

ScholarWorks@GSU

Essays on Unintended Consequences: Policy, Incentives and Behavior

Authors	Chatterjee, Chandrayee
Citation	Chatterjee, Chandrayee. Essays on Unintended Consequences: Policy, Incentives and Behavior. 11 Aug. 2020, Georgia State University. https://doi.org/10.57709/18336797 .
DOI	https://doi.org/10.57709/18336797
Download date	2026-04-13 11:30:15
Link to Item	https://hdl.handle.net/20.500.14694/1781

ABSTRACT

ESSAYS ON UNINTENDED CONSEQUENCES:
POLICY, INCENTIVES AND BEHAVIOR

By
Chandrayee Chatterjee
August, 2020

Committee Chair: Dr. James C. Cox

Major Department: Economics

The central theme of this research is understanding the underlying incentive structures of public policies that have behavioral implications, with particular focus on unintended consequences. The three chapters of the dissertation examines three such policies in the domain of labor, health and public economics with the aim of understanding how incentives are created or distorted, leading to unintended consequences.

The first chapter uses a laboratory experiment to test affirmative action variants that differ on the basis of their nature and timing. Using a rank-order tournament, the experiment tests the relative impact of ex-ante (developmental) and ex-post (preferential) affirmative action policies on performance and diversity.

The second chapter uses secondary data and quasi-experimental empirical techniques to study the impact of the Medicaid expansion under the 2014 Affordable Care Act on marital decision making. The Medicaid expansions provided an alternative source of health insurance other than through spousal dependent coverage, thus changing the relative benefit and costs associated with marriage.

The third chapter is a field experiment that uses information on Arizona's state income tax credit for donations to qualifying charities to understand whether the intention to give individuals the freedom to donate to their preferred cause leads to an increase in the overall charitable pie or reallocates funds away from contributions to other causes.

ESSAYS ON UNINTENDED CONSEQUENCES:
POLICY, INCENTIVES AND BEHAVIOR

By

Chandrayee Chatterjee

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy
in the
Andrew Young School of Policy Studies
of
Georgia State University

Georgia State University

2020

Copyright by
Chandrayee Chatterjee
2020

ACCEPTANCE

This dissertation was prepared under the direction of the candidate's Dissertation Committee. It has been approved and accepted by all members of that committee, and it has been accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics in the Andrew Young School of Policy Studies of Georgia State University.

Dissertation Chair: Dr. James C. Cox

Committee: Dr. James H. Marton
Dr. Barry T. Hirsch
Dr. Vjollca Sadiraj
Dr. Glenn W. Harrison

Electronic Version Approved:
Sally Wallace, Dean
Andrew Young School of Policy Studies
Georgia State University
August, 2020

DEDICATION

To
my mother who sacrificed her dreams so I could chase mine,
and
Ayan, who ensures that I never give up on them.

ACKNOWLEDGEMENTS

I am grateful to my committee members, Dr. James Cox, Dr. James Marton, Dr. Barry Hirsch, Dr. Vjollca Sadiraj and Dr. Glenn Harrison for their valuable feedback, guidance and support. To my advisor, Dr. James Cox, thank you for sharing your enthusiasm and knowledge of experiments with me, for always being supportive and inspiring me with your exemplary work ethic. Dr. James Marton, thank you for getting me interested in health economics, challenging me to do better and encouraging me at every step. Dr. Barry Hirsch, I cannot thank you enough for your kindness and encouragement throughout this process and for proving by example that being a great academic and a great human being are not mutually exclusive. Dr. Glenn Harrison, thank you for your valuable suggestions and critiques that have improved my dissertation immensely.

I am grateful to Dr. James Cox and ExCEN for providing funding and travel support for numerous conferences. Thank you, Bess and Kristy for helping us through all the administrative hurdles at different stages of the program. Lucy, thank you for taking such good care of us at ExCEN and for all your love and support.

Sei and Maria, this journey has been excruciatingly hard at times and I cannot imagine going through this without you two. I cherish our friendship, our endless conversations and our best "co-authored" project, our food blog. To Prithvi, Lily and Ano, thank you for all your help navigating this journey and for all the encouragement and support, especially during the job market.

Finally, I could not have made it this far without the most important people in my life, my "shelter from the storm". Ayan, I cannot express in words what your support has meant to me. You have kept me sane, been patient and have often been the only source of happiness in difficult times. My parents, thank you for always pushing me to be ambitious, your unwavering faith in me and for all the sacrifices you have made for me. My grandparents, thank you for your unconditional love and teaching me the importance of being a decent human being above everything.

Contents

1	Introduction	1
2	The “When” and “Why” of Affirmative Action	3
2.1	Introduction	3
2.2	Review of Literature	7
2.2.1	Affirmative Action in the Laboratory	7
2.2.2	Affirmative Action and its Timing	9
2.3	Hypotheses	11
2.4	Experimental Design	14
2.4.1	Timing and Nature of Intervention	15
2.4.2	Magnitude of Intervention	16
2.4.3	Experimental Procedures	18
2.5	Results	19
2.5.1	Descriptive Statistics	19
2.5.2	Effect on Performance of Alternate Affirmative Action Policies	21
2.5.3	Heterogeneous Effect by Ability	26
2.5.4	Effect on Representation	28
2.5.5	Effect of Magnitude of Interventions	29
2.6	Conclusion	31
3	ACA Medicaid Expansion and Marital Decisions	34
3.1	Introduction	34
3.2	Review of Literature	36
3.3	Theoretical Framework	39
3.4	Methodology	43
3.4.1	Data and Sample Selection	43
3.4.2	Identification Strategy	45

3.5	Results	48
3.5.1	Results for Marriage	48
3.5.2	Results for Divorces	50
3.5.3	Robustness Checks	51
3.6	Conclusion	52
4	State Tax Credits and Impact on Charitable Giving	54
4.1	Introduction	54
4.2	Program Description	57
4.3	Experimental Design and Procedure	60
4.3.1	Experimental Design	60
4.3.2	Procedures	64
4.4	Experimental Results	67
4.4.1	Impact of Information on Aggregate Pattern of Giving	69
4.4.2	The Allocation of Funds across Types of Charites	72
4.4.3	Impact of Information on Charity Selection and Allocation	75
4.4.4	Robustness Checks	81
4.5	Discussion	83
5	Conclusion	85
6	Appendices	87
6.1	Appendix A	87
6.2	Appendix B	94
6.3	Appendix C	99
	References	113
	Vita	119

List of Tables

1	Balance of Observables	18
2	Summary Statistics of Performance	20
3	Linear Regression Estimates: Effects on Performance Improvement	23
4	Ordered Logit Regressions: Effects on Performance Improvement	26
5	Linear Regressions: Heterogeneous Effects by Ability Types	27
6	Linear Regressions: Effect on Magnitude of Intervention	30
7	Summary Statistics	45
8	Newly Married Difference-in-Differences Estimates	48
9	Newly Divorced Difference-in-Differences Estimates	50
10	Description of Treatments in the Experiment	62
11	Sample Sizes across Treatments in the Experiment	66
12	Summary Statistics for the Pooled Sample	68
13	Effect of Information Treatment on Donation	70
14	Effect of Information Treatment on Participation (Likelihood of being a Donor)	72
15	Effect of Information Treatment on Fraction of Donation to Qualifying Charities	74
16	Effect of Information Treatment on Selecting a Qualifying/ Non-Qualifying Charity for T4 & T5	77
17	Effect of Information Treatment on Selecting a Specific Number of Qualifying Charities	78
18	Effect of Information Treatment on Fraction of Donation to Qualifying Charities	80
19	Newly Married Event Study Estimates	94
20	Newly Divorced Event Study Estimates	95
21	Newly Married Difference-in-Differences Estimates for Low Income Sample	96

22	Logit Model Estimates for Newly Married	96
23	Difference-in-Differences Estimates with Alternate Definition of Newly Married	97
24	Newly Married Difference-in-Differences Estimates by Dropping Early Expansion States	98
25	Effect of Information Treatment on Making a Donation to a Qualifying Charity	99
26	Effect of Information Treatment on Making a Donation to a Non-Qualifying Charity	100
27	Effect of Information Treatment on Donation	101
28	Effect of Information Treatment on Participation (Likelihood of being a Donor)	102
29	Effect of Information Treatment on Fraction of Donation to Qualifying Charities	103
30	Effect of Information Treatment on Selecting a Qualifying/ Non-Qualifying Charity	104
31	Effect of Information Treatment on Selecting a Specific Number of Qualifying Charities	105
32	Effect of Information Treatment on Fraction of Donation to Qualifying Charities	106
33	Charities in the Experiment, Part 1	110
34	Charities in the Experiment, Part 2	111
35	Charity Lists by Treatment	112

List of Figures

1	Experimental Design	17
2	Representation of each Category in Winners' Pool	28
3	Trends for Newly Married: Full Sample	47
4	Trends for Newly Married: Females	47
5	Trends for Newly Married: Males	47
6	Criteria for Qualifying Charities	58
7	Charitable Giving Trends	59
8	Example Allocation Task with Two Charities	61
9	Overview of Timeline and Procedures in the Experiment	65
10	Screenshot from Experiment	93
11	Information about the CTC	107
12	Example Receipt for Donation made to a Qualifying Charity	108
13	e-Rewards Currency Portal	109

1 Introduction

“There is only one difference between a bad economist and a good one: the bad economist confines himself to the visible effect; the good economist takes into account both the effect that can be seen and those effects that must be foreseen.”

— Frédéric Bastiat

The erroneous conflation of the terms unintended and unanticipated consequences has led to the understanding that unintended consequences may well be anticipated and failure to address it may subvert the understanding of true policy impacts and effective policy-making (De Zwart, 2015). This dissertation studies the unintended consequences of three different policies in the fields of labor, health and charitable giving to understand the underlying economic incentives associated with these policies that drive behavior.

The first chapter examines the impact of variants of affirmative action policies on performance and diversity. The variants are related to the nature and timing at which a policy is implemented in the life cycle of an individual. Thus, the focus is on whether a policy is developmental in nature in that it is implemented before the realization of productivity or it is preferential in nature in that it is implemented at the assignment margin after productivity has been realized. Using a controlled laboratory experiment, which replicates skill acquisition and production stages, I compare ex-ante (developmental) and ex-post (preferential) policies to investigate their impact on performance in a real effort task. The number of winners in the treatments are kept the same by design, such that any observed differences in outcomes across treatments will result from different underlying incentives. Results show that ex-ante interventions lead to greater performance improvement for the disadvantaged category compared to ex-post interventions, without any significant difference in representation. Thus, while both policies fulfill the aim of improved representation for the disadvantaged, the unintended consequence of the quota lies in limiting the performance achievement of the disadvantaged category.

Moving from a stylized laboratory experiment, my second chapter uses observational data to study the impact of the 2014 Affordable Care Act (ACA) Medicaid expansion on marital decision-making. The Medicaid expansions reduced the value of a benefit of marriage, the possibility of dependent health care coverage, by expanding eligibility for Medicaid. I use the American Community Survey data to study the causal impact of the Medicaid expansion on marital decision making. I exploit the variation across time and state Medicaid expansion status to estimate a difference in differences model. Results show that in the target population, there is a reduction in the likelihood of a marital union. Thus, underlying incentives of health insurance reforms can impact seemingly unrelated outcomes such as marital decision making.

The third chapter addresses a recent trend in the use of state income tax credits to reduce the cost of giving to qualifying causes. We use an online framed field experiment with a modified dictator game to test how Arizona's state income tax credit for donations to qualifying charities affects donation decisions. Arizona's tax code offers a suitable setting to study this question. Residents can claim a 100% tax credit for donations made to charities that provide services to vulnerable populations in the state. The study addresses the question whether tax incentives that intend to give individuals the freedom to donate to their preferred cause, leads to an increase in the overall charitable pie or reallocates funds away from contributions to other causes. We find that average giving is unaffected by the information provision and composition of the choice set. The distribution of dollars between qualifying and non-qualifying charities, however, changes dramatically across treatments highlighting a potential unintended consequence of such policies.

The dissertation concludes with broad implications of the research on understanding unintended consequences of policies and accounting for the role of incentives in understanding behavior. Future avenues of research are also discussed.

2 The “When” and “Why” of Affirmative Action ¹

2.1 Introduction

“You do not take a person who, for years, has been hobbled by chains and liberate him, bring him up to the starting line of a race and then say, ‘You are free to compete with all the others,’ and still justly believe that you have been completely fair. Thus, it is not enough just to open the gates of opportunity. All our citizens must have the ability to walk through those gates.”

— President Lyndon Johnson

The Civil Rights Act of 1964, particularly Title VII, stipulated that only the characteristics of workers relevant to employment should be considered for personnel decisions. Thus, workers should have the opportunity to be hired based on merit and not be affected by prejudice. Affirmative action goes beyond the apparent identity-blind nature of equal opportunity laws. It seeks to end employment discrimination to include remedial practices for certain protected groups to correct historical discrimination (Epstein, 1995). Despite the widespread prevalence of affirmative action policies around the world, the debate surrounding the efficacy and fairness of affirmative action policies is as old as the policies themselves. The points of contention have ranged from the fairness and efficiency perspectives of such policies to potential stereotypes and discrimination associated with them.

An overarching question associated with affirmative action has been whether it should be aimed at guaranteeing equality of opportunity or equality of outcome. This central question can be related to the nature of the affirmative action intervention being practiced. A distinction that is fundamental in this regard is provided by Loury (1997) in

¹This paper has benefited from the valuable feedback provided by James Cox, Vjollca Sadiraj, Daniel Kreisman, Danila Serra and Matthias Sutter at various stages of the project. Feedback from participants at the Spring School in Behavioral Economics at UC San Diego 2017 and Southern Economic Association Annual Meetings in 2019 is greatly appreciated. The experiments were funded through the Andrew Young School Dissertation Fellowship.

terms of “developmental” versus “preferential” policies. A preferential policy seeks to achieve target representation by having different selection criteria for different groups. On the other hand, a developmental policy addresses the issue of representation directly by making efforts to improve the skills of the disadvantaged, under-represented category, while maintaining the same criteria for selection across categories.

This distinction is also inherently related to the timing at which such a policy is introduced in an individual’s life cycle. Are such policies introduced ex-ante to production at the developmental margin or ex-post to production at the assignment margin by giving the disadvantaged category preferential access to a limited number of slots (Fryer & Loury, 2013)? For example, a developmental policy which is ex-ante, operates at the skill acquisition stage before production has taken place and bridges the gap at the developmental margin. These may be in the form of subsidized tutoring lessons, summer workshops and curriculum development, fee waivers for taking ACT/SATs and vocational training programs particularly for under-represented minorities (URMs). A preferential ex-post policy is one that bridges the gap at the assignment margin. This could operate as a quota or a lower cut off for the disadvantaged category. The varying nature and timing of such policies while targeting representation can also effect performance of its recipients and non-recipients. Not accounting for the incentive structures inherent in these variants can have the unintended consequence of limiting performance achievement of recipients.

Fryer and Loury (2013) develop a theoretical model addressing the welfare implications of the variants under consideration in this study. The model looks at ex-ante and ex-post policies as those introduced before and after the realization of productivity respectively. The theoretical predictions of their model outline the desirability of one variant over the other depending on whether the agent’s identity is visible (i.e. blind versus sighted policy). However, empirical evidence studying the comparative effects of the variants under consideration have not been explored.

Arcidiacono et al. (2015) study the role of affirmative action in undergraduate edu-

cation for both its beneficiaries and non-beneficiaries. They find that while race-based affirmative action does improve the quality of colleges attended by URMs (i.e. entry to the competitive arena), bans on affirmative action have no effect on URM completion rates (i.e. achievements in the competition itself). Moreover, the “mismatch hypothesis” is operative for major choice as there is consistent evidence showing that among URMs those who persist with selective majors (such as the sciences) have a better academic background than those who switch out of them. This implies that mere entry into the competitive arena does not guarantee improved performance for URMs, which eventually is the determinant of economic achievement.² The intended objective of affirmative action was to improve economic opportunities of URMs as laid down by Executive Order 11246 issued by President Johnson. This was to be achieved by correcting under-representation of minorities by increasing their entry into the competitive arena.³ However, the underlying incentives such policies create could eventually affect the performance of URMs, thereby having unintended consequences on their achievements.

These results raise questions about whether the usual affirmative action policies while fulfilling representation targets fall short of improving actual economic achievements for the disadvantaged minorities which eventually affect economic outcomes. The primary objective of this paper is therefore to investigate the comparative effects of developmental (ex-ante) and preferential (ex-post) policies on the performance and representation of its beneficiaries and non-beneficiaries. The central question is whether affirmative action has the same impact irrespective of when and how such policy is introduced in an individual’s life cycle. A secondary objective is to understand the impact of varying the magnitude of the intervention i.e. whether varying the degree of benefit accorded by

²The “mismatch hypothesis” proposes that affirmative action can negatively affect outcomes of URM students by bringing less-academically-prepared students into more-elite schools, where they have trouble competing with their peers who come from stronger academic backgrounds. Additionally it implies a situation where a student is matched to a school that does not optimize her chance of success (Arcidiacono et al., 2015).

³For example, the Labor Department, under President Nixon, issued Order No.4, authorizing flexible goals to correct the “underutilization” of minorities by federal contractors.

affirmative action has a differential impact on performance.

I use a laboratory experiment which includes a real effort task of answering sample ACT (American College Testing) examination questions that induces a competitive environment through a rank-order tournament. The treatments used in the experiment make it salient that the affirmative action variants are operative at different margins. The ex-ante intervention operates at the developmental margin, prior to production, by giving the disadvantaged category access to skill acquisition in the form of a practice round. The ex-post intervention works at the assignment margin, after productivity has been realized and is preferential in nature through the form of a quota. Thus, the design enables one to compare and identify the impact of these two variants of affirmative action policies on performance. Moreover, additional treatments vary the magnitude of the interventions to study whether the extent of the asymmetry has a differential impact on the performance of individuals.

The results indicate that the nature and timing of intervention does have a differential impact on performance. I find higher performance improvement caused by ex-ante interventions which allow for skill acquisition compared to ex-post quotas, particularly for the disadvantaged category. The results are consistent with prior literature which find that quotas do not hurt performance in the laboratory (Balafoutas and Sutter, 2012; Niederle and Vesterlund 2007, 2010). This study adds to it by showing that they do not improve performance either. Moreover, the ex-ante intervention in addition to improving representation of the disadvantaged category relative to the baseline, also shows no significant difference in the diversity of the winners' pool from the ex-post quota intervention. Varying the magnitude of ex-ante intervention yields a negative impact of over-compensating or under-compensating the disadvantaged category on performance, relative to an intervention that compensates both categories equally. This is consistent with the literature on discouragement effects which associates lower aggregate effort with greater heterogene-

ity between players (Dechenaux et al., 2012).⁴

2.2 Review of Literature

The effect of affirmative action on the employment outcomes of disadvantaged groups was the matter of primary interest in the seminal work by Leonard (1990), who reported affirmative action successfully promotes the representation of minorities and women in employment. Subsequent studies by Holzer and Neumark (1999), Rodgers and Spriggs (1996) have confirmed the same conclusion. Understanding the mechanisms through which affirmative action works is not only important for successful and improved policy targeting, but also for understanding other important outcomes that may be affected by such policies directly, or indirectly, through spillover effects. Experimental studies have often gone beyond impact evaluation of such policies and sought to bridge the gap in the literature by exploring the behavioral impacts of affirmative action policies for a better understanding of underlying mechanisms. In the following section I summarize some of the important outcomes that have been studied in the context of affirmative action in the laboratory.

