



# Survey Response Motivation



By: Alexandra Drossos, Carlos Moreno, Spencer  
Weston, Anand Patel, Madeline Whitlow



# What is the best way to motivate graduate level students to participate in an online survey?

- Previous research has shown that various techniques can be appropriate to motivate individuals to participate in a survey although these methods can differ depending on the individuals of interest.
- Research has shown that monetary incentives can be effective in motivating individuals.
- The goal of our experiment was to determine the most effective method of motivation for graduate students between:
  - No monetary incentives
  - Guaranteed \$5 reward
  - Entry into a \$100 lottery reward
- Our hypotheses
  - \$5 guarantee will increase response rates over control
  - \$100 lottery entry will increase response rates over control
  - \$5 guarantee will increase response rates over \$100 lottery entry

# Experiment Design

- Sent 3 different emails to pool of participants requesting that they respond to the online survey.
- Each participant received an email from Madeline Whitlow (fellow MIDS student) requesting their participation in the survey.
- Participants had 2 weeks to respond during which NO reminders were sent.
- The outcome variable of interest is whether or not the person completed the survey. (0 = no response, 1 = completed survey).

# Treatment 0 (control): No monetary incentive

Response Requested: Complete MIDS Course Time Requirement  
Survey Inbox x



**Madeline Whitlow** <madeline\_94@berkeley.edu>

Sun, Oct 31, 2:14 PM



Dear MIDS Student,

To help better inform MIDS students on which classes they should take, we're launching a survey for *you* to provide perspective on the time requirement that certain program classes take. The survey should take no longer than 10 minutes total to complete. Individual feedback will remain anonymous. However, aggregated data will be made available to current students looking for class selection guidance. We will be collecting data until November 14th, please submit a response by then if you would like to participate. Please follow this link [here](#) to complete. Thanks!

Best,

Madeline Whitlow  
(MIDS Spring '21 Cohort)

# Treatment 1 : Guaranteed \$5 Incentive



# Treatment 2: \$100 Lottery Entry

Response Requested: \$100 Lottery Chance to Complete Course  
Time Survey Inbox x



**Madeline Whitlow** <madeline\_94@berkeley.edu>

Sun, Oct 31, 2:16 PM



Dear MIDS Student,

To help better inform MIDS students on which classes they should take, we're launching a survey for *you* to provide perspective on the time requirement that certain program classes take. The survey should take no longer than 10 minutes total to complete. Individual feedback will remain anonymous. However, aggregated data will be made available to current students looking for class selection guidance. To show our gratitude, upon completion of the survey, you will be entered into a lottery with the chance to win a \$100 Amazon gift card. We will be collecting data until November 14th, please submit a response by then if you would like to participate. Please follow this link [here](#) to complete. Thanks!

Best,  
Madeline Whitlow  
(MIDS Spring '21 Cohort)

# Survey Questions

Of the MIDS courses you've taken please indicate how many hours a week each course took you (this includes asynchronous material, studying, homework and office hours).

	0-10	10-20	20-30	30+
W200	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
W201	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
W203	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
W205	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
W207	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
W209	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
W210	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
W231	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
W233	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
W241	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
W251	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
W261	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
W266	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
W271	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

For the classes that took the most time, what requirement was the most time consuming? \*

- ☐ Homework
- ☐ Async Material
- ☐ Studying
- ☐ Other: \_\_\_\_\_

Which courses did you find the most useful? \*

- ☐ W200: Introduction to Data Science Programming
- ☐ W201: Research Design and Applications for Data and Analysis
- ☐ W203: Statistics for Data Science
- ☐ W205: Fundamentals of Data Engineering
- ☐ W207: Applied Machine Learning
- ☐ W209: Data Visualization
- ☐ W210: Capstone
- ☐ W231: Behind the Data: Humans and Values
- ☐ W233: Privacy Engineering
- ☐ W241: Experiments and Causal Inference
- ☐ W251: Deep Learning in the Cloud and at the Edge
- ☐ W261: Machine Learning at Scale
- ☐ W266: Natural Language Processing with Deep Learning
- ☐ W271: Statistical Methods for Discrete Response, Time Series, and Panel Data

