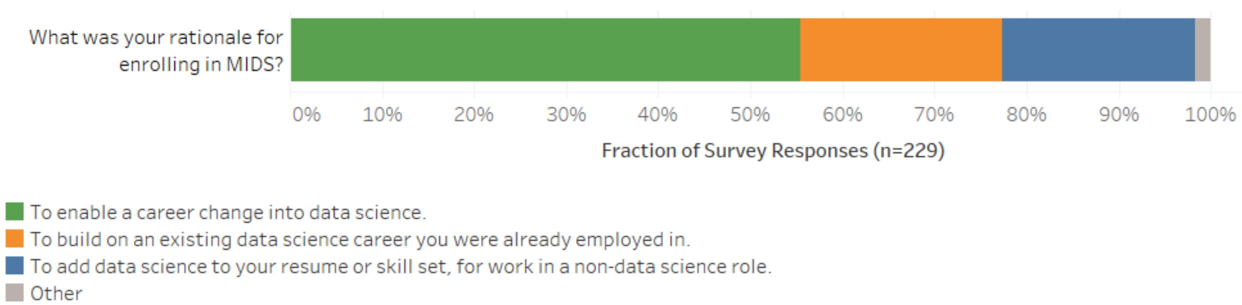

Figure 9: Summary of responses to the last question from the Qualtrics survey

Appendix D: HTE Regression with Interaction Between Treatments and Cohorts

Heterogeneous Treatment Effect

	Dependent variable:	
	Survey Completion	
	(1)	(2)
Graduation: 2016	-0.250	-0.250
	(0.177)	(0.178)
Graduation: 2017	-0.328*	-0.323*
	(0.170)	(0.171)
Graduation: 2018	-0.383**	-0.379**
	(0.165)	(0.166)
Graduation: 2019	-0.326**	-0.324*
	(0.165)	(0.166)
Graduation: 2020	-0.361**	-0.357**
	(0.162)	(0.164)
Graduation: 2021	-0.213	-0.208
	(0.183)	(0.184)
Group: Direct	-0.583***	-0.585***
	(0.155)	(0.157)
Group: Philanthropic	-0.492***	-0.489***
	(0.182)	(0.185)
2016 + Direct	0.341*	0.342*
	(0.184)	(0.185)
2017 + Direct	0.589***	0.590***
	(0.183)	(0.184)
2018 + Direct	0.620***	0.621***
	(0.175)	(0.176)
2019 + Direct	0.558***	0.560***
	(0.172)	(0.174)
2020 + Direct	0.563***	0.565***
	(0.168)	(0.170)
2021 + Direct	0.385*	0.387*
	(0.197)	(0.198)
2016 + Philanthropic	0.402*	0.398*
	(0.215)	(0.217)
2017 + Philanthropic	0.487**	0.483**
	(0.207)	(0.209)
2018 + Philanthropic	0.542***	0.538***
	(0.199)	(0.201)
2019 + Philanthropic	0.471**	0.469**
	(0.197)	(0.199)
2020 + Philanthropic	0.496**	0.494**
	(0.194)	(0.196)
2021 + Philanthropic	0.408*	0.405*
	(0.224)	(0.226)
Constant	0.583***	0.559***
	(0.155)	(0.158)

Gender Fixed Effects?	No	Yes
Excl. Gender Unknown?	Yes	Yes
Observations	962	962
R2	0.024	0.025
Note:	*p<0.1; **p<0.05; ***p<0.01	