2.2.1 Affirmative Action in the Laboratory

Some of the major factors that have been studied in this context are statistical discrimination and stereotype threats. The Coate and Loury (1993) model predicts that affirmative action may lead either to erosion or strengthening of negative stereotypes. Empirically, the benign equilibrium which eradicates the negative stereotype has been found to be prevalent (Beaman et al., 2009 ; Kidd et al. 2007). While such results can be expected to reinforce the confidence of recipients of such policies, a valid concern has been whether

⁴An under-compensating or over-compensating intervention retains the asymmetry between two groups of heterogeneous players. A discouragement effect is thus operative in both of these cases since there is a group of players who are stronger than the other group by virtue of the unequal nature of the intervention.

efficiency losses arising from rejecting more able employees outweigh the benefits. Additionally, it has also raised concerns about whether the incentives that such policies create lead to variations in outcomes across beneficiaries and non-beneficiaries.

A few studies have particularly looked at academic performance of beneficiaries and non-beneficiaries of affirmative action. Robles and Krishna (2012) find that affirmative action may have detrimental impact on its beneficiaries by delaying the graduation of the disadvantaged group. Moreover, they find larger gaps in GPA for selective majors reflecting the challenges faced in coursework and curriculum by the disadvantaged categories in a premier institution in India. However, Bagde et al. (2014) find different results with respect to a private engineering college in India where they find that college completion rates and time to graduation did not vary between beneficiaries and non-beneficiaries. Arcidiacono et al. (2015) find that while affirmative action increases the likelihood of graduating by the college quality effect, there is little evidence that affirmative action bans negatively influence URM collegiate attainment rates. Thus the impact of affirmative action on the performance of its beneficiaries has remained mostly inconclusive.

There have been several experimental studies that aim at overcoming the limitations of reduced form impact evaluation of affirmative action to identify the underlying mechanisms through which the policies change outcomes.

Schotter and Weigelt (1992), address the efficiency implications of affirmative action policies through a laboratory experiment with rank-order tournaments. They find that the perceived trade-off between efficiency and equity does not always exist and identify conditions that explain these behavioral implications. Another important metric that has been studied is the effect of affirmative action policies on competition. Studies have shown that affirmative action has a positive impact on encouraging competition and participation by women (Gneezy et al., 2003; Niederle and Vesterlund, 2007, 2010; Sutter and Rützler, 2010; Niederle et al., 2013; Balafoutas and Sutter, 2012). Calsamiglia et al. (2013) further investigates the impact of affirmative action on incentive effects in real-effort tour-

naments between pairs of children from two similar schools who systematically differed in how much training they received ex-ante on the task at hand (which was Sudoku in this case). They too find no significant performance loss and a more equitable pool of winners through the affirmative action intervention. Bracha et al. (2019) find that in the case of gender based affirmative action policies there is heterogeneous effect by ability for women, with the policy lowering (increasing) the performance of high (low) ability women.

2.2.2 Affirmative Action and its Timing

While the previous section discusses studies regarding the efficacy of affirmative action policies compared to scenarios with no intervention, a rarely asked question is whether there are variants of affirmative action that are more desirable than the others. It is important to note all these studies treat affirmative action in the form of quotas only. However, affirmative action and diversity enhancing policies in general have moved beyond the typical quotas. Thus there is need to consider variants of such policies in discussions of affirmative action. In this context, one variation that can be studied is the timing of such affirmative action policies which is also inherently tied to the nature of such policies.

One of the motivations for studying the timing of affirmative action interventions comes from the fact that about half the inequality in the present value of lifetime earnings is due to factors determined by the age of eighteen (Cunha and Heckman, 2007). Thus, ability gaps between advantaged and disadvantaged individuals are formed early in the life cycle of individuals. Moreover, early life interventions for disadvantaged children have much higher economic returns than later life interventions (Heckman et al., 1999). Thus, there is reason to study whether affirmative action interventions have the same impact irrespective of when such policies are introduced in an individual's life cycle.

Fryer and Loury (2013) develop a theoretical model addressing this question and the welfare implications of such variants. The model looks at ex-ante and ex-post policies

as those introduced before and after the realization of productivity respectively. Their model predicts that when an agent's identity is visible, the preferred policy operates at the assignment margin without providing any direct assistance to skill acquisition. However, when the affirmative action is not identity contingent, the ex-ante policy involves giving a universal subsidy to the disadvantaged or imposing a tax on skill acquisition for the advantaged, depending on whether the disadvantaged group is represented better at the development or assignment margin.

However, empirical explorations of the impact of varying the timing of affirmative action interventions is nearly absent. An ideal experiment to study such variations of affirmative action policies would randomly assign the variants across otherwise similar universities (employers) and study the before and after quality/performance of the admitted class (hired workforce) across these universities (employers), to determine the impact of each affirmative action regime. However, this is unlikely to happen and thus makes analysis with observational data implausible. Moreover, even if the data existed, there is possibility of selection bias arising as there is no data on those who are rejected and their counter-factual performance. This makes estimating the true treatment effect of such policies problematic. Moreover, information that is available to employers and admission committees are much more detailed than that available to a researcher. (Fryer, et al., 2008). A recent study by Cassan (2019), exploring affirmative action on the educational attainment of scheduled castes in India, finds that while affirmative action increases educational attainment, the main improvements are concentrated in literacy and secondary schooling while there is little evidence of increases in higher education. Thus, there is a possibility that the timing of affirmative action interventions can affect its success in improving representation but may also lead to unintended consequences on performance and achievements.

Given the lack of empirical findings regarding comparative efficacy of variants of affirmative action interventions and the challenges of carrying out estimation with obser-

vational data, this study circumvents these problems and contributes to the literature by designing a laboratory experiment. It also contributes to our understanding of how individuals respond to policies that bridge asymmetries between an advantaged and disadvantaged category at two different margins.

2.3 Hypotheses

Theoretically, heterogeneity in subjects, competing in tournaments leads to lower effort exerted by the disadvantaged category (Bull et al., 1987). Asymmetric tournaments have been reported to reduce effort and performance in the literature through a discouragement effect (Dechenaux et al., 2012). Weaker (disadvantaged) player gets discouraged in the presence of a stronger (advantaged) player who has a higher probability of winning and thereby finds it unprofitable to exert effort. In my framework, both variants of policy work at removing (or reducing) the asymmetry operating in the form of an “uneven” tournament (O’Keefe et al., 1984), although at different margins.

The ex-ante intervention operates at the skill acquisition stage. It reduces the asymmetry in opportunity by providing the disadvantaged category the opportunity to acquire skills. Even if skill acquisition for both categories is not of the same extent (as in this study and explained in the next section), the ex-ante intervention reduces heterogeneity between players of two categories. The ex-post intervention operates at the assignment margin, by ensuring that despite the heterogeneity in skills between the two categories, the asymmetry is addressed in context of outcomes by a quota. Under a quota, since individuals are essentially competing with players of the same type (i.e. with same skill level in this framework) as themselves, the asymmetry is no longer binding. Thus both variants of affirmative action should be successful in reducing the discouragement effect that arises out of player heterogeneity.

As a result of this and consistent with several laboratory studies such as Schotter and Weigelt (1992), Niederle et al. (2010) and Balafoutas and Sutter (2012) which study affir-

mative action in context of tournaments, I hypothesize the following :

Hypothesis 1: *Affirmative action interventions lead to higher aggregate performance in tournaments relative to the baseline of no intervention. This effect is stronger for disadvantaged subjects.*

In this framework, the performance score can be conceptualized as a positive function of both effort and skill. Recall that ex-ante interventions are developmental in nature and attempts to improve the skill levels in the disadvantaged category. This can be expected to strictly improve the performance of disadvantaged subjects relative to the baseline since it leads to skill development as well as reduces discouragement effect. Considering winning is strictly preferred to losing, advantaged subjects should have no change in performance or may have a performance increase when competing with more competitors undertaking skill acquisition than before i.e. they will have an improvement in score that is greater than or equal to zero.

Ex-post interventions in the form of a quota however bridges the asymmetry without skill development. For advantaged subjects, it can be expected to keep the effort unchanged or increase it relative to the baseline.⁵ Additionally, for disadvantaged subjects, the quota can increase performance relative to the baseline by encouraging effort. However, given that performance is a function of both effort and skills in this framework, a quota by itself will not be successful in raising performance as much as an ex-ante skill acquisition intervention which in addition to encouraging effort also develops skills. Thus, while a quota can improve performance relative to the baseline, it is reasonable to assume that it cannot substitute for skills that additionally improve performance. Thus, I hypothesize the following:

⁵Concerns of a "reverse discouragement" effect whereby advantaged category reduces their performance in the presence of quotas are not supported empirically. In fact for the advantaged category too, competing for fewer spots from within one's type, in the presence of a representation target may raise performance as in Michelitch (2009) who finds that quota ineligible category increase their performance in the presence of quota. While this study is cited in Dechenaux et al. (2012), the study itself is unavailable. Thus, I do not know the exact experiment design employed.

Hypothesis 2: *Skill acquisition, i.e. an ex-ante intervention, leads to a greater performance improvement compared to a quota, i.e. ex-post intervention, for disadvantaged subjects. However, performance improvement of advantaged subjects does not differ across forms of intervention.*

By design, the quota or ex-post intervention has equal representation. Assuming that there is no difference in average ability between the advantaged and disadvantaged subjects, skill acquisition in the form of ex-ante intervention should lead to equal or lower representation of the disadvantaged category.⁶ Thus, I hypothesize the following:

Hypothesis 3: *Skill acquisition i.e. an ex-ante intervention leads to lesser or equal representation of the disadvantaged category compared to a quota i.e. ex-post intervention.*

Another question of interest is to understand whether the size of the intervention matters. An intervention can compensate to equalize opportunities between the two categories i.e. by removing the asymmetry completely. Alternatively, it can compensate to reduce the asymmetry by under-compensating or reverse the asymmetry by over-compensating the disadvantaged category. Given that the unequal compensation preserves the asymmetry in the tournament, we would expect to see an overall lower performance in the unequal compensation case as compared to the equal compensation case. Over-compensating the disadvantaged category may dis-incentivize the advantaged group to exert maximum effort since the policy favors the previously disadvantaged group (positive or reverse discrimination). On the other hand, under-compensation may cause the discouragement effect to work on the disadvantaged category who despite being compensated relative to the baseline are still at a disadvantageous position. Thus, I hypothesize the following:

Hypothesis 4: *Equalizing affirmative action interventions leads to higher aggregate performance in tournaments relative to unequal i.e. over-compensating and under-compensating interventions*

⁶I consider the possibility of an ex-ante intervention that reduces but does not remove the asymmetry, as will be discussed in the design. This results in the possibility of lower representation compared to the advantaged category.

of the same nature.

2.4 Experimental Design

The experiment replicates a skill acquisition and a production stage in order to compare ex-ante and ex-post interventions. Thus, the design has a practice round where subjects acquire skills as the skill acquisition stage. A tournament stage where subjects compete in a rank-order tournament replicates the production stage. Subjects take part in a real effort task which involves answering sample ACT (American College Testing) math questions. This is done to closely replicate an environment where performance can be affected by affirmative action policies such as college admissions.

To test for affirmative action, we require one group of subjects to have an advantage over the other in order to have an advantaged and a disadvantaged category. Subjects are randomly assigned to one of two types in a group of fixed size, in this case groups of six members. In the baseline, subjects of one type (Type A) can take the practice round but those of the other type (Type D) cannot. Thus we will refer to advantaged subjects as Type A and disadvantaged subjects as Type D subjects henceforth.⁷

The practice round has questions which come with feedback, i.e. correct answers are shown to the subjects who take it. Moreover, some questions from the practice round are repeated in the actual tournament which determines the final winners and payoffs. Thus, subjects who take the practice round are at an advantage to those who cannot when participating in the tournament since they have access to the answers for the questions that are being repeated from the practice round. This is how an advantaged category is induced in the laboratory. In the tournament stage, subjects take part in the rank-order tournament where they answer sample ACT questions. The prize spread is \$10 ; winners get a higher reward (\$15) than those who lose (\$5).

⁷While the disadvantaged category is referred to as Type D, in the instructions for the experiment they are referred to as Type B to avoid a possible experimenter demand effect by over-emphasizing the disadvantage.

2.4.1 Timing and Nature of Intervention

Two factors are varied in the experimental design: (1) option to practice for Type D subjects and (2) existence of a quota. The opportunity to practice is in the ex-ante stage, operating at the developmental margin, while the preferential quota relates to the ex-post stage and operates at the assignment margin. Each subject takes part in the baseline with no intervention and one other treatment only.

In the baseline, Type A takes the practice round while Type D does not. Subjects of both types compete in the tournament and the top two scorers are chosen as winners. Thus, the baseline has no affirmative action intervention. In the ex-ante intervention, Type D subjects are now given an option to take the practice round. In the tournament stage the top two scorers are again chosen as winners, thus keeping the competition across types. However, to account for the fact that affirmative action policies do not eliminate the disadvantage altogether, Type D has a smaller number of questions repeated in the tournament stage from the practice round, compared to Type A. A parallel can be drawn with students with college-educated parents who can guide them in the process versus first-generation URM students applying to college who often work with guidance counselors who are assigned to numerous students and thus do not receive guidance as effective as students with college-educated parents. In the ex-post treatment, the productivity gap is assumed to exist in that Type D still does not get to practice. However, now there is a quota such that the top scorer from each type is chosen as a winner, which makes the competition essentially within type. A third treatment looks at the joint implementation of both interventions whereby Type D subjects can practice as well as have a quota i.e. one winner from each type as the winner choice criterion. This treatment enables me to disentangle the effect of the quota from the opportunity to practice on performance as well as understand affirmative actions that bridge the asymmetry at both margins.

The number of winners is held constant across all treatments. It is 2 out of a group of 6 in ex-ante intervention and 1 out of 3 of each type in a group with an ex-post quota. This

enables us to differentiate between the ex-ante and ex-post type of interventions while holding the number of winners constant across treatments. While the ACT questions across all three treatments are the same, the questions in the baseline are different from the treatment since subjects take the baseline before taking the treatment stage. However, the questions are chosen randomly from a pool of ACT questions of the same level of difficulty to maintain the same difficulty level for the baseline and treatments. I can thus compare the difference in the performance score between the treatment and baseline for each subject across treatments.

Additionally, subjects go through a third stage which is used to measure their ability and control for it in the econometric analysis. In this stage, subjects are asked to answer as many questions they can in five minutes. The compensation is piece-rate as they now earn 50 cents for each correct answer and lose 25 cents for each wrong answer. They can also choose to skip a question. This is done to get a true measure of ability and discourage guessing. Subjects complete a questionnaire at the end of the experiment where information about their self reported ACT and SAT math scores and GPAs as well as demographic characteristics.

For each baseline-treatment pair, I reverse the order in keeping with an AB-BA design to account for decision task order effects, if any. Subjects are paid for one of the two parts they participate in, either the baseline or a treatment based on a coin toss after their decisions are entered. Since either part is equally likely to be chosen, subjects have the incentive to do their best in each part and potential wealth effects are avoided. The third stage is introduced after the first two stages are played out so as not to interfere with the main outcomes. Subjects are also paid a show-up fee of \$5.

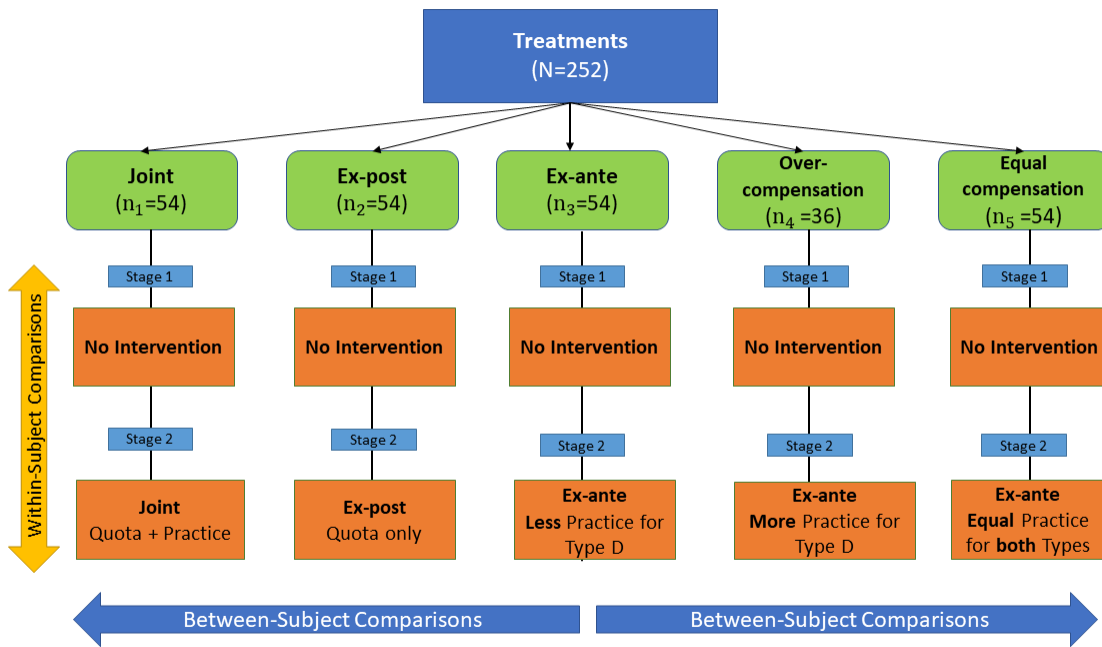
2.4.2 Magnitude of Intervention

To understand whether subjects respond differently to equalizing versus unequal compensation I compare the ex-ante intervention which under compensates individuals with

an equal compensation and over-compensating variations of the ex-ante treatment. In the over-compensating variation, I allow the Type D to have more questions repeated from the practice round than the advantaged category Type A, in the tournament stage. In the equally-compensating variation of the ex-ante treatment, I allow Type D to see the same number of questions repeated from the practice round as Type A in the tournament stage. In all these cases, the number of winners is held constant, so that any treatment effect arises from the magnitude of the intervention. Everything else remains the same as earlier with subjects taking a third task for measuring their ability and are paid for one of the two parts as before.

The experimental treatments are further summarized in Figure 1.

Figure 1: Experimental Design



In a tournament setting, subjects face outcome risk of winning or losing in a tournament. If subjects are risk neutral, subject behavior should not be affected by outcome risk. However, the population at large is typically risk averse (Dohmen et al., 2011). This experiment does not elicit risk preferences of the subjects.

2.4.3 Experimental Procedures

I use Z-Tree version 4.1.6 (Fischbacher, 2007) to operationalize the experiment. The sessions were conducted in Georgia State University's Experimental Economics Laboratory between June to September, 2019. A total of 252 subjects participated over 14 experimental sessions. The Experimental Economics Center (ExCEN) at Georgia State University has a database where undergraduates can voluntarily sign up to participate in economic experiments. These subjects were randomly selected through the ExCEN recruiter system to participate as subjects in this experiment. There was no selection on demographics for this experiment since we induce disadvantage in the laboratory. Balancing tests indicates that the randomization was successful based on key observables as the treatment samples within each condition are similar across key observables (i.e. no significant difference in means across treatments) such as gender, race, age and self-reported SAT/ACT scores as seen in Table 1.

Table 1: Balance of Observables

	Ex-Ante	Ex-Post	Joint
Age	19.98	20.89	20.46
Female	0.61	0.63	0.61
White	0.07	0.06	0.09
Black	0.61	0.69	0.67
Asian	0.19	0.09	0.11
Other	0.09	0.07	0.13
High School GPA	3.67	3.69	3.78
SAT (Quantitative)	565.4	549.3	602.05
ACT(Math)	22.94	21.82	24.68

Subjects read and signed a consent form on reaching the lab. They picked a card from a set of numbered cards (laid out with the numbered side down). They were assigned the computer terminal which matched the number on the card that was drawn. After all participants were seated, the experimenter handed out printed instructions. The experimenter also reviewed the instructions publicly after the subjects finished reading

their instructions and answered any questions that subjects had. Subjects were handed instructions to each part of the experiment after the preceding part ended. The experimenter reviewed instructions with subjects before the beginning of each part, so that subjects were clear about the difference between parts. Instructions were also re-iterated on screen.