Please explain your answer from above in more detail. Why were the classes you selected useful? \*

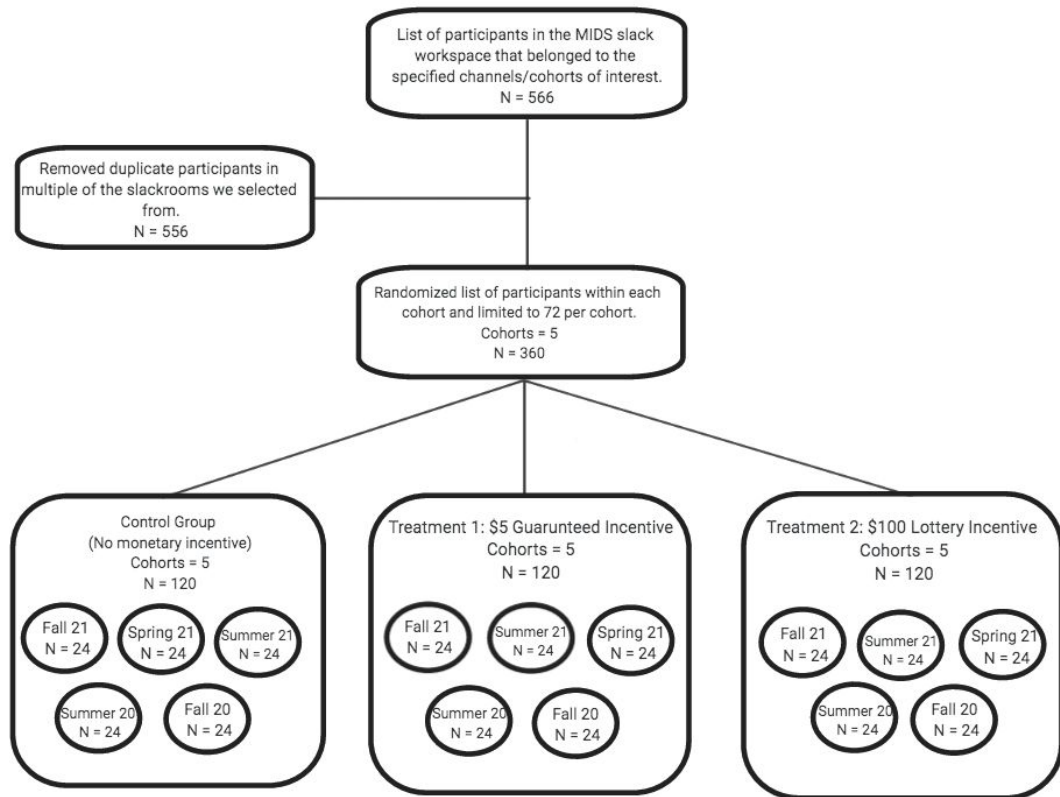
Your answer \_\_\_\_\_

What kind of new course would you like to see in the program?

Your answer \_\_\_\_\_

We did not assess any answers to the survey questions, just whether or not it was completed. We kept the survey brief and relevant to student's interests.

# Randomization



- Identified 5 slack channels to collect participants from (designated by cohort).
- Systematically randomized participant emails in R - identified the selected individuals to participate (not all individuals within a slack channel were chosen)
- Randomly selected participants to be assigned to each treatment group (same size per cohort)