At the end of the experiment subjects filled out an un-incentivized questionnaire. Subjects were paid in a separate room and according to their computer number such that no one other than the subject and the experimenter knew their earnings. Subject instructions are included in Appendix A. Figure 10 in Appendix A also provides an example of a question and the interface subjects view.

2.5 Results

To investigate the comparative effect on performance of the variants of affirmative action policies discussed, I use a competitive tournament using ACT math questions. I focus on their effect on the performance scores. More specifically I am interested in the improvement in the performance score in each of the treatments relative to the baseline. Since the underlying incentives and baseline conditions are different across the two categories, the results are shown for the two types separately.

2.5.1 Descriptive Statistics

Descriptive statistics on the effects of the variants of affirmative action under study on performance are provided in Table 2. The means of scores and score improvement are reported for the three main treatments as well as the baseline.

As expected, I find the advantaged, Type A category has a higher mean baseline score (4.70) than the disadvantaged, Type D category (3.69) in the absence of any intervention. A Mann-Whitney test rejects the null hypothesis of no difference in the treatment scores at 1% level of significance ($p < 0.01$). Thus, in the absence of an intervention, I can reason-

Table 2: Summary Statistics of Performance

	Full Sample	Type A	Type D
Baseline Score	4.19	4.70	3.69
Treatments			
Ex-ante:			
Score	5.06	5.33	4.78
Score improvement	1.02	0.67	1.37
Ex-post:			
Score	4.83	5.37	4.30
Score improvement	0.39	0.56	0.22
Joint:			
Score	5.17	5.04	5.30
Score improvement	1.07	0.41	1.74
N	162	81	81

ably conclude that the design is successful in inducing disadvantage in the lab as reflected in the higher performance score of Type A subjects.

In the ex-ante treatments with skill acquisition, there is a significant difference between the mean scores between the two categories (5.33 vs. 4.78 for Type A and D respectively, $p < 0.1$). For ex-post intervention the mean scores are again significantly different across types with Type A mean being 5.37 and Type D mean score being 4.30 ($p < 0.01$). The joint treatment which allows for both skill development and a quota, bridges the gap in performance means as there is no significant difference in the means ($p > 0.1$). However, in each treatment, the average scores are higher than the baseline, driven particularly by the score improvement of disadvantaged subjects. Thus affirmative action leads to higher mean performance overall relative to the baseline for each treatment as hypothesized earlier.

Alternatively, looking at the score improvements across treatments I find that the mean score improvement is higher for the ex-ante and joint treatments compared to the ex-post quota. Table 2 shows that there is no significant difference in the score improve-

ment for the ex-post intervention for the two types ($p > 0.1$). However, for both treatments with skill development i.e. the ex-ante and joint treatments, Type D records a higher mean score improvement than Type A ($p < 0.1$ and $p < 0.01$, respectively).

Comparing score improvement of ex-post quota with the joint treatment shows a significantly higher mean in the latter ($p < 0.05$) implying that bridging the gap in skill acquisition makes a significant contribution to improving performance. The improvement is particularly substantial for Type D. Comparing the means of score improvement for the ex-ante intervention and the joint intervention shows no significant difference ($p > 0.1$) thus implying that quotas do not have a substantial impact on improving performance.

Thus, the descriptive statistics seem to reflect that the ex-ante intervention does better at improving performance for the disadvantaged category as compared to an exogenously imposed quota.

In the following sections, I disentangle the role of quotas versus skill acquisition through practice to formally measure their respective impact on the performance improvement of the subjects.

2.5.2 Effect on Performance of Alternate Affirmative Action Policies

I estimate the causal impact of ex-ante practice versus ex-post quota on the performance improvement across the different types of subjects.

Linear Regressions

I first estimate the impact of alternative affirmative action policies on performance improvement using a series of linear regressions.⁸ Separate regressions are estimated by types since the underlying incentives and conditions vary by types. Model 1 specifies the model without any controls. Model 2 controls for ability by using subjects' scores from the third stage. Model 3 introduces additional controls such as gender and race.

Note that the treatment variable is a categorical variable with three levels for the three

⁸Robust standard errors are estimated and normality assumptions are met in the data.

treatments i.e. ex-post quota only, ex-ante practice only and joint treatment. The joint treatment is considered the base for comparison in my analysis.⁹ The estimation model for each individual is therefore:

$$Y_i = \beta_0 + \beta_1 T_{1i} + \beta_2 T_{2i} + \gamma X_i + \epsilon_i \quad (1)$$

where Y_i is individual i 's score improvement i.e difference in their performance scores in the treatment and the no intervention scenario. T_{1i} and T_{2i} are indicator variables for the ex-post quota only and ex-ante practice only treatments respectively. X_i represents characteristics specific to individual i such as race, gender, and score in ability measure round. The regression compares the score improvement across treatments. Specifically, the joint treatment is set as the baseline for comparison to disentangle the individual effects of quota and practice on score improvement.

The estimates provide the difference in the predicted score improvement for a treatment (ex-ante only or ex-post only) from the joint treatment which is the baseline. Thus, β_1 gives the difference in the mean score improvements between ex-post quota only treatment and joint treatment capturing the impact of the absence of practice on score improvement. Similarly, β_2 gives the difference in the mean score improvements between ex-ante practice only treatment and joint treatment capturing the impact of the absence of quota on score improvement.¹⁰

Table 3 provides the estimates. Type A subjects show a higher improvement when one of the interventions is absent. However, the estimates are not significant. Thus, neither interventions to bridge the asymmetry for Type D subjects have a differential impact on

⁹This is represented by two indicator variables T_{1i} and T_{2i} which equals 1 for ex-post quota only treatment and ex-ante practice only treatment respectively. We do not need to create an indicator variable for the joint treatment because when both T_{1i} and T_{2i} equal to zero, we know that the treatment is joint.

¹⁰For interpreting Table 3 in terms of the regression equation in (1), T_{1i} and T_{2i} may alternatively be interpreted as the absence of ex-ante practice and absence of ex-post quota, respectively. The base joint treatment reflects the situation where both quota and practice are present, that is when both T_{1i} and T_{2i} equals zero. The coefficient β_0 , as represented by the constant term in the regression, provides the estimates of the mean score improvement in the joint treatment.

the performance of the Type A subjects. Moreover, given that the effects, though positive are not statistically different from zero, I cannot conclude if this reflects any evidence of a "reverse" discouragement effect on Type A that may operate in the presence of either intervention. This is summarized in Result 1.

Result 1: *The joint treatment does not lead to a performance improvement for Type A subjects relative to the no intervention case. Neither the removal of ex-ante practice nor the removal of ex-post quota leads to any significant change in performance improvement for Type A compared to the joint treatment.*

Table 3: Linear Regression Estimates: Effects on Performance Improvement

	Model 1		Model 2		Model 3	
	Type A	Type D	Type A	Type D	Type A	Type D
Absence of Ex-ante Practice	0.1482 (0.3980)	-1.5190*** (0.3956)	0.1410 (0.3960)	-1.5285*** (0.3950)	0.0549 (0.3998)	-1.4934*** (0.3645)
Absence of Ex-post Quota	0.2593 (0.3630)	-0.3704 (0.3365)	0.2652 (0.3633)	-0.3804 (0.3301)	0.2103 (0.3888)	-0.3887 (0.3646)
Constant	0.4074 (0.2574)	1.7407*** (0.2537)	0.4715 (0.3273)	2.0304*** (0.3367)	-0.4181 (0.4963)	0.8960*** (0.4550)
Ability	No	No	Yes	Yes	Yes	Yes
Demographics	No	No	No	No	Yes	Yes
N	81	81	81	81	79	80

Notes: Linear regressions. Dependent variable: difference in score between treatment and no intervention baseline. Joint treatment is the base (hence excluded) category set for the categorical treatment variable. Robust standard errors are reported in parentheses. * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level.

When we consider Type D subjects, I find differential effects of the interventions. The difference in the predicted score improvement for the ex-post only treatment from the joint treatment is negative implying that the score improvement was higher under the joint treatment. From Table 3, the coefficients thus imply that taking away the ex-ante opportunity to practice, reduces the mean performance improvement of Type D subjects

by around 1.5 points. Thus, I can conclude that the opportunity to take the practice round has a significant impact on the performance improvement of Type D subjects. 97% of Type D subjects chose to take the practice round when they were provided the opportunity in the skill acquisition interventions. Results show that subjects who take the practice round utilize and benefit from this opportunity, even though it does not completely remove the asymmetry.

However for Type D subjects, the difference in the predicted score improvements for the ex-ante only treatment from the joint treatment, albeit negative, is not statistically significant. Thus, following the similar logic as before, the estimates show that we cannot conclude that taking away the quota has a significant impact on the mean score improvement of Type D subjects. The results are in line with prior experiments studying affirmative action which have shown that quotas do not hurt performance in the lab, (Balafoutas and Sutter, 2012; Calsamiglia et al., 2013) but this experiment adds to it showing it does not improve performance either. These findings are summarized in Result 2.

***Result 2:** The joint treatment leads to a significant performance improvement for Type D subjects relative to the no intervention case. While the removal of ex-ante practice leads to a significant reduction in performance improvement, the removal of the quota has no significant impact on performance improvement compared to the joint treatment.*

The linear regressions give us strong evidence that the affirmative action variants under consideration have differential impact on the performance of the two categories. However, it assumes that the treatment and performance improvement have a linear relationship. I carry out a robustness check with an alternative regression model, the ordered logit model to allow for non-linearity.

Ordered Logit Regressions

For the Ordered Logit specification, I create three categories from the difference in scores between the treatments and baseline, 0,1 and 2 which represents negative score improvement (or score deterioration), no improvement and improvement, respectively.

The dependent variable can thus be thought of as a categorical, ordered variable where a higher category represents a higher level of improvement.

Table 4 below gives the estimates from the Ordered Logit Regressions reporting the coefficients, standard errors and odds ratios. Coefficients represent the effect of absence of practice and absence of quota as obtained by treatment comparison and explained earlier. The negative coefficients are represented by odds ratios of less than one. This implies that the odds of having an improvement in scores relative to no or negative improvement is higher in the joint treatment.¹¹ However, the estimates are not significant for Type A for either interventions similar to the results from the linear regression. Comparing the ex-post only treatment to the joint treatment for Type D implies that the absence of practice leads to lower odds of a performance improvement relative to no or negative improvement. This implies that for Type D, practice has a significant impact in increasing the odds of a score improvement. Comparing ex-ante only treatment with joint treatment, and interpreting the odds ratio similarly shows that removing quotas do not have a significant impact on the odds of a performance improvement.¹²

These results are similar to what we observe in the linear regression where skill acquisition leads to Type D subjects showing significant score improvement compared to quotas.

Thus, removal of asymmetry at the developmental margin by providing disadvantaged subjects the opportunity to practice, even if to a lesser extent than the advantaged category, leads to a considerable improvement in performance of the disadvantaged category, unlike the effect of quotas. Neither of the interventions impact the performance of

¹¹The odds ratio measures the ratio of the odds of a higher order category (e.g. $Y = 2$) to the combined categories below it ($Y = 0$ to $Y = 1$) due to a unit increase in the covariate. Here, since $Y=2$ signifies performance improvement, the odds ratio provides the odds of a performance improvement due to the treatment, relative to the categories below it (no or negative improvement). The joint treatment remains the baseline for comparison.

¹²Interpreted alternatively, this means that Type D subjects in the joint treatments have greater odds of having a performance improvement compared to the ex-post only treatment with a quota implying that practice has a significant impact in performance improvement. However, when comparing ex-ante only treatment to the joint treatment, we find that the odds ratios are not significantly higher for the joint treatment, implying that the quota by itself does not have a significant impact on performance.

Table 4: Ordered Logit Regressions: Effects on Performance Improvement

	Model 1		Model 2		Model 3	
	Type A	Type D	Type A	Type D	Type A	Type D
Absence of Ex-ante Practice	-0.2904	-1.9838***	-0.3214	-2.0077***	-0.5740	-2.2792***
	(0.5244)	(0.6298)	(0.5310)	(0.6303)	(0.6506)	(0.7122)
<i>Odds Ratios</i>	0.7480	0.1376	0.7252	0.1343	0.5633	0.1025
Absence of Ex-post Quota	-0.2904	-0.4000	-0.2791	-0.4260	-0.5663	-0.5001
	(0.5244)	(0.6450)	(0.5765)	(0.6491)	(0.5740)	(0.7007)
<i>Odds Ratios</i>	0.7480	0.6704	0.7565	0.6531	0.5677	0.6065
Ability	No	No	Yes	Yes	Yes	Yes
Demographics	No	No	No	No	Yes	Yes
N	81	81	81	81	79	80

Notes: Ordered Logit Regressions. Joint treatment is the base (hence excluded) category set for the categorical treatment variable. Robust standard errors reported in parentheses. * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level.

Type A thus corroborating Hypothesis 1 and 2 that were laid out earlier.

On analyzing answers to the survey questionnaire, it appears that among those who believed that the interventions affected their performance success, 37% responded that the quota helped their performance compared to only 12% who believed that the skill acquisition helped their performance. This lends itself to concerns that the presence of a quota (which results in one winner from the disadvantaged category to be chosen for certain), may be perceived as a direct success enhancer. Quotas may be viewed more as a substitute for effort than complementing it to meet the selection criteria.

2.5.3 Heterogeneous Effect by Ability

Although I control for ability in the previous specifications, it is worth investigating heterogeneous effects of ability on performance. Affirmative action has been found to have heterogeneous effect on performance by ability among women in the case of gender-based affirmative action. (Bracha et al., 2019).

I use two measures of ability here. First, I use the score from Round 3 of the experiment as the ability measure within the experiment. Additionally I use self-reported high school

GPA. I categorize those above the median ability score or median high school GPAs as high ability and those under, as low ability.¹³

Results are shown in Table 5. Neither the absence of quota nor the absence of practice has any significant impact on performance improvement by ability type. While the estimates for absence of quota is negative implying a possible discouragement effect for high ability types at the assignment margin, they are not significant. The estimates for the absence of practice is positive implying higher score improvement for high ability types than low ability types. However, these estimates are also not significant. Thus I do not find any conclusive evidence of heterogeneous effect by ability for either interventions.

Table 5: Linear Regressions: Heterogeneous Effects by Ability Types

	Stage 3 Score	High School GPA
Absence of Ex-ante Practice	-0.7990** (0.3788)	-0.4792 (0.3821)
Absence of Ex-post Quota	0.1444 (0.3325)	0.6450 (0.3230)
High Ability	0.0468 (0.4290)	0.4251 (0.3869)
Absence of Ex-ante Practice x High Ability	0.3706 (0.6328)	0.3772 (0.5760)
Absence of Ex-post Quota x High Ability	-0.3990 (0.5373)	-0.1468 (0.4895)
Constant	0.1372 (0.3294)	-0.1503 (0.3410)
Demographic Controls	Yes	Yes
N	159	159

Notes: Linear regressions. Dependent variable: difference in score between treatment and baseline no intervention. Robust standard errors are reported in parentheses. * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level.

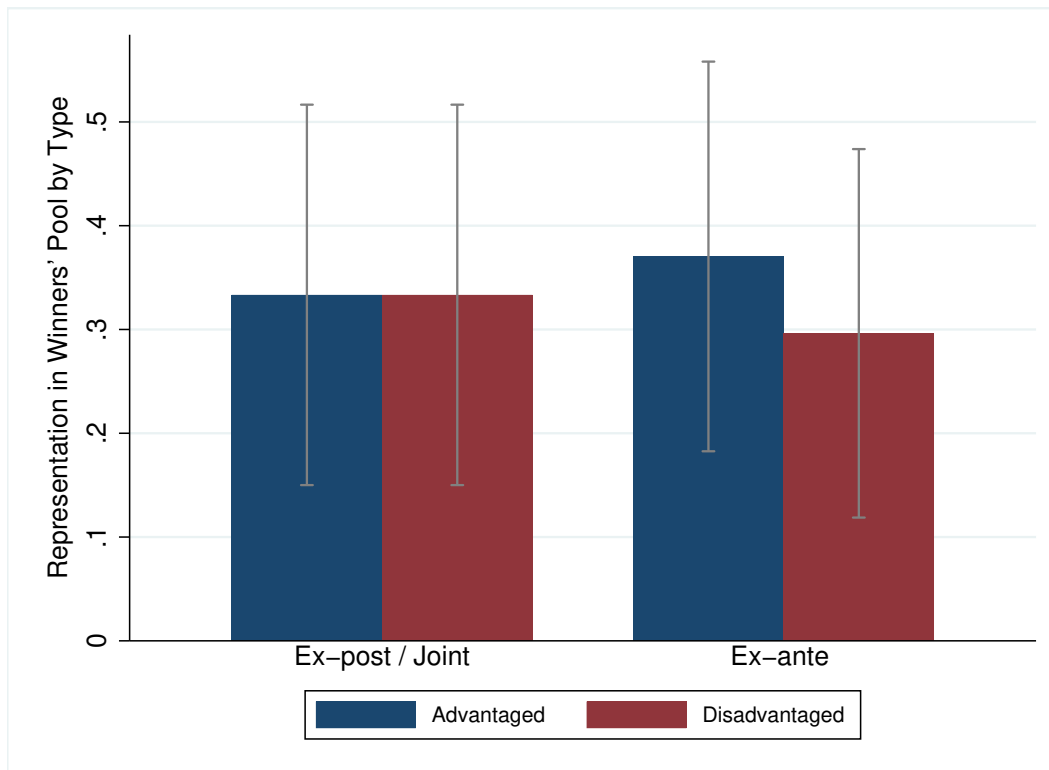
¹³While the questionnaire at the end of the experiment collected information of subjects' ACT and SAT Math scores, very few subjects reported it. This would reduce the sample substantially if considered as a measure of ability in the analysis shown above. Moreover, subjects who responded generally had high ACT and SAT scores, which may bias estimates. Thus, I choose to use the above reported measures as the measure of ability.

2.5.4 Effect on Representation

A secondary outcome of interest is the eventual representation of each category in the winner's pool. Recall, that in the ex-post and joint treatment, which have quotas the representation is designed to be equal. However, in ex-ante treatment, the representation of a particular category can be equal, none or all of the winner's pool. The graph shown below in Figure 2 shows that the representation of both categories in the ex-ante treatment is not significantly different from each other ($p > 0.1$) and from the average representation ratio with quota ($p > 0.1$) as laid out in Hypothesis 3. Thus we have the following result:

***Result 3:** The representation as measured by the distribution of types in the final winners' pool does not vary significantly across treatments.*

Figure 2: Representation of each Category in Winners' Pool



2.5.5 Effect of Magnitude of Interventions

Given that we find ex-ante interventions to be more effective in raising performance, it is worth investigating the optimal magnitude of such an intervention for performance improvement. Therefore, I consider three types of ex-ante interventions. The original ex-ante intervention considered for the analysis is one that is under-compensating in that it bridges the opportunity gap but not completely. Now, I consider interventions that are equal and over-compensating as explained in the design. Our baseline for comparison is the equalizing intervention. Table 6 provides the estimates. The equalizing intervention seems to improve performance in the full sample and for Type D while leading to a marginally significant score deterioration for Type A (as shown by the estimates for the constant).

Interestingly, an over-compensating variation does not significantly improve performance compared to the equal compensation in the sample. The estimates, though insignificant, are negative. There is a possibility that an over-compensating variation leads to Type D subjects not exerting as much effort as in the equal case since they perceive this over-compensation as an advantage in itself. However, in the equal case, since both categories see the same number of questions, there is no asymmetry in opportunities, and thus subjects are incentivized to exert maximum effort. For Type A subjects, the over-compensation for Type D may create an asymmetry which may operate as a "reverse discouragement effect" that results in them reducing effort and having lower score improvements than equalizing case.

The results show that performance improvement with under-compensating ex-ante intervention, that we have considered so far, is lower than an equalizing ex-ante intervention in the full sample as well as for Type D subjects. This is consistent with Hypothesis 4 that the complete removal of asymmetry between two groups leads to better performance overall. An under-compensating intervention preserves the asymmetry and thereby results in a lower performance improvement for the disadvantaged relative to the equal

Table 6: Linear Regressions: Effect on Magnitude of Intervention

	Full Sample	Type A	Type D
Over-compensating Intervention	-0.5836* (0.3245)	-0.3845 (0.3505)	-0.7125 (0.4290)
Under-compensating Intervention	-0.3517 (0.2735)	0.2855 (0.2898)	-0.8516** (0.3669)
Constant	0.4401 (0.5005)	-1.1232* (0.5910)	1.5681** (0.3669)
Ability	Yes	Yes	Yes
Demographic Controls	Yes	Yes	Yes
N	144	72	72

Notes: Linear regressions. Dependent variable: difference in score between treatment and baseline no intervention. The base category for the categorical treatment variable is the equalizing intervention. Robust standard errors are reported in parentheses. * indicates significance at the 10% level; ** indicates significance at the 5% level; *** indicates significance at the 1% level.

compensation which completely removes the asymmetry. The results are significant for the disadvantaged category. For Type A, the under-compensating intervention (although insignificant) shows a lower performance deterioration compared to an equalizing intervention. A possible explanation could be that when the initial advantage that Type A subjects have, by the opportunity to practice, is taken away as in the equalizing intervention, these subjects view this as a reduction in the likelihood of winning (given that the number of winners remain the same) and exert lower effort. The results are summarized in Result 4 below.

Result 4: (i) *Over-compensating interventions lead to lower performance improvement compared to equalizing intervention overall.*

(ii) *The equal ex-ante intervention leads to higher score improvement than the under-compensating ex-ante intervention, particularly for Type D subjects.*

(iii) *Type A subjects show a marginally significant score deterioration with equalizing interventions that does not vary significantly with either the over-compensating or under-compensating interventions.*

2.6 Conclusion

This paper investigates the impact of varying the timing and nature of affirmative action intervention on an important metric that can be affected by such policies, i.e. the performance of its beneficiaries and non-beneficiaries. The results show that ex-post quotas which are preferential in nature, and aim to achieve equality in outcomes, lead to lower performance improvement as compared to ex-ante interventions that are developmental in nature and are implemented to bridge any asymmetry in skills. Moreover, the representation target that is set by the quota is not significantly different from the representation achieved by the ex-ante policy. Thus, the opportunity to develop skills in the practice round leads to an equal representation of both types in the winners pool, even without a quota.

In this framework, it does matter at which margin the asymmetry is bridged for incentivizing performance. A possible explanation is the way benefits of such policies are perceived. Given these results, it appears that while the ex-ante intervention is seen as an opportunity to improve performance, the quota in the ex-post case is perceived as more of a crutch which substitutes for performance. Moreover, among ex-ante interventions, the one which completely removes the asymmetry by providing equal opportunity has higher performance improvement associated with it than under-compensating or over-compensating alternatives. This shows that policies aimed at narrowing the skill acquisition gap between categories can lead to higher performance improvement for the disadvantaged category without hurting representation or the performance of the advantaged category.

Universal testing, whether in the context of admission to gifted education programs (Card and Giuliano, 2016) or as standardized tests for college admissions such as ACT/SAT (Hyman, 2017), has been found to lead to discernible impact in narrowing income, racial and ethnic gaps in representation (Dynarski, 2017). Given the importance of such objective measures of performance, the results from this study raise two important points

worth considering in relation to affirmative action policies. First, affirmative action policies should consider performance as a key target variable in addition to the objective of representation. Second, given that performance is a primary target variable, such policies should aim at understanding the most effective nature and timing of affirmative action interventions that can be implemented to achieve better outcomes. In doing so, such policies should extend the narrative of affirmative action beyond their preferential variants which work in terms of later life compensations and include developmental variants in terms of early life interventions. Investing in policies that target skill development among minorities before selection to higher educational institutions and targeted vocational training before selection to jobs, can potentially reduce the incidence of “mismatch”. For affirmative action to have long term effects on the economic outcomes of minorities, it is important to ensure that minorities can follow through after being selected for a role. Eventually, labor market outcomes would depend on performance and bridging the skill acquisition gap may help facilitate performance improvement and in turn representation, without the explicit use of quotas.

Comparative evaluation of variants of affirmative action policies is thus important for making optimal policy choice. However, it is important to remember that the design does not account for the possibility of differential costs of different treatments in my analysis which can lead to different welfare implications. Moreover, while the experiment is particularly comparable to college admissions based on an objective scoring criteria, the experiment cannot address the issue of taste-based discrimination that may happen in a principal agent framework given that the selection criteria in the tournament is blind. An additional limitation of the design is the inability to study possible effects on behavior by risk-attitudes.

Future research should be aimed at understanding the response of employers to the variants considered here. It is possible that affirmative action which targets skill development can alleviate the problem of statistical discrimination or tokenism that is often

associated with such policies. Moreover, given that the success of such policies depend on the general perception across the board, it would be worth understanding group interactions and prevalence of cooperation, retaliation or sabotage in the context of these variants.

3 ACA Medicaid Expansion and Marital Decisions¹⁴

3.1 Introduction

“But marital relationships, parent-child relationships, decisions to marry and divorce, etc., are also profoundly economic acts. Becker blasted through the Victorian detritus of all that bourgeois romantic ideology to analyze the ways in which marital and reproductive behaviors are fundamentally rooted in a utilitarian economic calculus.”

— Kathleen Geier

Gary Becker (1973), famously argued that positive gains from marriage and negative consequences of getting divorced are what motivates two individuals to stay in a marriage. The legal status of marriage itself bestows a wide range of social, economic and legal benefits for those who choose to participate in a marital contract. Marriage is thus often associated with economic stability and security (Gibson-Davis, Edin, & McLanahan, 2005). One of the broad objectives of the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA) of 1996 was in fact the promotion of marriage in addition to promoting work and reducing childbirth out of wedlock. Promotion of marriage and a two-parent household was considered to be a means to reduce poverty under welfare reform (Hu, 2003). While policies that affect the costs and benefits of marriage can be reasonably expected to affect an individual’s decision to marry, stay married and to divorce, a longstanding question of interest to economists is whether economic incentives created by policies can truly affect marital decision.

One benefit of marriage is access to dependent health insurance coverage through one’s spouse. In the United States, around 152 million non-elderly people are covered

¹⁴I want to thank James Marton for his guidance and valuable feedback at every stage of this project. I would also like to thank participants at the Southern Economic Association Annual Meetings in 2018, the American Society of Health Economists Conference 2019 and the Fall 2019 Conference of the Association of Public Policy Analysis and Management for their valuable feedback. The project also benefited from discussions and presentations at the Georgia State University Health and Labor Economics Summer brown bag seminars.

by employer-provided health insurance according to the Kaiser Family Foundation 2018 Employer Health Benefits Survey and among these a substantial proportion are covered as dependents. Thus, it is reasonable to consider the prospect of obtaining health insurance coverage to be an important consideration for individuals' decisions to marry as well to stay in a marriage. While this may not be important for those who have access to health insurance coverage on their own (employer or government provided), many individuals are not eligible for these forms of coverage and thus resort to dependent coverage. Low-income childless adults who did not qualify for state provided coverage such as Medicaid, did not have ESHI (Employer Sponsored Health Insurance) or could not afford private coverage are people who would especially fit into this vulnerable category. Moreover, dissolution of marriage makes a person vulnerable to the prospect of losing coverage, reiterating the link between health insurance and marital status.

Policies that change the eligibility for different sources of health insurance can alter their relative costs. If such policies provide an alternative source of coverage outside of marriage, or such policies remove any barriers to marriage that had previously made marriage costly, one can expect them to impact marital decision making as well. The Affordable Care Act (ACA) aimed to achieve nearly universal health insurance coverage in the United States through a number of major provisions that took effect in 2014 including the expansion of Medicaid. The purpose of this paper is to identify the causal impact of the ACA Medicaid expansion on the propensity to marry and divorce, using data from the American Community Survey and using a difference-in-differences (DD) identification strategy that exploits variation across time and state Medicaid expansion status.

While popular media has reported anecdotal evidence on the association between the ACA and marital decision making, this paper is the first to my knowledge that causally estimates the impact of the ACA Medicaid expansion on marriage and divorce decisions. By considering divorces, I also try to address the possibility of the incidence of "marriage-lock" whereby individuals decide to stay married to access dependent coverage through

their spouse's health insurance. Since women (25%) are still more likely to be covered as dependents on ESHI policies than men (13%), it is important to stratify the analysis by gender (Peters et al., 2014). Thus, this analysis provides a basis to examine whether such economic incentives affect marriage and family structure.

The results indicate that among low educated (defined as those with a high school degree or lower) non-elderly adults over 26 years of age (i.e. excluding young-adults), the Medicaid expansion results in a statistically significant decrease in the likelihood of being newly married. There is a total reduction of 6.92% which is particularly driven by females who show a reduction of 11.6% in the likelihood of being married. As the theoretical predictions suggest, a plausible explanation is that Medicaid serves as an alternative source of coverage in place of dependent coverage such that people substitute away from dependent coverage (which would entail marriage when they can attain Medicaid when single). The results for divorces also support this explanation although they are more associational in nature. However, my results broadly support the narrative that incentives to gain coverage can affect marital decisions and thereby family structure. The results being larger for women thereby validates the vulnerability of this population in losing insurance due to marriage dissolution.

3.2 Review of Literature

A substantial body of literature has now accumulated which looks at how health insurance within marriage affects labor market decisions. The effects of ESHI on labor supply of married women has shown the spouse's coverage affecting the number of hours women worked negatively (Buchmueller & Valletta, 1999). Wellington & Cobb-Clark (2000) also finds that even husbands who receive health insurance coverage through their wife's employment work fewer hours than husbands who do not, even though the impact is larger for wives covered by their husband's insurance. Abraham & Royalty (2005) argue that having a second earner in the household can improve household health insur-

ance options, access, coverage and generosity, particularly for vulnerable workers (part-time, self-employed and workers in small firms).

One of the earliest empirical works which attempts to quantify the relation between marital disruption and loss of health insurance by Zimmer (2007) finds that marital separation increases the rate of insurance loss by approximately 20 percentage points among wives who are dependent on their husbands' policies and is immediate. Lavelle & Smock (2012) in their seminal work find that approximately 115,000 American women lose private health insurance each year immediately after divorce and that slightly more than one-half become uninsured as a result. However, baseline factors are found to moderate loss of coverage after divorce (i.e. factors such as employment status, source of coverage when married, education, job, poverty status etc.), thus accounting for subgroup heterogeneity in evaluating these results.¹⁵

Subsequent work using the Survey of Income and Program Participation (SIPP) data by Peters et al.(2014) extends the analysis to include both divorced and separated to find that individual based private coverage increases irrespective of gender, after marriage dissolution.¹⁶ However, the decreases in dependent coverage are much larger and offset the increase in private coverage. Children and women show an increase in public coverage around the time of separation or divorce. Sohn (2015) applies a hazard model to married individuals and finds that on average, people who were covered as dependents through their spouse's health plans had lower rates of divorce showing some support for the incidence of "marriage lock". Moreover, not having an alternative source of health insurance outside their current plan as dependents, diminished risks of divorce or sepa-

¹⁵Their estimates are smaller than those provided by Zimmer (2007), primarily because they control for selection bias in the data which a lot of the earlier studies including Zimmer (2007) fail to do. Zimmer (2007) fail to account for the baseline disparity whereby even married women who later divorce, are more likely to be uninsured than women who remain married. Lavelle and Smock (2012) account for the role of selection in their study which may arise from the fact that divorced women may be at a more disadvantageous position relative to married women, to begin with. Thus, they start off by measuring to what extent married women differ from divorced women on pre-divorce characteristics and rates of health insurance coverage. They then apply a multivariate fixed effects model that controls for time-invariant characteristics, subgroup heterogeneities and time heterogeneity.

¹⁶They use the 1996, 2001 and 2004 panels of the SIPP.

ration further.

Health care policies, especially those relating to coverage eligibility can affect marital decisions. Yelowitz (1998) finds that the expansion of Medicaid eligibility in the 1980s and 1990s to married parent families who were not on Aid to Families with Dependent Children (AFDC) program led to an increase in the probability of marriage. A recent paper by Abramowitz (2016) has looked at the impact of the young adult provisions (YAP) of the ACA in 2010 on the probability of marriage given that the YAP opens up a new avenue to access health insurance outside marriage. This paper uses a difference-in-differences analysis and exploits variation across age groups and over time to identify the impact of ACA young adult provision on the likelihood of marriage. The results find a decrease in marriage, cohabitation and spousal health insurance among the treatment group (individuals aged 23-25 years) and an increase in divorces. Heim et al. (2018) use administrative panel data on taxes to also study the impact of the ACA YAP on childbearing and marriage. In addition to finding a decline in marriage among individuals aged 24-25, they particularly find reduced childbearing among young women who are unmarried, those with fewer than two prior children, and those not in post-secondary school.

Barkowski & McLaughlin (2017) studies the influence of U.S. state and federal health insurance coverage mandates on the marriage of young adults. They find that pre-ACA, marriage rates of eligible young adults in states with coverage mandates were lower than ineligible young adults in the same states. This pattern reversed upon the passage of the ACA, with marriage becoming more likely among eligible young adults than ineligible ones. In contrast to Abramowitz (2016), they find that the effect of the ACA on marriage was not uniformly negative, with the complete picture of how the law changed marital behavior depending on the interaction of both the federal and state mandates. Chen (2019) also studies marriage and divorce decisions caused by the introduction of the Massachusetts health care reform of 2006 using the data from ACS and finds a reduction in divorce rates and increase in marriage rates.

This literature review shows that there exists a link between health insurance coverage and marital decisions and family structure. Some recent studies have looked at the impact of specific provisions of the ACA such as the young adult provision on marital decision making. However, these studies do not necessarily generalize to other sections of the population and other provisions of the ACA, one of the most important one being Medicaid expansion.

3.3 Theoretical Framework

The goal of the Patient Protection and Affordable Care Act (ACA) of March 2010 was to achieve nearly universal health insurance coverage in the United States through a “three-legged stool approach” involving a combination of insurance market reforms, mandates, subsidies, health insurance exchanges, and Medicaid expansions (Gruber, 2011). The major components of the ACA took effect in 2014 and these provisions were especially important for low income individuals, women and childless adults. The three legged stool approach addressed the affordability of individual mandates by subsidies and Medicaid expansion. In this section, I discuss how the Medicaid expansion may affect decisions to marry or divorce through a stylized model of marital decision making.

Previously, Medicaid eligibility was typically tied to those with low income among specific groups such as children, pregnant women, elderly and disabled individuals, and some parents, but excluded other low-income adults. With the ACA Medicaid expansion, Medicaid eligibility was no longer categorical. In Medicaid expansion states, Medicaid was made available up to 138 % of FPL. Since the Supreme Court verdict in 2012 allowed states to opt out of the requirement to expand Medicaid, Medicaid eligibility in non-expansion states still remained limited by category with childless adults remaining ineligible in most states and only some parents being eligible. For example, the KFF notes that in 2014, for non-expansion states, the median eligibility limit for parents was 46% of FPL which was about \$9000 for a family of three in 2014. Moreover, in some states, the

eligibility was less than 20% of FPL.

It is important to understand the possible theoretical channels through which the ACA Medicaid expansion can impact marital decisions. Two primary impacts of the Medicaid expansion with respect to health insurance coverage that is relevant to the research question would be a reduction of the cost of coverage and improving access to coverage (through changes in eligibility). This can alter the relative costs of different sources of insurance and thereby the relative benefits and cost of marriage initiation as well as termination. It is important to note here that my question is essentially a cross-sectional question which aims to study how expanding the generosity of the Medicaid program impacts transition in and out of marriage.

I model the decision to marry or remain married to understand how having health insurance coverage through the Medicaid expansion affects marital decision making among low income individuals below 138% of FPL. The model is based on the theory of marriage by Becker (1981) and subsequent adaptation by Chen (2018). The central idea of this model is that people enter into marriages or remain in marriages if the expected utility derived from being single is lower than that of being in a marital union.

Consider a model of marital decision making with identical agents who seek each other in the marriage market, with strictly quasi-linear preferences as follows:

$$U_k = V_k + (H_k - \pi_k), k = M, S \quad (2)$$

where M denotes the state of being married and S denotes the state of being single (or divorced). V denotes the utility gain (measured in dollar units) from marriage (such as having children, income, companionship, love, security etc.) and H is the utility (measured in dollar units) derived from having health insurance coverage. π is the premium or cost of health insurance.

I make the following three simplifying assumptions. First, I assume that individuals

always have health insurance of some form since having coverage is strictly preferred to not having coverage. Second, I assume that the utility derived from health insurance is identical across all plans. Third, that without Medicaid expansion, low income individuals had only ESHI (Employer Sponsored Health Insurance) in marriage through spouses and obtain non-group insurance, when single. ACA Medicaid expansion, expanded eligibility for Medicaid such that single low income individuals now qualified for Medicaid. Moreover, households with income below 138% of FPL also qualified for Medicaid after the expansion. Recall, previously Medicaid eligibility was quite limited in that it covered low-income children, pregnant women, elderly and disabled individuals, and some parents, but excluded other low-income adults. Thus, we have the following:

$$H_S = H_M \quad (3)$$

$$\pi_M = \begin{cases} \pi_{ESHI} & \text{without Medicaid expansion} \\ (\pi_{Medicaid}, \pi_{ESHI}) & \text{with Medicaid expansion} \end{cases} \quad (4)$$

$$\pi_S = \begin{cases} \pi_{Non-group} & \text{without Medicaid expansion} \\ \pi_{Medicaid} & \text{with Medicaid expansion} \end{cases} \quad (5)$$

$$\pi_{Medicaid} < \pi_{ESHI} < \pi_{Non-group} \quad (6)$$

An individual decides whether to enter or leave a marriage. In order to do this they undertake the following optimization :

$$Max[U_M - U_S, 0] \quad (7)$$

According to the model, if $U_M - U_S \geq 0$ the agent prefers to marry or remain married. However, if $U_M - U_S < 0$, the individual prefers to be single.

Substituting (2) and (3) in (7) we have,

$$U_M - U_S = V_M - V_S + \pi_S - \pi_M \quad (8)$$

Now I assume that the marginal agent is indifferent between marriage and staying single such that the utility derived from marriage (independent of health insurance) is equal across the two states i.e. $V_S = V_M$. This implies from equation (8) that,

$$U_M - U_S = \begin{cases} \pi_{Non-group} - \pi_{ESHI} > 0 & , \text{ without Medicaid expansion} \\ \pi_{Medicaid} - (\pi_{Medicaid} \text{ or } \pi_{ESHI}) \leq 0 & , \text{ with Medicaid expansion} \end{cases} \quad (9)$$

According to this model, individuals will choose a marital union without Medicaid expansion since the cost of insurance when single outweighs the cost of insurance when married, leading to a positive net utility from marital union.