# Experiment Findings

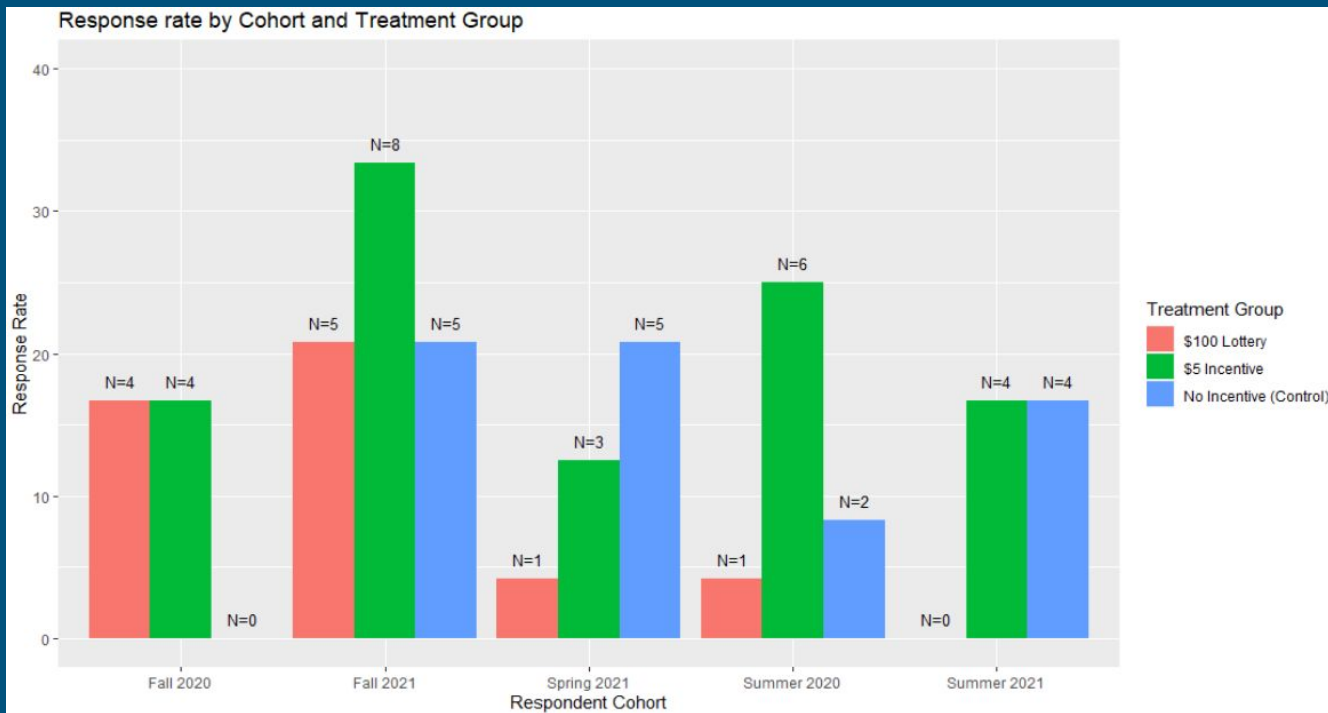
Table 1: Response Rate by Treatment.

Treatment	Group Size	Response (N)	Response (%)
No Incentive (Control)	120	16	13.33
\$5 Incentive	120	25	20.83
\$100 Lottery	120	11	9.17

*Note.* Summary response rate by treatment.

- Outcome Measurement = Response Rate.
- Calculated as the number of respondents that completed the survey divided by the total number of participants in each group.