However, with Medicaid expansion, if the household income is more than 138% of FPL as a result of marriage such that the $\pi_M = \pi_{ESHI}$ and $U_M - U_S < 0$, an individual will prefer being single to a marital union. Thus, in this case individuals substitute away from using spousal coverage when a cheaper alternative is present in the form of Medicaid. However if the household income is less than 138% of FPL after marriage, $\pi_M = \pi_{Medicaid}$ and $U_M - U_S = 0$. Marriage is then a preferred choice given that it does not increase the cost of insurance. Thus Medicaid expansion can have varying impacts on an individual's incentive to marry post ACA depending on what the post-marriage household income would be. Thus the combined effect of Medicaid expansion is theoretically ambiguous.

Thus, the main takeaway is that the effect of Medicaid expansion on marriage and divorce decisions remains an empirical question since a new source of coverage can increase or decrease marriage incentives relative to the pre-expansion phase. From the review of literature, I find inconclusive evidence about how such coverage expansions affect mar-

riage with some studies finding an increase while others find a decrease. This suggests that effects could potentially be contingent on specific policies and the target population.

3.4 Methodology

This section describes the choice of the sample for my analysis and the methodology implemented for identifying the impact of the ACA Medicaid expansion on marital decision making.

3.4.1 Data and Sample Selection

The primary data source is the American Community Survey (ACS), a nationwide survey administered by the Census Bureau asking detailed questions about population and housing characteristics. The ACS is well suited for this study since it has variables that allow for the measures of marriage initiation and termination (not just the stock of marriage or divorces).

The dependent variable for marriage initiation captures all those who are newly married. Defined specifically, the dependent variable is “whether an individual got married in the calendar year prior to the survey year” constructed from the “year last married” variable from ACS. This is same as the measure used by Abramowitz (2016) in evaluating the impact of the Young Adult Provision of the ACA on marriage. It does not look at the stock of married people in the sample since that would reflect both current and past conditions and instead concentrates on those who initiated marriage recently. For divorces, there is no similarly defined variable as “year last married”. So instead I use the variable measuring whether one got divorced in the last 12 months as an indicator for newly divorced.¹⁷

¹⁷A similar variable for marriage asks whether a person was married in the past 12 months for which data is also available in the ACS. A drawback to using this variable is that, given that the ACS is conducted throughout the year, it is not possible to clearly identify individuals’ date of marriage precisely enough to identify whether they married during the pre-treatment or post-treatment period. However, I use this variable as a robustness check.

The main sample includes low educated (high school degree or lower) adults between the ages of 27 and 64, who got married or remained unmarried in a calendar year previous to the survey year for studying marriage initiation. For divorces, the sample includes low educated (high school degree or lower) adults between the ages of 27 and 64 who got divorced in the past 12 months and all those who remained married, for each calendar year. Recall that the Medicaid expansion is targeted at low-income individuals. Following Kaestner et al. (2015) I limit the sample to low educated individuals as a proxy for low income since education and income are strongly correlated and selecting a sample based on income can lead to biases as Medicaid can affect both marital decisions and income. Young adults of ages 18-26 are removed from the sample as they may be affected by the Young Adult Provision of the ACA.

I consider the time period consisting of ACS survey years 2011 to 2017. Since the outcome variable is lagged, I use the survey years 2011 to 2014 as the pre-treatment period which captures new marriages between 2010 to 2013. The post treatment period includes survey years 2015 to 2017 which captures new marriages between 2014 to 2016 calendar years. The categorization of the expansion states is done in accordance to KFF and Centers for Medicare and Medicaid Services (CMS) which puts the number of states that expanded Medicaid by December 2017 at 31 states and Washington DC (i.e. AK, AZ, AR, CA, CO, CT, DE, DC, HI, IL, IN, IA, KY, LA MD, MA, MI, MN, MT, NH, NJ, NY, ND, NM, NV, OH, OR, PA, RI, VT, WA, and WV). Since the ACA allowed states the option to extend their eligibility prior to 2014, my results should be interpreted as pertaining only to the effects from states that expanded Medicaid between January 2014 to December 2017, and may be underestimating the effect of the total expansion from 2010 to 2017.¹⁸

Table 7 presents the summary statistics for the sample of new marriages used in my analysis i.e. all non-elderly adults above the age of 26, with a high school degree or lower who got married or remained unmarried in the calendar year prior to the survey year.

¹⁸In robustness checks discussed later, I will consider early expanders, and the degree of expansion among these early expanders to mitigate potential confounds in the results arising from early expanders.

Expansion states show a slight increase in the proportion of new marriages post expansion, although the difference is not significant. Non-expansion states show an increase in marriage in post expansion period and the difference is significant at 1% level. There are some differences in the demographic characteristics of the expansion and non-expansion states that I control for in the regression analysis.

Table 7: Summary Statistics

	Expansion States		Non-expansion States	
	Pre-treatment	Post-treatment	Pre-treatment	Post-treatment
Proportion married in calendar year prior to survey year	0.0347 (0.1827)	0.0353 (0.1845)	0.0369 (0.1885)	0.0406 (0.1971)
Percentage Female	0.4766 (0.4995)	0.4651 (0.4988)	0.4875 (0.4998)	0.4749 (0.4993)
Percentage Black	0.0698 (0.2547)	0.1321 (0.3386)	0.1196 (0.3244)	0.2262 (0.4183)
Percentage Hispanic	0.2111 (0.4081)	0.2215 (0.4152)	0.1713 (0.3768)	0.1851 (0.3884)
Percentage Other Race	0.0346 (0.1828)	0.0736 (0.2611)	0.0186 (0.1349)	0.0421 (0.2007)
Percentage Unemployed	0.0957 (0.2941)	0.0640 (0.2448)	0.0908 (0.2873)	0.0577 (0.2332)
N	572,427	421,781	365,587	268,685

Source: American Community Survey-1-year-estimates. Survey years 2011 to 2017

3.4.2 Identification Strategy

My identification strategy addresses the Medicaid expansion of the ACA and its impact on marital decisions by exploiting the variation in time and state Medicaid expansion status as in Simon, Soni & Cawley (2017). Thus, I compare changes in outcomes in the treatment states to the same outcomes in the control states.

The treatment states are the ones that expanded Medicaid to low-income adults, between January 2014 to December 2017. The control consists of the rest of the states which had not yet expanded Medicaid to this population. Formally, I estimate the following

difference in differences(DiD) regression:

$$Y_{ist} = \beta_0 + \beta_1(\text{MEDICAID}_s * \text{POST}_t) + \beta_3 X_{ist} + \lambda_s + \delta_t + \epsilon_{ist} \quad (10)$$

Y_{ist} is the binary marriage outcome for individual i in state s in year t , POST_t is an indicator for whether period t is in the post-treatment period and MEDICAID_s is an indicator for whether state s expanded Medicaid. X_{ist} is a vector of control variables that includes demographic characteristics, family characteristics and household income. λ_s and δ_t are state and time fixed effects respectively. The outcome variable of interest is binary in whether an individual was newly-married or newly-divorced based on the definitions given earlier. I estimate a linear probability models for our binary dependent variable of interest since they are typically known to give reliable estimates of average effects (Angrist & Pischke, 2008). Standard errors are clustered by states.

The identifying assumption of the DD model is that the outcomes would follow similar trends in Medicaid expansion and non-expansion states in the absence of the ACA, conditional on the covariates. Given that this assumption holds, the coefficient β_1 identifies the impact of Medicaid expansion on the outcome.

A state's Medicaid expansion decision is highly political in nature with predominantly Republican states less likely to undertake Medicaid expansions. I test this assumption first informally by graphically comparing the trends. Figures 3, 4 and 5 provides the trends in the proportion of newly married in our sample across treatment and control groups separately. This provides a visual test of the parallel trend assumption for the DD model to be valid for causal interpretation. The trends are predominantly parallel for the pre-treatment years which implies that in the absence of the intervention (i.e. the Medicaid expansion), I would expect to see similar trends in the expansion and non-expansion states. In addition to this I formally test the parallel trend assumption by carrying out an event study analysis. The event study is discussed under the robustness checks.

Figure 3: Trends for Newly Married: Full Sample

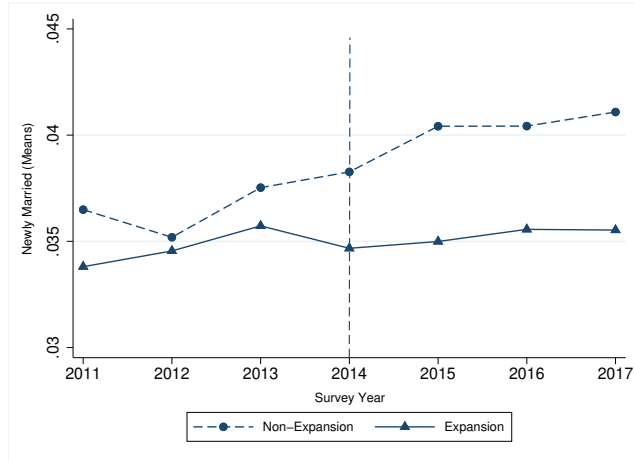


Figure 4: Trends for Newly Married: Females

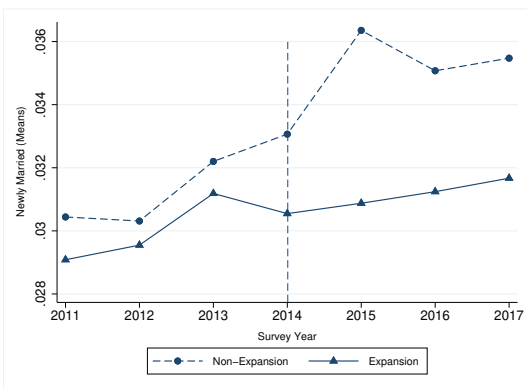
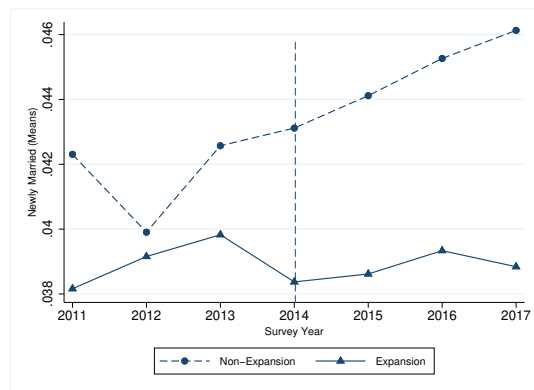


Figure 5: Trends for Newly Married: Males



3.5 Results

This section discusses the main results of the two outcomes of interest, new marriages and new divorces obtained from the DD specification. Additionally, robustness checks and sensitivity analyses are also reported.

3.5.1 Results for Marriage

This section summarizes the impact of the Medicaid expansion on the decision to marry as estimated by the DD model. Table 8 presents the results of the baseline DD model. Column 1 presents results for the entire sample while Columns 2 and 3 presents the results for females and males respectively.

Table 8: Newly Married Difference-in-Differences Estimates

	Full Sample	Female	Male
Medicaid x Post	-0.0024** (0.0011)	-0.0035*** (0.0008)	-0.0015 (0.0015)
Pre-treatment mean	0.0347	0.0301	0.0389
% Change	-6.92%	-11.6%	-3.85%
N	1,628,480	774,843	853,637

Source: American Community Survey

Notes: The sample is restricted to non-elderly adults above 26 years of age with a high school degree or lower who got married or stayed unmarried in the calendar year previous to the survey year. Regression includes demographic and unemployment controls.

*** Significant at 1 per cent level. ** significant at 2 per cent level. * significant at 10 per cent level.

It should be mentioned here that I carry out an event study analysis and the parallel trend assumption holds. Thus the results for marriages can be interpreted causally. I discuss the event study analysis in detail later.

The results show that the expansion of Medicaid eligibility reduced the propensity to marry significantly. In the full sample, the probability of being newly married reduced by 0.2 percentage points which represents a 6.9% decrease in marriage rates among low educated non-elderly adults above 26 years of age compared to before Medicaid expansion.

While there is an overall negative effect of Medicaid expansion on the propensity to marry, the provision may have differential impact on different subsections of the population. For example men and women may have differential access to health insurance or varying degrees of dependence on spousal coverage. Thus, I first stratify the sample by sex and find that Medicaid expansion reduces propensity to marry among both men and women. I find that expansion in Medicaid eligibility leads to a decrease of 0.35 percentage points i.e. 11.6% drop in new marriages in the sub-sample of women relative to their pre-treatment mean. The reduction for men is 0.15 percentage points or 3.9% from their pre-treatment mean and is not statistically significant. However, the estimates for men and women are not statistically significantly different from each other. Thus there is no heterogeneous effect by sex on the propensity of being newly married due to Medicaid expansion ($p=0.14$).

To the best of my knowledge, this is the first study which looks at the impact of ACA Medicaid expansion on the likelihood of new marriages. As a result, I cannot compare my estimates to similar studies. However, the studies on Young Adult Provision of the ACA and its impact on marriage do show a similar substitution away from marriage. Both Abramowitz(2016) and Heim et al. (2018) find a 0.5 percentage point reduction among 23-25 year olds which is a 9.3 and 2.1 percent reduction in the respective samples of these studies.

Thus, the results support the theoretically plausible explanation that in the presence of an alternate source of coverage, individuals substitute away from dependent coverage through marriage, thereby showing a reduction in the propensity to marry. As discussed in the theoretical section, incentives that can cause individuals to substitute away from

marriage as a source of health insurance coverage, can also affect other marital outcomes such as divorce. An alternate source of coverage can also reduce the cost of divorce.

3.5.2 Results for Divorces

Table 9 presents the results of new divorces. The results indicate that the Medicaid expansion is associated with an increase in the likelihood of having divorced in the past 12 months. Although statistically insignificant, there is an increase in divorces in the sample by 0.11 percentage points i.e. a 5.95% increase. The sub-sample for women shows an increase of 0.16 percentage points or a 8.74% increase. Men show an increase in the likelihood of new divorces by 0.07 percentage points. However, there is no heterogeneous effect on the likelihood of divorces by sex as the estimates are not statistically significantly different from each other ($p=0.42$).

Table 9: Newly Divorced Difference-in-Differences Estimates

	Full Sample	Female	Male
Medicaid x Post	0.0011 (0.0007)	0.0016* (0.0010)	0.0007 (0.0010)
Pre-treatment mean	0.0185	0.0182	0.0188
% Change	5.95%	8.74%	3.72%
N	2,280,159	1,106,657	1,173,502

Source: American Community Survey

Notes: The sample is restricted to non-elderly adults above 26 years of age with a high school degree or lower who got divorced in the past 12 months and all those who remained married, for each calendar year. Regression includes demographic and unemployment controls.

*** Significant at 1 per cent level. ** significant at 2 per cent level. * significant at 10 per cent level.

However, these results are associational in nature since the parallel trends assumption does not hold for divorces given our model. In the event study presented in Table 20, I

reject the null hypothesis that all pre-2014 interaction coefficients are jointly equal to zero. Thus, I cannot attribute increase in divorces to change in Medicaid eligibility alone since the underlying assumption of the DiD model is violated.

3.5.3 Robustness Checks

I test the sensitivity of the results to validity of the model assumptions, modifications of the sample, definition of the dependent variable and model specification. Robustness checks are carried out for new marriages since the model is not valid for causal interpretation for divorces. The results are provided in Appendix B.

First, I test for the parallel trend assumption more formally with an event study. The results are provided in Table 19 for new marriages. I jointly test the null hypothesis that all pre-2014 interaction terms are equal to zero using an F-test. I cannot reject the hypothesis that all pre-2014 interaction coefficients are jointly equal to zero for the full sample and the sample restricted to women. However, I can reject the null for males since one of the pre-2014 coefficients is significant. These results give us confidence in a causal interpretation of the regression estimates.

Second, in Table 21, I define the eligibility of non-elderly adults above 26 years of age using low income (i.e. less than 138% of FPL) instead of low education as I have used in the baseline model. The estimates are qualitatively similar compared to those from Table 8, although they are statistically insignificant for the full sample. Women still show a decrease in new marriages by 8.2% although it is smaller than the estimate in the low education sample.

Third, I estimate a logit model instead of the baseline linear probability model for the binary outcome variable. The odds-ratio of an interaction term is problematic for the interpretation of the size of the treatment effect due to non-linearity of the logit model (Karaca-Mandic et al.,2012). However, Puhani (2012) has shown that the coefficient of the interaction term in a non-linear DiD model represents the treatment effect under obser-

vation as in the linear DiD model. Thus, I report the estimates in Table 22. The statistical significance and magnitude of the marginal effects are similar to the baseline results using the linear probability model.

As a fourth sensitivity check, I use an alternate measure of new marriages. I use the variable which records whether a person got married in the last 12 months from the time they were surveyed. As discussed before this is a less accurate measure than the one used in my baseline model, but can be reasonably expected to give qualitatively similar results. Table 23 shows, the results are qualitatively same with women driving the decrease in new marriages.

Finally, I check the robustness of the results by dropping states having different degrees of early expansion prior to 2014 as per Courtemanche et al. (2017). I first drop five states (DE, DC, MA, NY, VT) which had a relatively complete expansion prior to 2014 according to Kaestner et al. (2015). I then drop states which had partial expansion prior to 2014, 14 in the treatment group and 4 in the control group. I then drop all early expanders to have only states that expanded in or after 2014. The results are robust to this sample selection implying that the effects are not differentially driven by states that expanded Medicaid (partially or more completely) prior to 2014. The estimates are provided in Table 24.

3.6 Conclusion

This paper examines the spillover effects of Medicaid expansion on marital decision making. I find that the ACA Medicaid expansion leads to a substitution effect whereby low educated adults over 26 years of age substitute away from coverage as dependents through marriage in the presence of an alternative source of coverage. The estimated figures for the sample is a 6.92% reduction in marriage initiation. This adds to the broader discussion on whether economic incentives can have an impact on marital decision making. Results indicate that economic incentives that change the costs and benefits asso-

ciated with entering marriage can affect outcomes. This study is the first to look at the impact of the ACA Medicaid expansion on marital decision making. Moreover, by extending the discussion to non-elderly adults other than young adults, this study is not restricted to studying marital behavior for a particular age group only.

Even though I do not find heterogeneous effect by sex, the fact that the effects are significant for the sub-sample of women (at a 11.6% drop in new marriages) is suggestive of the dependence that women have on their spouses for coverage. It also underlines their vulnerability to loss of coverage in an event of marriage dissolution that has been discussed in earlier studies. While I cannot discern whether the increase in divorces can be completely attributed to Medicaid expansion, the decrease in new marriages can be interpreted as an indirect validation of the vulnerability of women to loss of coverage owing to marriage dissolution.

The Medicaid expansion may also affect same-sex marriages. They can be explored in conjunction with the state-specific timing of legalization of same-sex marriages to investigate if states which had legalized same-sex marriages at the time of Medicaid expansion differed in outcomes from states which did not. The results may also have implications on fertility which has not been explored as yet. If individuals choose to substitute away from marriage or defer marriage it may have effects on fertility and childbearing. Another area for further research could be aimed at studying the impact of other regulations of the ACA, such as insurance market reforms and sliding scale subsidies, to get a more holistic understanding of the impact of the ACA on marital decision making and family structure. As Abramowitz (2016) points out, the impact of the multiple moving pieces of ACA may counteract each other making identification challenging. Thus, use of a novel methodology to disentangle the effects of the various limbs of the ACA may be an area of future research.

4 State Tax Credits and Impact on Charitable Giving¹⁹

4.1 Introduction

“To give away money is an easy matter and in any man’s power.

But to decide to whom to give it, and how large, and when, and for what purpose and how, is neither in every man’s power nor an easy matter”

— Aristotle

Total giving to charitable organizations in the United States in 2017 exceeded \$410 billion or approximately 2.1% of GDP. As nearly 80 percent of these dollars come from individual donors, there has been a substantial amount of research on the primitives of the economics of charity and the relationship between charities and potential donors (see, e.g., List, 2011 or List and Price, 2012 for overviews of work in this area). To date, much of this work has focused on impacts of different fund-raising techniques on both the number of donors and overall contribution levels to a given cause. While this literature has advanced our understanding of the interaction among nonprofit organizations and potential donors, there is another important player in the sector – the government. Research looking at the role of government in the nonprofit sector has focused on two areas: (i) crowd-out and how government grants impact fund-raising effort (e.g., Kingma, 1989; Payne, 1998; Andreoni and Payne, 2003, 2011a; Andreoni et al., 2014); and (ii) the effect of tax policy and the rate of deductibility on individual donations (Feldstein and Clotfelter, 1976; Clotfelter, 1980; Randolph, 1995; Auten et al., 2002; Fack and Landais, 2010; Duquette, 2016).

This study merges these two strands of literature to explore how state level tax credits

¹⁹This chapter is co-authored with James C. Cox of Georgia State University, Michael K. Price of University of Alabama and Florian Rundhammer of Cornerstone Research. This research was made possible through grant number SES-1658743 from the National Science Foundation. The statements and opinions expressed herein are those of the authors’ alone and do not necessarily reflect those of the National Science Foundation.

impact overall giving and the allocation of donations across qualifying and non-qualifying causes. Prior work has successfully documented the effectiveness of various fund-raising strategies and identified ways for a given nonprofit to increase dollars raised. However, we know very little about whether and how increased giving to one charity impacts giving to others in the sector and the size of the overall “charitable pie.” Do tax incentives that intend to give tax-payers the freedom to determine which organizations receive their tax revenues generate new dollars thereby increasing the size of the charitable pie? Or do such incentives result in an unintended consequence that prompts donors to reallocate funds among the causes they already support?

Such questions are of first-order importance given that the charitable sector is comprised of millions of organizations competing for dollars from a finite set of budget constrained donors. In such an environment, substantial changes in giving to one organization should impact overall patterns of giving. What remains unknown is whether private dollars attracted via government policies (or other fund-raising mechanisms) increase total contributions or simply prompt a substitution of funds across charities. It is this gap in the literature that this study aims to fill. To do so, we implemented a framed field experiment that embedded information about the nation’s largest state income tax credit program for donations to charity which is Arizona’s Credit for Contributions to Qualifying Charitable Organizations – within a modified dictator game. The program provides a Charitable Tax Credit (CTC) which is a dollar-for-dollar tax credit on a tax payer’s state income tax return for donations to qualifying charities. The program structure implies that donations up to the specified threshold are free to the donor since they reduce the individual’s tax liability by the amount claimed.

Participants in the experiment were recruited via Qualtrics to complete an on-line questionnaire (which included the modified dictator game) and were randomized into one of six treatments. The participants in the experiment were adults living in the state of Arizona and were selected to reflect a random sample of this population. In the first-stage

of the modified dictator game, subjects selected either one or two charities as potential recipients in a second stage dictator game. In the second stage, subjects determined how to allocate their \$80 endowment among themselves and the charities selected in the first-stage. Subjects could choose to keep all, some or none of the \$80 for themselves. Experimental treatments varied along three main dimensions: (i) whether, prior to making the first stage decision, subjects were provided detailed information about the credit program and which organizations qualified for the credit; (ii) the number of potential recipients (one or two) that could be selected in the first-stage of the experiment; and (iii) whether the lists from which potential recipients were selected were comprised of only qualifying or non-qualifying charities or a mix of both types. Our experiment was designed to isolate the effect of the tax credit program on five distinct outcomes of interest: (i) the likelihood of making a donation; (ii) the aggregate amount shared with the selected recipients; (iii) the amount of the credit used to increase private good consumption; (iv) the likelihood and number of qualifying and non-qualifying organizations selected as recipients; and (v) the allocation of donations across qualifying and non-qualifying organizations.

To the best of our knowledge, ours is the first experiment to separately identify the effects of tax incentives for giving to a subset of qualifying causes on the choice of charities included in a donor's portfolio and the allocation of funds among qualifying and non-qualifying charities.²⁰ As more than 30 states currently offer some form of tax credit for contributions to qualified causes, understanding how incentives designed to encourage giving to select causes impacts giving to other causes in the sector is important for both policymakers and practitioners in the charitable sector.

Results from our experiment highlight that information has no effect on either the likelihood a subject contributes or the aggregate amount shared with charity. However,

²⁰For example, Filiz-Ozbay and Uler (2019) examine how incentives affect giving amongst a fixed set of charitable recipients (a homeless individual or animal at a shelter) or goods (a toothbrush and a tube of toothpaste). A related body of literature examines the impact of incentives on the allocation of donations amongst a fixed set of "projects" in the context of crowd-funding (e.g., Corazzini et al., 2015; Meer, 2015; Cason and Zubrickas, 2018). Importantly, all of this work holds constant the set of potential recipients and is thus unable to identify the effect of incentives on the composition of a donor's portfolio.

awareness of the charitable tax credit program does affect both the mix of qualifying and non-qualifying charities selected as potential recipients and the allocation of funds among a fixed set of qualifying and non-qualifying organizations. For example, in treatments where subjects are forced to select one qualifying and one non-qualifying charity, information on the tax credit program causes an approximately 8 percentage point increase in the fraction of contributions allocated to the qualifying cause. In treatments where subjects are free to select any mix of qualifying and non-qualifying charities as potential recipients, information on the tax credit program has an even greater effect on the allocation of funds across qualifying and non-qualifying causes due to changes in the composition of the donors' portfolios. In such instances, our information treatment leads to an approximately 22 percentage point increase in the likelihood that a subject selects two qualifying organizations and an approximately 12 percentage point increase in the fraction of contributions allocated to qualifying causes.

4.2 Program Description

In this study, we focus on Arizona's Credit for Contributions to Qualifying Charitable Organizations which is the most generous of all state tax credit programs. The CTC was enacted in 1998 and provides a one-to-one (or 100 percent) tax credit for contributions to qualifying charities. This incentive structure implies that taxpayers can contribute to charity at no cost because giving reduces their state income tax burden by the amount given. The program thus allows taxpayers to redirect dollars from the general tax fund to a charity — or multiple charities — of their choosing.

The CTC offers a tax credit for contributions up to \$400 (single or head of household) or \$800 (married filer) to qualifying charities. Charities must fulfill several criteria to qualify, which we summarize in Figure 6. As described by the Arizona Department of Revenue, to be designated an eligible recipient under this program, the charity must be one that receives community services block grant program monies that spends at least 50%

of its budget on services to Arizona residents who either receive temporary assistance for needy families (TANF) benefits, are low income residents whose household income is less than 150% of the federal poverty level, or are chronically ill or physically disabled children and shows that the charity plans to continue spending at least 50% of its budget on services to those described above.²¹

Figure 6: Criteria for Qualifying Charities



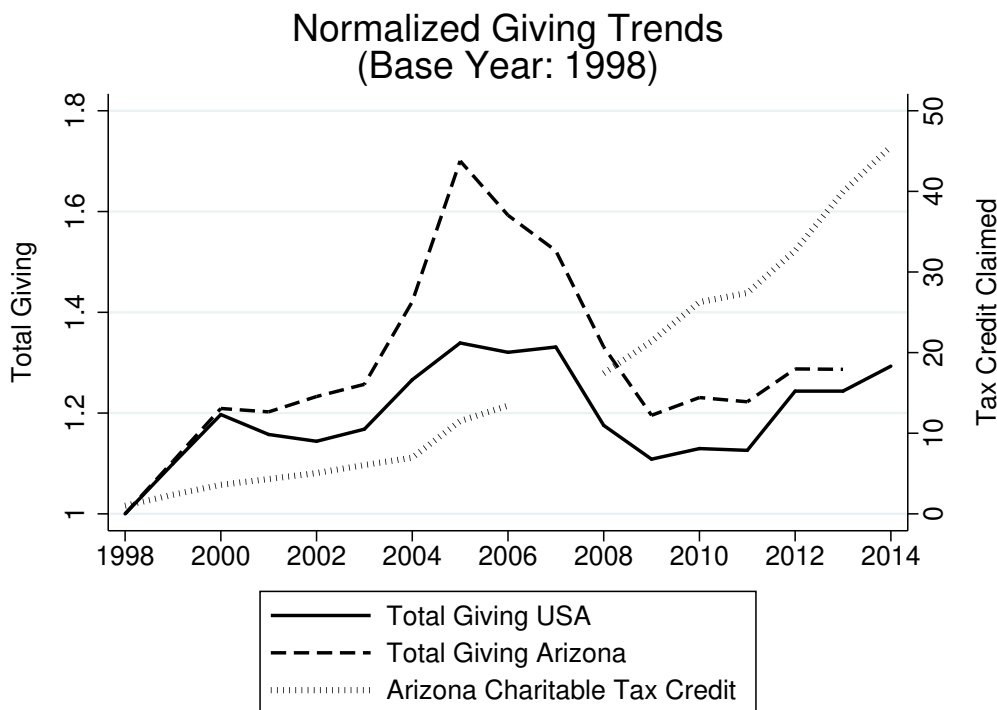
Notes: Overview of the criteria used by the Arizona Department of Revenue to determine the qualifying status of charities for the CTC. Qualifying charities must meet all criteria and they must be registered with the Department of Revenue.

Contributions claimed through the CTC have increased dramatically since the inception of the program in 1998. Figure 7 provides a visual depiction of the trends in the total value of donations claimed through the CTC and trends in aggregate giving to all registered 501(c)(3) charities in Arizona over the period 1998-2014. Values in the figure are normalized relative to the corresponding amount given in 1998. The left hand axis of the figure corresponds to the normalized value of contributions to all registered nonprofits in the state of Arizona whereas the right hand axis of the figure corresponds to the normalized value of contributions claimed through the CTC.

Figure 7 highlights divergent trends in total tax credits and total contributions: while

²¹More details about the requirements can be found at <https://azdor.gov/news-notice/news/changes-tax-credits-contributions-qualifying-charities>

Figure 7: Charitable Giving Trends



Notes: We plot giving in three categories from 1998 to 2015: (i) all giving to registered charities in the US; (ii) giving to registered charities in Arizona; and (iii) giving claimed through the CTC. All annual totals are normalized to giving in 1998 in the respective category.

the total value of contributions claimed through the CTC has increased by a factor of 45 since the inception of the program, there has been little change in aggregate giving statewide over this same time horizon. Hence, while the introduction of the CTC appears to have had a great impact on giving to qualifying organizations, it does not show similar impact on the overall giving to non-profits in Arizona. This poses an interesting question: where are the added dollars flowing to qualifying organizations coming from? Are they additional dollars flowing into qualifying organizations but whose value are too small to impact aggregate trends in giving statewide? Or do they reflect a redistribution of funds from non-qualifying organizations and/or an increased propensity for donors to claim the contributions they would otherwise make to qualifying organizations? Below

we describe a field experiment designed to address these questions.

4.3 Experimental Design and Procedure

4.3.1 Experimental Design

In this section, we describe details of the field experiment. The main component of the experiment is an allocation task whereby subjects decide how to split a fixed amount of money between themselves and charitable organizations selected as recipients. Our allocation task follows Eckel et al. (2005) by implementing a modified dictator game where the recipient is a charitable organization selected from a predetermined list of causes. The modified dictator game proceeded in two stages. In the first stage, the subject was provided a set of charitable organizations and asked to select either one or two organizations from this set as a recipient in the second stage allocation task.

The second stage was the allocation task. In the allocation task, subjects receive an endowment of \$80 and are provided an opportunity to share any portion of the endowment with the organizations selected as recipients in the first stage. Figure 8 provides an example of the allocation task as observed by the subject. Importantly, the decision has consequences. Subjects are paid the amount they elect to keep to themselves and the chosen charities receive the shared allocation as a donation. By prior arrangement with the Arizona Department of Revenue, receipts for charitable contributions made through the experiment can be used for claiming tax credits on Arizona income tax returns.

We vary three dimensions of the modified dictator game across experimental treatments. First, some subjects receive detailed information about the CTC, its history, and qualification requirements for charities prior to selecting recipients in the first stage.²² Second, we vary whether subjects are asked to select one recipient from a set of ten potential recipients or two recipients – one each from two distinct sets of five potential recipients. In treatments that provide information about the CTC, the lists explicitly indicate

²²We show an example of the information in Figure 11 in Appendix C.

Figure 8: Example Allocation Task with Two Charities

How much (out of your \$80) do you want to donate to Southwest Behavioral Health Services [qualified for the tax credit] AND/OR Make-A-Wish Foundation? You will be paid the amount that you do not donate to charity in e-rewards currency within two weeks of this survey. If you donate to a qualified charity, your donation can be refunded to you when you file your state income tax return. Please only enter whole dollar values between 0 and 80 (no cents). Please do not enter the "\$" sign.

Donation to Charity Southwest Behavioral Health Services [qualified for the tax credit]	\$	<input type="text" value="30"/>
Donation to Charity Make-A-Wish Foundation	\$	<input type="text" value="10"/>
Keep to yourself (do not donate) [in e-rewards currency]	\$	<input type="text" value="40"/>
Total	\$	<input type="text" value="80"/>

Next

Notes: Example allocation task. Subjects receive an endowment of \$80 and can freely allocate the amount between themselves and one or two charities that they chose on the previous page of the survey. The survey software indicates whether a charity qualifies for the CTC only in information treatments.

whether donations to a potential recipient would qualify under the CTC. Importantly, across all treatments subjects have the opportunity to select among a mix of qualifying and non-qualifying organizations. Third, in treatments where the subject selects two recipients, we vary the composition of the lists and the corresponding mix of qualifying and non-qualifying organizations that are selected as recipients. In some treatments, subjects select recipients from one list that contains five qualifying organizations while the other list contains five non-qualifying organizations. In such treatments, subjects are “forced” to select one qualifying and one non-qualifying organization as recipients. In other treatments, both lists contain a mix of qualifying and non-qualifying organizations. In these treatments, subjects are thus able to select zero, one, or two qualifying organizations as recipients.²³

²³Table 33 and 34 in Appendix C present all charities in the experiment. Table 35 shows the charity lists by treatment.

In total, the experiment consists of six treatments that we describe in Table 10. The six treatments are arranged in three pairs, where the only difference between groups within a pair is the provision of information about the CTC. The first two treatments, B and T1, have subjects select a single recipient from a list of ten charities before proceeding to the allocation task. In these treatments, the list from which subjects select the recipient organization contains five qualifying and five non-qualifying charities that are arranged in random order. Subjects in treatment B were not provided any information about the CTC or whether donations to a particular organization qualified under this program. Subjects in T1, in contrast, were provided information about the CTC in the first stage of the experiment and, prior to selecting the recipient organization, observed whether or not donations to alternative recipients qualified for a tax credit under the CTC.

Table 10: Description of Treatments in the Experiment

Treatment	Charity Recipient(s)	Choice Set	Charity Types	Information about CTC and Eligibility
Baseline (B)	1	1 list of 10	mixed; 5 qualifying and 5 non-qualifying	No
Treatment 1 (T1)	1	1 list of 10	mixed; 5 qualifying and 5 non-qualifying	Yes
Treatment 2 (T2)	2	2 lists of 5	list 1: qualifying; list 2: non-qualifying	No
Treatment 3 (T3)	2	2 lists of 5	list 1: qualifying; list 2: non-qualifying	Yes
Treatment 4 (T4)	2	2 lists of 5	both lists mixed	No
Treatment 5 (T5)	2	2 lists of 5	both lists mixed	Yes

Notes: The table describes the six treatments of the experiment. Each subject was assigned to only one of the six treatments. Thus, it is a between subject design.

The remaining four treatments provide two lists of five charities each. In these treatments, subjects select one charity from each list as recipients in the allocation task. In treatments T2 and T3, subjects are “forced” to select a recipient of both types as one list contains only qualifying charities, while the second list is made up exclusively of non-qualifying charities. These treatments do not allow subjects to alter the number of qualifying and non-qualifying organizations selected as recipients in response to information about the CTC. The only channel for response is to adjust the amounts allocated to a qual-

ifying or non-qualifying organization.²⁴ Finally, both lists in treatments T4 and T5 contain a mix of qualifying and non-qualifying charities. Subjects in these treatments are thus free to select zero, one, or two qualifying organizations as recipients and adjust the number of qualifying organizations selected as recipients in response to information about the CTC. Hence, subjects in these treatments have two channels for response to information about the CTC; they can reallocate the amount allocated to a fixed mix of qualifying and qualifying recipients and they can adjust the number and mix of each type of organization selected as recipients.

This experimental design allows us to study the impact of the CTC on four dimensions of giving:

- i. the likelihood of making a donation;
- ii. the overall dollar amount given;
- iii. the choice of charity recipient(s); and
- iv. the allocation of donations across qualifying and non-qualifying organizations.

A comparison of B and T1 helps us understand whether knowledge of the tax credit affects the type of organization, qualifying or non-qualifying, selected as the recipient and the resulting amount donated to the selected recipient. The remaining treatments are designed to explore whether and how information about the CTC impacts the total amount given and the corresponding allocation of dollars across qualifying and non-qualifying charities. T2 and T3 are designed to isolate how information about the CTC allocate money across the two types of charities in the second-stage when the subject is “forced” to select one recipient of each type in the first-stage. T4 and T5 allow us to assess whether the CTC affects both the number of qualifying organizations chosen as recipients in the first-stage along with the subsequent amounts donated to the different charity types in

²⁴The design of these treatments thus shares similarity with past experiments exploring how incentives for giving to a subset of potential recipients impact the allocation of funds amongst a fixed set of potential recipients (see, e.g., Null, 2011; Meer, 2015; Cason and Zubrickas, 2018; Filiz-Ozbay and Uler, 2019).

the second-stage.

We chose the charities for this experiment from the published list of all qualifying charities and the universe of non-qualifying 501(c)(3) charities that operate in Arizona. The choice of charities for each treatment pair was based on the size of a charity and its cause. We obtained information about total donations and the official designation of a charity's cause from IRS form 990 filings.²⁵ We then constructed lists containing charities of similar size and causes to avoid corner solutions, where subjects donate (almost) exclusively to a popular charity irrespective of information about the CTC. As a second measure to avoid extreme outcomes, we created two different sets of charity lists within each treatment pair and randomly assigned subjects to one of these lists.²⁶

The final component of the experiment is the provision of a receipt for subjects who made a positive donation. We designed the receipt in partnership with the Arizona Department of Revenue to ensure that it could be used by subjects to claim the CTC. In doing so, we reinforce that a subject's decisions are consequential and impact not only immediate earnings but also future tax liabilities should they donate to a qualifying cause.²⁷

4.3.2 Procedures

We partnered with the company Qualtrics to implement the experiment. The experiment was conducted as an online survey using the software provided by Qualtrics. In addition to the allocation task, the survey contains two question blocks that are identical across all treatments. We show the resulting structure of the survey and experiment in Figure 9. The main purpose of these questions was to ensure that the sample was representative of Arizona's adults and to ensure balance in factors that could impact response to our information treatments. We also address the possibility of field price censoring

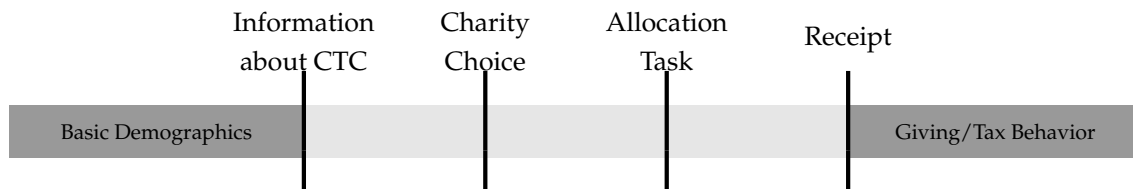
²⁵Income tax-exempt organizations, such as 501(c)(3) charities, file this form to provide the IRS with basic information about the organization and its financial situation, including donations received.