# Response Rate by Cohort & Treatment Group



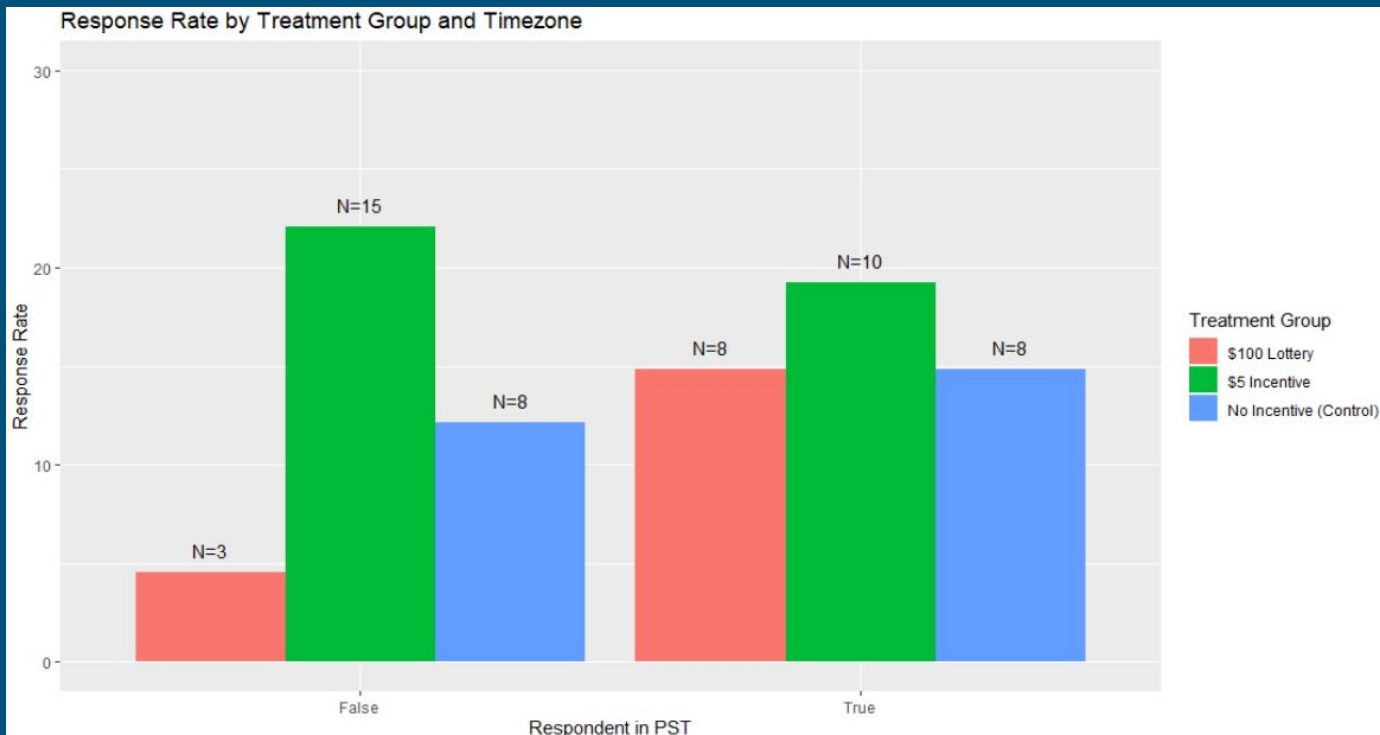
More responses overall in the Fall 2021 cohort.

\$100 Lottery Incentive treatment is the lowest for  $\frac{4}{5}$  cohorts.

\$5 Incentive treatment highest in the Fall 2021 and Summer 2020, tied for highest in two, and 2nd to No Incentive for Spring 2020.

No incentive for survey completion only worse for Fall 2020.

# Response Rate by Treatment Group & Timezone



Equal responses between students in PST and not in PST; 26 each.

Students in PST's response rate did not change much based on incentive.

Students not in PST, the \$5 Incentive yielded a greater response rate than both no incentive and the \$100 lottery assignment.

# A \$5 guaranteed incentive is not significantly better than giving no incentive.

- Not statistical significant  
Average Response Rate increase of 7.5% for “\$5 Incentive” vs. “No Incentive”.
- The treatment effect is equal in all models as covariate for Cohort and Pacific Time (PST) are added.
- None of the covariates helped to increase precision.

Table 1: Linear Regression - No Incentive vs. 5 Per Survey Incentive

	Dependent variable:			
	Response Rate			
	Simple (1)	Cohort Included (2)	PST included (3)	Cohort and PST included (4)
5 Dollar Incentive	0.075 (0.049)	0.075 (0.049)	0.075 (0.049)	0.075 (0.049)
Fall 2020		-0.083 (0.067)		-0.083 (0.068)
Fall 2021		0.104 (0.085)		0.104 (0.085)
Spring 2021		0.000 (0.078)		-0.0002 (0.078)
Summer 2021		0.000 (0.078)		-0.0003 (0.078)
In PST			-0.001 (0.049)	0.002 (0.049)
Baseline	0.133*** (0.031)	0.129** (0.055)	0.134*** (0.037)	0.129** (0.057)
Cohort fixed effects	No	Yes	No	Yes
PST fixed effects	No	No	Yes	Yes
Baseline is:	No Incentive (Control)			
Observations	240	240	240	240
R <sup>2</sup>	0.010	0.035	0.010	0.035
Adjusted R <sup>2</sup>	0.006	0.014	0.002	0.010
Residual Std. Error	0.376 (df = 238)	0.374 (df = 234)	0.377 (df = 237)	0.375 (df = 233)
F Statistic	2.386 (df = 1; 238)	1.694 (df = 5; 234)	1.188 (df = 2; 237)	1.406 (df = 6; 233)

Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Note: Uses Robust Standard Errors.

A \$100 lottery entry incentive is not significantly different than giving no incentive.

- No statistical significant Average Response Rate decrease of -4.2% for “\$100 Lottery” vs. “No Incentive”.
- The treatment effect is equal in all models as covariate for Cohort and Pacific Time (PST) are added.
- None of the covariates helped to increase precision.

Table 2: Linear Regression - No Incentive vs. 100 Lottery Win

	Dependent variable:			
	Response Rate			
	Simple (1)	Cohort Included (2)	PST included (3)	Cohort and PST Included (4)
100 Dollar Lottery Incentive	-0.042 (0.041)	-0.042 (0.041)	-0.042 (0.041)	-0.042 (0.041)
Fall 2020		0.021 (0.055)		0.021 (0.055)
Fall 2021		0.146** (0.070)		0.147** (0.070)
Spring 2021		0.062 (0.060)		0.053 (0.059)
Summer 2021		0.021 (0.054)		0.018 (0.054)
In PST			0.065 (0.042)	0.067 (0.042)
Baseline	0.133*** (0.031)	0.083* (0.043)	0.104*** (0.036)	0.055 (0.049)
Cohort fixed effects	No	Yes	No	Yes
PST fixed effects	No	No	Yes	Yes
Baseline is:	No Incentive (Control)			
Observations	240	240	240	240
R <sup>2</sup>	0.004	0.031	0.015	0.042
Adjusted R <sup>2</sup>	0.0002	0.011	0.006	0.018
Residual Std. Error	0.317 (df = 238)	0.315 (df = 234)	0.316 (df = 237)	0.314 (df = 233)
F Statistic	1.039 (df = 1; 238)	1.521 (df = 5; 234)	1.775 (df = 2; 237)	1.720 (df = 6; 233)

Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Note: Uses Robust Standard Error.

People are more likely to respond to a survey when offered guaranteed compensation than when offered to enter in \$100 lottery.

- Statistical significant Average Response Rate decrease of -11.7% for “\$100 Lottery” vs. “\$5 Incentive” (p-value = 0.011).
- This result holds even when using adjusted  $\alpha$  of 0.017 (Bonferroni correction, 3 tests).
- The treatment effect is equal in all models as covariate for Cohort and Pacific Time (PST) are added.
- None of the covariates helped to increase precision.

Table 3: Linear Regression - 5 Per Survey Incentive vs. 100 Lottery Win

	<i>Dependent variable:</i>			
	Response Rate			
	Simple (1)	Cohort Included (2)	PST included (3)	Cohort and PST Included (4)
100 Dollar Lottery Incentive	-0.117** (0.046)	-0.117** (0.046)	-0.117** (0.046)	-0.117** (0.046)
Fall 2020		0.021 (0.075)		0.022 (0.076)
Fall 2021		0.125 (0.082)		0.122 (0.082)
Spring 2021		-0.062 (0.065)		-0.065 (0.065)
Summer 2021		-0.062 (0.064)		-0.068 (0.063)
In PST			0.037 (0.047)	0.042 (0.046)
Baseline	0.208*** (0.037)	0.204*** (0.061)	0.192*** (0.043)	0.188*** (0.066)
Cohort fixed effects	No	Yes	No	Yes
PST fixed effects	No	No	Yes	Yes
Baseline is:	5 dollar Incentive	5 dollar Incentive	5 dollar Incentive	5 dollar Incentive
Observations	240	240	240	240
R <sup>2</sup>	0.027	0.064	0.029	0.067
Adjusted R <sup>2</sup>	0.023	0.044	0.021	0.043
Residual Std. Error	0.354 (df = 238)	0.350 (df = 234)	0.354 (df = 237)	0.350 (df = 233)
F Statistic	6.526** (df = 1; 238)	3.200*** (df = 5; 234)	3.590** (df = 2; 237)	2.807** (df = 6; 233)

Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Note: Uses Robust Standard Error.