²⁶Table 35 in Appendix C that presents the resulting charity lists and their combinations.

²⁷We also provide a receipt for donations to non-qualifying charities because such donations are generally eligible for federal tax deductions for charitable giving if subjects itemize. Appendix C, Figure 12 provides an example of the receipts provided to subjects in the experiment.

which arises due to the availability of substitutes to the chosen commodity in the laboratory (Harrison et al., 2003). In this case, a rational subject may allocate the entire endowment of \$80 to themselves and then allocate those funds outside the experiment to their preferred qualified charity if it is not in the restricted list. This may be a problem particularly when subjects can choose at most one charity of a particular type from the given list. Questions on past giving behavior and past donation to charities in current choice list are thus posed to be used as controls in the analysis to account for this issue.

Figure 9: Overview of Timeline and Procedures in the Experiment



Notes: The experiment consists of three main phases. First, after consenting to participation, subjects answer basic questions about individual characteristics. Second, subjects face the randomized decision task, which include information about the CTC in some treatments, the choice of one or two recipient charities depending on the treatment, the allocation of the endowment between the subject and the recipient(s), and a receipt in case of a positive donation. Third, we ask subjects about their past donation behavior, knowledge and use of tax credits, and their tax filing behavior. The two question blocks are identical across treatments.

We conducted the survey across two deployment waves that followed identical procedures and contained all six treatments. The first wave took place in December 2017 as giving typically increases substantially around Christmas and its proximity to the end of a tax year. The second wave took place in late April and early May of 2018, shortly after the 2018 Tax Day. Our design thus allows us to study program impacts at different points of the tax year when the incentives provided by the CTC may be more or less salient.

Based on power calculations following List et al. (2011), our experimental sample of 904 subjects across six treatment cells is designed to detect treatment effects of approximately one-third of a standard deviation. Table 11 provides an overview of the resulting sample by experimental treatment and wave. Table 11 further splits the data into com-

plete and incomplete responses where the subject exited the survey after completing the allocation task. In total, we observe 904 completed surveys – 454 in the first wave and 450 in the second – and an additional 347 incomplete responses – 205 in the first wave and 142 in the second.

Table 11: Sample Sizes across Treatments in the Experiment

	Total	B	T1	T2	T3	T4	T5
<i>Wave 1:</i>							
Complete	454	79	76	80	74	84	61
Incomplete	409						
Useful Incomplete	205	40	28	31	37	28	41
<i>Wave 2:</i>							
Complete	450	82	69	82	71	75	71
Incomplete	361						
Useful Incomplete	142	21	28	23	19	30	21
<i>Total:</i>							
Complete	904	161	145	162	145	159	132
Incomplete	770						
Useful Incomplete	347	61	56	54	56	58	62

Notes: This table splits the sample by treatment and wave. We also categorize subjects as “Complete” when they finish the entire online survey. Subjects are categorized as “Incomplete” if they exit the survey without completing it fully. Among the Incomplete subjects, we categorize those who exit after having made the allocation decision as “Useful Incomplete”. We do not have demographic information for subjects who did not complete the survey.

Our partner panel pays subjects in “e-Rewards currency”, an online currency that can be used to purchase goods as well as gift cards from several retailers in an online portal. While e-Rewards currency is not cash-equivalent, its many uses and familiarity of panel participants with the currency make it an attractive payment method. We thus relied on this existing infrastructure to pay subjects at the end of each wave of the experiment.²⁸ All payments to the charities were handled by Georgia State University by tallying the

²⁸We show an example of current options as of April 2018 in Figure 13 in the Appendix C.

total contributions received by each charity.

4.4 Experimental Results

Table 12 provides summary statistics from our experiment. The upper panel of the table presents data on aggregate patterns of giving across our various experimental treatments. Specifically, the upper panel of Table 12 provides data on three metrics of interest; (i) the likelihood that the subject donated to at least one of the selected recipients, (ii) the average contribution level, and (iii) the fraction of all dollars contributed that were given to a qualifying organization. In total, 81.4% of the subjects in our experiment donated to at least one of the selected recipients. The average donation in our experiment was \$49.50 with approximately 59.3% of this amount allocated to qualifying causes.

The middle panel of Table 12 restricts attention to donations made to qualifying causes and summarizes four metrics of interest; (i) the number of qualifying causes selected as recipients in the first-stage, (ii) the likelihood that the subject donated to a selected qualifying cause, (iii) the number of qualifying causes that the subject donated to in treatments T4 and T5, and (iv) the average contribution to qualifying causes. In total, approximately 82% of our subjects selected at least one qualifying organization as a recipient in the first-stage of the modified dictator game. Of these, 80.2% (592 out of 738) made a positive donation to the selected qualifying cause with an average gift of \$29.33 to selected qualifying causes.

If we restrict attention to the final two treatments, T4 and T5, approximately 81% of the subjects (236 out of 291) select at least one qualifying cause as a recipient. Of these, approximately 51.3% (or 121 subjects) select two qualifying causes as recipients in the first-stage. Amongst the 236 subjects who selected at least one qualifying cause, 80.5% (or 190) make a positive donation to at least one of the selected causes. Amongst the 121 subjects that select two qualifying causes, 79.3% make a positive allocation to both of the selected causes.

Table 12: Summary Statistics for the Pooled Sample

	B	T1	T2	T3	T4	T5	Total
<i>All Decisions:</i>							
Subjects	161	145	162	145	159	132	904
Subjects Donated	137	111	130	119	133	106	736
Pr(Donated)	85.1%	76.6%	80.2%	82.1%	83.6%	80.3%	81.4%
Fraction to Qualifying	60.5%	65.0%	51.4%	58.6%	52.8%	70.5%	59.3%
Fraction to Non-qualifying	39.5%	35.0%	48.6%	41.4%	47.2%	29.5%	40.7%
Mean Donation	\$49.11	\$44.47	\$49.73	\$50.95	\$52.23	\$50.30	\$49.50
Conditional Mean	\$57.71	\$58.09	\$61.98	\$62.08	\$62.44	\$62.64	\$60.79
<i>Qualifying Charities:</i>							
Can Select	0 or 1	0 or 1	1	1	0, 1 or 2	0, 1 or 2	
Subjects Selected Any	103	92	162	145	124	112	738
Subjects Selected One	103	92	162	145	75	40	617
Subjects Selected Two					49	72	121
Subjects Donated	87	70	127	118	102	88	592
Pr(Donated)	84.5%	76.1%	78.4%	81.4%	82.3%	78.6%	80.2%
Subjects Donated to One					64	30	
Pr(Donated to One)					51.6%	26.8%	
Subjects Donated to Two					38	58	
Pr(Donated to Two)					30.6%	51.8%	
Mean Donation	\$29.70	\$28.92	\$25.57	\$29.86	\$27.58	\$35.45	\$29.33
Conditional Mean	\$54.97	\$59.91	\$32.62	\$36.69	\$42.99	\$53.18	\$44.79
<i>Non-Qualifying Charities:</i>							
Can Select	0 or 1	0 or 1	1	1	0,1 or 2	0,1 or 2	
Subjects Selected Any	58	53	162	145	110	60	588
Subjects Selected One	58	53	162	145	75	40	533
Subjects Selected Two					35	20	55
Subjects Donated	50	41	122	100	89	45	447
Pr(Donated)	86.2%	77.4%	75.3%	69.0%	80.9%	75.0%	76.0%
Subjects Donated to One					61	29	
Pr(Donated to One)					55.5%	48.3%	
Subjects Donated to Two					28	16	
Pr(Donated to Two)					25.5%	26.7%	
Mean Donation	\$19.40	\$15.54	\$24.16	\$21.10	\$24.65	\$14.85	\$ 20.20
Conditional Mean	\$62.48	\$54.98	\$32.08	\$30.59	\$44.04	\$43.56	\$40.79

Notes: : Figures in the table represent summary statistics across the different treatments.

4.4.1 Impact of Information on Aggregate Pattern of Giving

We begin by exploring the effect of our information treatment on aggregate patterns of giving. Specifically, we explore the effect of information about the CTC program on two metrics of interest: (i) average contribution levels and (ii) the likelihood of making a donation. As summarized in the upper panel of Table 12, average contribution levels range from a low of \$44.47 in treatment T1 to a high of \$52.23 in treatment T4 and participation rates range from a low of 76.6% in treatment T1 to a high of 85.1% in our baseline treatment. If we restrict attention to a comparison across paired information and no information treatments, our raw data suggest no significant effect of information on either average gifts or the likelihood of giving. To formally test the effect of our information treatment on average contributions we estimate a linear regression model of the form:

$$Y_{it} = \alpha + \beta_1 T_i + \beta_2 M_i + \beta_3 T_i M_i + \gamma_t + \epsilon_{it} \quad (11)$$

where, Y_{it} is the aggregate amount donated by subject i in wave t , T_i is an indicator that equals one if subject i participated in an information treatment, M_i is an indicator that equals one if subject i participated in a treatment with two recipients, and γ_t are wave fixed effects. Given prior evidence showing that demographic factors such as age, gender, and income are correlated with charitable donations, we augment our baseline specification to include indicators for: (i) female subjects; (ii) subjects below age 35; (iii) subjects above age 65; (iv) subjects with reported annual income less than \$50,000; and (v) subjects with reported annual income above \$100,000.

The results from these models are presented in Table 13. As noted in the first column of the table, our information treatments have no discernible impact on the average amount donated in our experiment. Both the coefficient on the treatment indicator and the coefficient on the interaction of this indicator with the indicator for multiple-recipient treatments are statistically insignificant. We observe qualitatively similar effects in col-

umn 2 which includes demographic controls although the magnitude of the differences increases.

Table 13: Effect of Information Treatment on Donation

	Pooled	Pooled with Demographic Controls
1(Information Treatment)	-4.8603 (3.5652)	-6.7569* (3.5570)
1(Indicator for Treatment with Multiple Recipients)	1.7312 (2.9331)	0.4358 (2.9171)
Information Treatment \times Multiple Recipients Indicator	4.6883 (4.3992)	6.5963 (4.3751)
1(Female)		6.8362*** (2.1068)
1(Below Age 35)		-1.2378 (2.7190)
1(Above Age 65)		5.7966** (2.5532)
1(Annual Income below \$50,000)		-9.4282*** (2.4667)
1(Annual Income Above \$100,000)		-3.4799 (2.5900)
Fixed Effects:		
Wave Fixed Effects	Yes	Yes
R^2	0.017	0.044
N	904	904

Notes: The dependent variable is the aggregate amount donated by an individual in each wave, pooled over treatments. Robust standard errors are reported in parentheses.

*** denotes significance at the 1 percent level

** denotes significance at the 5 percent level

* denotes significance at the 10 percent level

Before proceeding, we should note that the estimates on our demographic controls are consistent with prior findings in the literature (e.g., List, 2004). Specifically, we find that (i) women donate approximately \$6.84 (or 13.9 percent) more to charity than do male counterparts, (ii) the elderly donate \$5.80 (or 11.8 percent) more to charity than do middle aged counterparts, and (iii) those with annual household income below \$50K donate approximately \$9.43 (or 19.2 percent) less to charity than do counterparts with income in the \$50-100K range. That our estimated demographic effects are consistent with prior findings suggests that our sample is representative of the broader population – at least

with regard to the determinants of giving.

We next explore the impact of our information treatments on the likelihood a subject makes a positive contribution. As noted in Table 12, this probability ranges from 76.6% in treatment T1 to 85.1% in our baseline condition with little difference across our information and no information treatments. To formally test the effect of information about the CTC on the likelihood of giving to a selected cause, we estimate probit models, with the dependent variable Y_{it} now being a binary indicator that equals 1 if subject i in wave t shared a positive amount in the second stage allocation game, and zero otherwise.

Average marginal effects from these models are reported in Table 14 and suggest that our information treatments had no discernible impact on the likelihood of giving. For example, as noted in the first column of the table, subjects in treatment T1 were actually 8.7 percentage points less likely to give than were counterparts in the baseline treatment – a difference that is marginally significant at the $p < 0.10$ level. Although we are unable to pinpoint the cause of this extensive margin effect, it drives the difference in average contributions across our baseline treatment and treatment T1.²⁹ We observe a less pronounced difference in treatments with multiple recipients that is not statistically significant. The qualitative nature of our findings are unchanged when we add demographic controls.

Viewed in its totality, the empirical estimates in Tables 13 and 14 suggest a first result:

Result 1: *Information about the CTC has no impact on aggregate behavior. Both the likelihood of giving and average contributions are unaffected by our information treatment – particularly when subjects select multiple recipients in the first-stage.*

That our information treatment has no impact on average contribution levels in our experiment is consistent with the data patterns observed in Figure 7 which shows that the introduction of the CTC has had no discernible impact on aggregate patterns of giving statewide. In what follows, we set forth to understand what drives this result. Perhaps

²⁹As noted in Table 12, there is no difference in the average conditional donation across our baseline (\$57.71) and treatment T1 (\$58.09).

Table 14: Effect of Information Treatment on Participation (Likelihood of being a Donor)

	Pooled	Pooled with Demographic Controls
1(Information Treatment)	-0.0877* (0.0450)	-0.1051** (0.0451)
1(Indicator for Treatment with Multiple Recipients)	-0.0349 (0.0373)	-0.0449 (0.0365)
Information Treatment \times Multiple Recipients Indicator	0.0814 (0.0545)	0.0973* (0.0543)
1(Female)		0.0882*** (0.0258)
1(Below Age 35)		0.0547 (0.0362)
1(Above Age 65)		0.0174 (0.0307)
1(Annual Income below \$50,000)		-0.0758** (0.0312)
1(Annual Income Above \$100,000)		-0.0478 (0.0321)
Fixed Effects:		
Wave Fixed Effects	Yes	Yes
R^2	0.010	0.029
N	904	904

Notes: The dependent variable is a binary indicator which takes the value 1 when a positive amount is allocated for donation, and 0 otherwise. Average marginal effects reported. Robust standard errors are reported in parentheses.

*** denotes significance at the 1 percent level

** denotes significance at the 5 percent level

* denotes significance at the 10 percent level

information about the CTC program and available tax credits for giving to select causes has no impact on donor choice. Alternately, the aggregate statistics could be misleading and the impact of the CTC is that it leads to a reallocation of donations across qualifying and non-qualifying causes.

4.4.2 The Allocation of Funds across Types of Charites

We next explore whether our information treatment leads subjects to shift donations from qualifying to non-qualifying charities. To do so, we first discuss summary statistics from Table 12. As noted in the table, the fraction of donations allocated to qualifying charities is greater in every information treatment than in its paired no information condition.

To evaluate whether these differences are statistically significant, we estimate a series of probit models with fractional response to estimate the effect of information on the fraction of donations to qualifying charities. We calculate the fraction of donations to a qualifying charity as the amount allocated by subject i to qualifying charities divided by subject i 's aggregate donation and set this fraction to zero should subject i keep the \$80 endowment for themselves.³⁰ By construction, our dependent variable thus takes a value between 0 and 1.

Estimates for these models are presented in Table 15. As noted in the first column of the table, there is no difference in the average fraction allocated to qualifying causes across our baseline treatment and treatment T1 – the estimated coefficient on our indicator for an information treatment is negative but not statistically significant at any meaningful level. There is, however, a significant increase in the amount allocated to qualifying causes when subjects receive information about the CTC and select multiple recipients. The estimated coefficient on this interaction term is 0.153 suggesting that subjects in these treatments allocate approximately 15.3% more to qualifying causes than do counterparts in the corresponding no information treatments .

The qualitative nature of these differences are unchanged when we add demographic controls. As noted in column 2 of Table 15, information has no impact on the fraction allocated to qualifying charities when subjects select a single recipient. However, information about the CTC leads to a significant increase in the fraction allocated to qualifying causes when subjects select multiple recipients.

Viewed in conjunction with the null effect of information on average giving, the estimates in Table 15 suggest a second result:

Result 2: *Providing subjects information about the CTC leads to a reallocation of donations towards qualifying causes but only when subjects select multiple recipients.*

³⁰The qualitative nature of our findings remain unchanged if we instead calculate our dependent variable as the amount allocated to qualifying causes divided by the maximum possible donation – the subject's \$80 endowment.

Table 15: Effect of Information Treatment on Fraction of Donation to Qualifying Charities

	Pooled	Pooled with Demographic Controls
1(Information Treatment)	-0.0581 (0.0572)	-0.0710 (0.0573)
1(Indicator for Treatment with Multiple Recipients)	-0.1142*** (0.0434)	-0.1253*** (0.0433)
Information Treatment \times Multiple Recipients Indicator	0.1534** (0.0644)	0.1672** (0.0647)
1(Female)		0.0391 (0.0284)
1(Below Age 35)		0.0121 (0.0359)
1(Above Age 65)		0.0642* (0.0346)
1(Annual Income below \$50,000)		-0.0500 (0.0330)
1(Annual Income Above \$100,000)		-0.0509 (0.0346)
Fixed Effects:		
Wave Fixed Effects	Yes	Yes
R^2	0.010	0.012
N	904	904

Notes: The fraction is defined as donation to qualifying charities over the total donation of a subject. This is closer to a measure of conditional giving to qualifying charities. For subjects who did not donate at all, we code the fraction as zero, while keeping them in the sample.

Average marginal effects reported. Robust standard errors are reported in parentheses.

*** denotes significance at the 1 percent level

** denotes significance at the 5 percent level

* denotes significance at the 10 percent level.

Result 2 shares similarity with findings in Null (2011) who shows that changes in the relative price of giving leads to weak substitution between causes and a partial reallocation of donations from higher to lower priced causes. Result 2 also shares similarity with findings in Filiz-Ozbay and Uler (2019) that increases in the rebate rate for one charity relative to that of a substitute cause leads to a reallocation of funds among the two causes. By design, subjects in our experiment should view qualifying and non-qualifying organizations as substitute causes so we should expect information about the CTC to cause a reallocation of donations towards qualifying causes.

For policy-makers and practitioners, Result 2 provides a potential explanation for the

patterns of giving illustrated in Figure 7. The introduction of the CTC likely caused a reallocation of donations towards qualifying causes. Hence, there can be a dramatic increase in the amounts claimed through the CTC program with no change in aggregate patterns of giving statewide. For researchers, Result 2 highlights the importance of examining the effects of incentives not only on giving to the targeted cause but also on giving to other causes i.e. it highlights the importance of modeling choice and testing behavior in a world with multiple public goods.

4.4.3 Impact of Information on Charity Selection and Allocation

We next set forth to explore the various channels through which our information treatment impacts the allocation of funds among qualifying and non-qualifying causes. In doing so we examine the effects of information along two distinct margins: (i) the extensive margin – the selection of recipients in the first stage; and (ii) the intensive margin – the allocation of funds in the second stage. Throughout, we restrict attention to the subset of treatments where subjects select multiple recipients as we do not observe significant differences in the allocation across the different cause types in treatments with a single recipient.

We begin by exploring the effect of our information treatment on the selection of recipients in the first-stage of the modified dictator game. Recall that by design, subjects in treatments T2 and T3 were “forced” to select one recipient of each type. We thus focus our analysis on treatments T4 and T5 where subjects faced mixed lists and could select a “portfolio” of recipients that includes zero, one, or two qualifying causes. In doing so, we focus on three metrics of interest: (i) the likelihood of selecting at least one qualifying cause as a recipient; (ii) the likelihood of selecting at least one non-qualifying cause as a recipient; and (iii) the number of qualifying causes selected as recipients.