# Heterogeneous Treatment Effect for being in PST

\$100 Lottery worse than \$5 Incentive for outside of PST students:  
95% CI: [-5.8%, -28.6%]

- In of PST students did not benefit from incentives
- Cohorts covariate increased precision, changed estimates

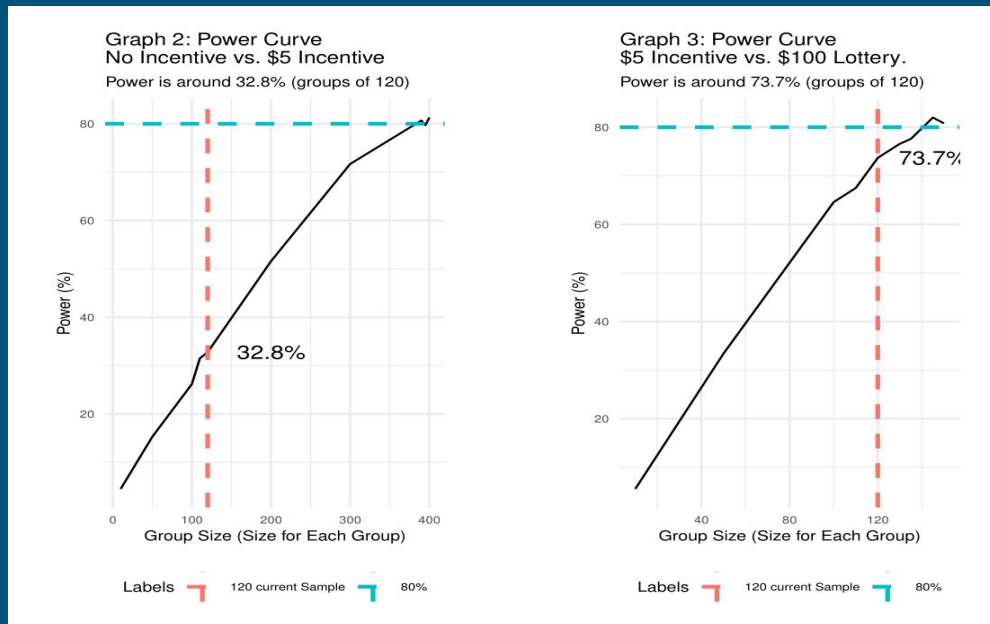
Table 5: Linear Regression - Heterogeneous Treatment Effects for being in PST

	Dependent variable:			
	Response Rate			
	Not PST (1)	PST (2)	Not PST, Cohorts (3)	PST, Cohorts (4)
100 Dollar Lottery Incentive	-0.175*** (0.057)	-0.044 (0.074)	-0.172*** (0.057)	-0.040 (0.074)
No Incentive	-0.099 (0.065)	-0.044 (0.074)	-0.103 (0.065)	-0.031 (0.073)
Fall 2020			-0.015 (0.071)	-0.005 (0.088)
Fall 2021			0.084 (0.082)	0.190* (0.107)
Spring 2021			-0.030 (0.074)	0.033 (0.085)
Summer 2021			-0.061 (0.067)	0.050 (0.088)
Baseline	0.221*** (0.051)	0.192*** (0.056)	0.222*** (0.070)	0.133 (0.085)
Cohort fixed effects	No	No	Yes	Yes
Baseline is:	5 dollar Incentive	5 dollar Incentive	5 dollar Incentive	5 dollar Incentive
Observations	200	160	200	160
R <sup>2</sup>	0.046	0.003	0.067	0.038
Adjusted R <sup>2</sup>	0.036	-0.010	0.038	-0.0001
Residual Std. Error	0.331 (df = 197)	0.372 (df = 157)	0.331 (df = 193)	0.370 (df = 153)
F Statistic	4.723*** (df = 2; 197)	0.248 (df = 2; 157)	2.293** (df = 6; 193)	0.998 (df = 6; 153)
Note:	* p<0.1; ** p<0.05; *** p<0.01 Note: Uses Robust Standard Error.			



# Concerns/Things to think about:

- Statistical power - larger sample size and more informative covariates would be desirable
- Expected value of lottery incentive is undefined for subjects
- Subjects may know Madeline, who sent out the survey request emails
- Matching of individuals from assignment to their responses was facilitated by our privileged access to Berkeley emails





# Conclusion:

## Results Summary

- We fail to reject the null hypotheses of no difference in average response rate between:
  - The \$5 guaranteed incentive and the control group
  - The \$100 lottery incentive and the control group
- We reject the null hypothesis of no difference in average response rate between the \$5 guaranteed incentive and the \$100 lottery incentive

## Suggestions for Future Research:

- Increase sample size and identify more informative covariates to increase power
- Gather more identifying data from survey responses to ensure data integrity
- Guarantee subjects are not affiliated with researchers
- Clarify expected value of lottery incentive

Thank you for listening! Questions?

# Appendix

# In\_PST Covariate Balance Check

Group Distribution by Treatment and In\_PST:

Treatment Group	In PST	Size	Response (N)	Response Rate (%)
No Incentive (Control)	1	54	8	14.81
No Incentive (Control)	0	66	8	12.12
\$5 Incentive	0	68	15	22.06
\$5 Incentive	1	52	10	19.23
\$100 Lottery	1	54	8	14.81
\$100 Lottery	0	66	3	4.55

in\_PST does not vary with treatment group assignment.

## Covariate Balance Check on PST: No Incentive vs. 5 Per Survey Incentive

<i>Dependent variable:</i>	
	in_PST
5 Dollar Incentive	-0.017 (0.065)
Baseline	0.450*** (0.046)
Observations	240
R <sup>2</sup>	0.0003
Adjusted R <sup>2</sup>	-0.004
Residual Std. Error	0.499 (df = 238)
F Statistic	0.067 (df = 1; 238)

Note: \* p<0.1; \*\* p<0.05; \*\*\* p<0.01

# Model Findings - Summary Table

Table 4: Linear Regression - Including All Treatments				
	Dependent variable:			
	Response Rate			
	Simple (1)	Cohort Included (2)	PST Included (3)	Cohort and PST Included (4)
100 Dollar Lottery Incentive	-0.117** (0.046)	-0.117** (0.046)	-0.117** (0.046)	-0.117** (0.046)
No Incentive	-0.075 (0.049)	-0.075 (0.049)	-0.076 (0.049)	-0.076 (0.049)
Fall 2020		-0.014 (0.054)		-0.014 (0.054)
Fall 2021		0.125* (0.064)		0.124* (0.065)
Spring 2021		-0.000 (0.055)		-0.004 (0.055)
Summer 2021		-0.014 (0.054)		-0.019 (0.053)
In PST			0.034 (0.038)	0.037 (0.037)
Baseline	0.208*** (0.037)	0.189*** (0.053)	0.194*** (0.041)	0.175*** (0.056)
Cohort fixed effects	No	Yes	No	Yes
PST fixed effects	No	No	Yes	Yes
Baseline is:	5 dollar Incentive	5 dollar Incentive	5 dollar Incentive	5 dollar Incentive
Observations	360	360	360	360
R <sup>2</sup>	0.019	0.042	0.021	0.044
Adjusted R <sup>2</sup>	0.013	0.025	0.013	0.025
Residual Std. Error	0.350 (df = 357)	0.348 (df = 353)	0.350 (df = 356)	0.348 (df = 352)
F Statistic	3.431** (df = 2; 357)	2.561** (df = 6; 353)	2.565* (df = 3; 356)	2.341** (df = 7; 352)
Note:				
*p<0.1; **p<0.05; ***p<0.01				
Note: Uses Robust Standard Error.				