As noted in Table 12, our information treatment impacts the selection of both recipient types. For example, information about the CTC increases the likelihood of selecting at

least one qualifying cause by approximately 6.9 percentage points (84.9 percent in T5 versus 78 percent in T4). In contrast, information about the the CTC reduces the likelihood of selecting at least one non-qualifying cause by 23.7 percentage points (45.5 percent in T5 versus 69.2 percent in T4).

Taken jointly, these differences lead to changes in the composition in the mix of cause types within the donors “portfolio” of selected recipients. For example, whereas 47.2 percent of subjects (75 out of 159) in T4 select one cause of each type, this fraction falls to approximately 30.3 percent (40 out of 132) in the paired information treatment. In contrast, the number of subjects selecting two qualifying causes as recipients increases from approximately 30.8 percent (49 out of 159) in the no-information condition to more than 54 percent (72 out of 132) in the paired information condition.

To evaluate whether these differences are statistically significant, we estimate a series of probit models of the form:

$$Y_{ilt} = \alpha + \beta T_i + \gamma_t + \nu_l + \epsilon_{it} \quad (12)$$

where, Y_{ilt} equals one if subject i facing list l in wave t selects at least one qualifying (non-qualifying) cause as a recipient, T_i is indicator for subjects that in our information treatment (T5), γ_t are wave fixed effects, and ν_l are list fixed effects. We also estimate an augmented version of this basic model that includes demographic controls.

Estimates for the marginal effects of the information treatment on the probability of selecting at least one recipient of the given type are presented in Table 16. The first two columns examine the likelihood of selecting at least one qualifying cause as a recipient whereas the last two columns examine the likelihood of selecting at least one non-qualifying cause as a recipient. Empirical results are largely consistent with the aggregate summary statistics. For example, subjects in an information treatment are approximately 3 percentage points more likely to select at least one qualifying cause as a recipient. However, this difference is not statistically significant at any meaningful level.

Table 16: Effect of Information Treatment on Selecting a Qualifying/ Non-Qualifying Charity for T4 & T5

	Qualifying Charity	Qualifying Charity with Demographic Controls	Non-qualifying Charity	Non-qualifying Charity with Demographic Controls
1(Information Treatment)	0.0274 (0.0564)	0.0257 (0.0567)	-0.2197*** (0.1525)	-0.2222*** (0.0612)
1(Female)		0.0102 (0.0583)		-0.0133 (0.0623)
1(Below Age 35)		-0.0216 (0.0771)		0.1043 (0.0836)
1(Above Age 65)		-0.0231 (0.0656)		-0.1279* (0.0710)
1(Annual Income below \$50,000)		-0.0194 (0.0690)		0.0171 (0.0747)
1(Annual Income Above \$100,000)		0.0219 (0.0715)		0.0224 (0.0747)
Fixed Effects:				
List Fixed Effects	Yes	Yes	Yes	Yes
Wave Fixed Effects	Yes	Yes	Yes	Yes
R^2	0.011	0.013	0.056	0.073
N	291	291	291	291

Notes: This table includes T4 and T5 treatments only since these treatments allow the choice of more than one charity of similar type.

The dependent variable is a binary indicator which takes the value 1 when a qualifying(non-qualifying) charity is chosen, and 0 otherwise. Average marginal effects reported.

Robust standard errors are reported in parentheses.

*** denotes significance at the 1 percent level

** denotes significance at the 5 percent level

* denotes significance at the 10 percent level

As noted in the last two columns of Table 16, our information treatment does have a significant impact on the likelihood of selecting at least one non-qualifying cause as a recipient. Subjects in treatment T5 are approximately 22 percentage points less likely to select at least one non-qualifying charity as a recipient than counterparts in the no-information condition (T4).

Viewed in conjunction, the above results suggest that information about the CTC influences the composition of the “portfolio” of cause types selected by subjects. To formally evaluate this conjecture, we estimate the effect of our information treatment on the number of qualifying causes selected as recipients in the first-stage of the modified dictator game. To do so, we estimate a series of probit models similar to those described above. However, now the dependent variable is an indicator for whether the subject selects a

specific number – zero, one, and two – of qualifying causes. Results from this exercise are presented in Table 17.

Table 17: Effect of Information Treatment on Selecting a Specific Number of Qualifying Charities

	Chose Zero	Chose Zero	Chose One	Chose One	Chose Two	Chose Two
1(Information Treatment)	-0.0628 (0.0458)	-0.0587 (0.0456)	-0.1695*** (0.0588)	-0.1784*** (0.0597)	0.2408*** (0.0601)	0.2463*** (0.0609)
1(Female)		0.0416 (0.0466)		-0.1338** (0.0610)		0.0887 (0.0621)
1(Below Age 35)		0.0527 (0.0598)		0.0753 (0.0809)		-0.1504* (0.0885)
1(Above Age 65)		0.0152 (0.0530)		-0.0913 (0.0686)		0.0734 (0.0691)
1(Annual Income below \$50,000)		0.0085 (0.0553)		0.0295 (0.0725)		-0.0422 (0.0738)
1(Annual Income Above \$100,000)		-0.0606 (0.0590)		0.0761 (0.0738)		-0.0134 (0.0752)
Fixed Effects:						
List Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Wave Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.023	0.037	0.0400	0.062	0.090	0.108
N	291	291	291	291	291	291

Notes: This table includes T4 and T5 treatments only since these treatments allow the choice of more than one charity of similar type.

The dependent variable is an indicator which takes values between 0 and 2 indicating the number of qualifying charity/charities being selected. Average marginal effect reported.

Robust standard errors are reported in parentheses.

*** denotes significance at the 1 percent level

** denotes significance at the 5 percent level

* denotes significance at the 10 percent level

Results from the table are consistent with our raw data summary and suggest a significant change in the composition of a subject’s “portfolio”. Specifically, we find that subjects in our information treatment are more than 17 percentage points less likely to select a mixed portfolio that includes one qualifying and one non-qualifying charity – a difference that is statistically significant at the $p < 0.01$ level and robust to the inclusion of demographic controls. In contrast, we find that subjects in our information treatment are approximately 24 percentage points more likely to select a portfolio that contains only qualifying causes – a difference that is statistically significant at the $p < 0.01$ level and

robust to the inclusion of demographic controls.

Estimates from Tables 16 and 17 suggest a third result:

***Result 3:** Information about the CTC program and whether donations to a given cause qualify under the program influences the mix of cause types supported.*

We believe this result is novel to the literature and suggests a new channel through which incentives and competition among charities influence donor behavior. This has implications for the importance of measuring social benefits of donations to a given cause for researchers. For policy-makers it highlights the importance of understanding that policies aiming to enhance donations to one cause should be designed keeping in mind not only who gains from the program but also who loses once the program is enacted.

As a final metric of interest, we revisit the effect of our information treatments on the allocation of donations across qualifying and non-qualifying causes – the intensive margin effect. To do so, we restrict attention to the subset of subjects who donated to at least one cause and estimate a probit model with fractional response, with Y_{it} now being the fraction of subject i 's total contribution that is directed to qualifying causes and the remaining variables are identical to those described above.

We estimate the model for two different pairwise comparisons – T2 vs T3 and T4 vs T5. The first comparison is akin to that explored in prior work (e.g, Null, 2011; Ek, 2017; Filiz-Ozbay and Uler, 2019) which report how incentives affect allocation of donations across a fixed set of cause types. The latter extends this prior work and captures the combined effect of changes in the composition of the donor's portfolio of causes and any changes in the allocation of funds across a fixed portfolio type.

Results for these models are presented in Table 18 which reports the average marginal effects. The estimates suggest that information about the CTC programs leads subjects to increase the fraction of donations allocated to qualifying charities. For example, when subjects cannot adjust the mix of qualifying and non-qualifying charities selected as recipients, information provision causes an approximately 8 percentage point increase in

the amount allocated to qualified causes. For perspective, the average amount donated in treatment T2 (the no information benchmark) is approximately \$49.73. The estimated treatment effect thus corresponds to an increase of approximately \$3.98 in the amount allocated to qualifying causes.

Table 18: Effect of Information Treatment on Fraction of Donation to Qualifying Charities

	T2 vs. T3	T2 vs. T3	T4 vs. T5	T4 vs. T5
1(Information Treatment)	0.0801** (0.0319)	0.0799** (0.0314)	0.1200** (0.0509)	0.1191** (0.0509)
1(Female)		0.0444 (0.0314)		0.0434 (0.0519)
1(Below Age 35)		0.0825** (0.0371)		-0.0578 (0.0661)
1(Above Age 65)		0.1034** (0.0428)		0.0253 (0.0607)
1(Annual Income below \$50,000)		-0.0672* (0.0367)		-0.0445 (0.617)
1(Annual Income Above \$100,000)		-0.0000 (0.0413)		-0.0002 (0.0638)
Fixed Effects:				
List Fixed Effects	Yes	Yes	Yes	Yes
Wave Fixed Effects	Yes	Yes	Yes	Yes
R^2	0.008	0.016	0.024	0.029
N	307	307	291	291

Notes: We consider treatments T2 to T5 here since subjects can allocate their endowment across two types of charities in these treatments.

The fraction is defined as donation to qualifying charities over the total donation of a subject. We consider only donors in this sample i.e. people who have made a positive donation.

Average marginal effects reported. Robust standard errors are reported in parentheses.

*** denotes significance at the 1 percent level

** denotes significance at the 5 percent level

* denotes significance at the 10 percent level

Such effects are enhanced when donors are allowed to both adjust the composition of their portfolios and reallocate funds within a fixed portfolio mix. As noted in the final two columns of the table, the effect of information provision of the share allocated

to qualifying causes is an approximately 12 percentage point increase. For perspective, the average amount donated in treatment T4 (the no information benchmark) is approximately \$52.23. The estimated treatment effect thus corresponds to an increase of around \$6.17 in the average amount donated to qualifying causes.

Viewed in its totality, the estimates in Table 18 suggest a fourth result:

***Result 4:** Information about the CTC program causes a reallocation of funds among qualifying and non-qualifying causes; an effect that is enhanced when donors are allowed to adjust the composition of cause types supported.*

Result 4 highlights the two channels through which the CTC program likely impacts donor choice. For those donors who support both types of causes, the program leads to a shift in donations away from non-qualifying causes and towards qualifying causes. However, there is an additional effect that enhances such reallocation. For a subset of donors, the CTC leads to a change in the types of causes supported. Importantly, this helps explain why contributions claimed through the CTC program have grown exponentially with no discernible impact on overall giving in the state; the program proverbially robs Peter to pay Paul.

4.4.4 Robustness Checks

As a first robustness check, we estimate the effect of our information treatment on the likelihood of donating to a given cause type. To do so, we estimate a series of linear probability models of binary indicators for whether or not the subject donated to a given cause type on our indicator for subjects assigned to an information condition and both list (choice set) and wave of survey fixed effects. As in the prior section, we restrict the analysis to the subset of treatments where subjects select multiple recipients and estimate the model separately for our two pairwise comparisons of interest (T2 vs T3 and T4 vs T5).

Results from these models are presented in Tables 25 and 26 of the Appendix C and

provide further insight into the channels through which information impacts the allocation of funds across cause types. For example, as noted in Table 25, information provision had a small, but statistically insignificant, impact on the likelihood of making a donation to a qualifying cause. Similarly, Table 26 provides evidence that information provision had a negative, but statistically insignificant impact, on the likelihood of donating to a non-qualifying cause in those treatments where subjects were “forced” to select one cause of each type. In contrast, information provision had a negative and statistically significant effect on the likelihood of donating to non-qualifying cause in situations where subjects were free to select any mix of qualifying and non-qualifying causes.

Viewed in conjunction with results from Tables 16-18, these results reinforce that information provision works through different channels across these two treatment types. In situations where subjects cannot adjust the mix of cause types supported, information provision works solely along the intensive margin – subjects shift a portion of what they would have otherwise given to non-qualifying causes to qualifying recipients. However, when subjects are free to adjust the mix of cause types supported, information provision works predominantly through selection of causes – the extensive margin.

As a second robustness check, we expand our sample to include incomplete responses and rerun our various econometric models. These are respondents who completed the first block of the survey and the donation decision but did not complete the second block of survey questions. The inclusion of such observations affords a way to check for potential selection effects and expand the power of our statistical tests.³¹ Results from these models are presented in Appendix C Tables 27-32 and provide qualitative support for our main findings highlighting the robustness of our main results. While the point estimates are muted when we include incomplete responses, they remain economically and

³¹Recall that our second block of survey questions focused largely on tax morale and prior charitable donations. Such questions may have lead subjects to infer that the experiment was designed to measure altruism or generosity. As such, it is possible that subjects who allocated less to selected causes in the second stage allocation game would disproportionately drop out of the survey while answering the second block of questions. If so, this would bias our sample of completed surveys in favor of more altruistic types. Ex ante, it is unclear whether and how this would impact our estimated treatment effects.

statistically significant.

4.5 Discussion

Policy makers frequently attempt to encourage donations to charity with special tax provisions such as income tax deductions and credits. A key feature of many such programs is that the provisions only apply for donations to select causes rather than any registered 501(c)(3) organization. We report an online field experiment around the largest such state program, Arizona's state income tax credit for donations to qualifying charities to ascertain how awareness of the program affects overall giving and the allocation of donations among causes.

Our research questions were motivated by two trends in giving following the introduction of the CTC: (i) claims for contributions to qualifying causes increased nearly 50 fold since the program's inception; and (ii) aggregate contributions statewide remained fairly constant over this same time period. We designed our experiment to ascertain whether targeted tax credits increase aggregate patterns of giving or lead donors to reallocate donations among causes. Our design further allows us to separately identify the effect of targeted tax credits along two margins of interests – the types of causes a donor elects to support and the allocation of funds among selected causes.

Empirical results from our experiment show that information provision has no impact on either the number of donors or the aggregate amount donated. However, information about the program does influence the allocation of funds among qualifying and non-qualifying causes but only when subjects select multiple recipients. Results from our experiment thus provide a potential explanation for the observed trends in giving in Arizona that motivated our experiment – the increased contributions to qualifying causes may be coming at the expense of donations to other, related causes and thus the program has had limited impact on aggregate giving statewide.

Exploring the mechanisms underlying the reallocation across cause type, we find that

it reflects changes along both the intensive and extensive margins. When subjects are unable to adjust the mix of cause types supported, information about the CTC leads them to reallocate some of the money that they would have given to non-qualifying causes to increased donations to qualifying causes. When subjects are allowed to adjust the mix of cause types supported, information about the CTC reduces the likelihood of selecting non-qualifying causes and increases the likelihood of selecting two qualifying causes.

Our findings should be of interest for both researchers and policymakers alike. For policymakers, our results provide a cautionary tale and highlight an unintended consequence of policies designed to encourage giving to a subset of targeted causes. Increased donations to targeted causes may serve to crowd out donations to other causes and thus have limited impact on overall patterns of giving. It is thus important to consider such effects when designing targeted policies and to understand not only who would gain from the program but also would lose if it were enacted. For researchers, our results suggest the importance of extending our theoretical models and empirical analysis to consider multiple public goods. Specifically, we should extend our models to explore whether and how targeted incentives such as tax credits or fund-raising mechanisms used by individual charities influence the selection of causes a donor supports and the allocation of donations across causes. Moreover, our findings suggest the need to develop methods to measure the social benefit of dollars allocated to different causes because the main effect of targeted charitable tax credits is re-allocation across causes, not increase in total donations.

5 Conclusion

The three chapters of the dissertation highlights an important issue in the discourse of social sciences : unintended consequences may not necessarily be unanticipated. Thus, to understand the complete effects of policies, it is important to account for and preempt such unintended effects. Underlying incentives of policies can impact behavior and decision-making with regard to factors not directly targeted by policies. While such impacts may be unintended, considering these possible spillover effects in designing policies would result in better policy-making.

The first chapter shows that when designing policies to promote diversity and inclusion we should consider dual objectives of representation and performance, especially since the latter can impact the former. The laboratory experiment shows that later life compensations for early life disparities can improve representation but policies that directly address bridging these disparities can improve both representation and achievements. The success of diversity enhancing policies depend on the perceptions of such policies across the board which can impact employer perceptions as well as inter-group interactions. The variants of the policies studied here thus provide opportunities of further research in understanding how such policies affect discrimination as well as interactions such as cooperation, stereotypes and statistical discrimination.

The second chapter shows that access to health insurance coverage can affect demographic outcomes such as marital decision making. When provided with an alternative avenue of health insurance through Medicaid, individuals substitute away from obtaining coverage through their spouses, leading to a reduction in marriages. Future research should aim at studying the effect of other provisions of the ACA such as the premium subsidies and also extend to other outcomes such as fertility.

The third chapter shows yet another example of unintended consequences. The failure to consider a model of multiple public goods can lead to policies that fulfill the intention of giving tax-payers more freedom in choosing which cause they want tax dollars to be

re-directed to, but leads to an unintended consequence of reallocating funds away from ineligible charities to eligible ones. Complementing the field experiment with observational data at the charity level would give us an idea of the extent of the effect of this policy.

The dissertation establishes a strong case of experimental methods as the first step in understanding behavioral channels which can inform public policy by clean identification and measurement, especially where observational data is absent or insufficient. However, experiments do have limitations of external validity to provide sufficient evidence for designing and evaluating policy. Thus it is important to complement our understanding of behavioral phenomena from experiments with other quasi-experimental and empirical studies with observational data for a better understanding of policy-making and evaluation.

6 Appendices

6.1 Appendix A

Instructions for Baseline Treatment

Thank you for agreeing to take part in this study. Please read the following instructions carefully. A clear understanding of the instructions will help you make better decisions and increase your earnings.

Please hold on to your station number. You will need this for payment.

Please do not talk. If you have a question, please raise your hand and someone will come over to answer you.

Anonymity

The decisions you make in this experiment and your outcome and earnings will remain anonymous.

Task

You will be answering math problems taken from the ACT math section. Each question will always have one correct answer. You can use only the paper and pencil provided. Calculators, cellphones, or any other helping device is not allowed.

There will be three parts in the experiment which will be explained in detail later.

Earnings:

You will be paid for Part 3 and either Part 1 or Part 2 of the experiment. The Part 1 or Part 2 chosen to be paid will be decided by a coin toss. If the coin toss yields Head, Part 1 will be chosen for payment. If the coin toss yields Tail, Part 2 will be chosen for payment.

Your final earnings will be: Earnings from Part 1 or Part 2 (based on coin toss) + Earnings

from Part 3 + \$5 (show up fees)

For the part chosen:

If you are a winner, your final payoff is: $\$15 + (\text{Your payoff from Part 3}) + \5 (show up fee) If you lose, your final payoff is: $\$5 + (\text{Your payoff from Part 3}) + \5 (show up fee)

PART 1

Groups & Types

You will be randomly put into **groups of 6**.

Within these groups, you will be randomly assigned to one of two types, **Type A or Type D** by the computer. You will retain your type assignment and group through all parts (parts 1,2 & 3) of the experiment. There will be 3 members of each type in a group.

Type A has an advantage over Type D in this part which will be explained below.

Stages

This part consists of two stages, a practice round and a tournament.

Stage 1: Practice Round

In this stage, if you are in **Type A**, you will take a practice round. This round will have questions that Type A subjects will take as practice for the next round. The questions will come with feedback i.e. the correct answers will also be shown to subjects in Type A on answering the questions.

If you are in **Type D** you will not have the opportunity to take the practice round.

Note that some of the questions from the practice questions will be repeated in the next round.

Stage 2: Tournament

Once the practice round is over, your computer will start the tournament. You will be asked to answer the questions on your screen correctly. You get 1 point for each correct answer. There is no penalty for incorrect answers. You have 1 minute to answer each question. If you do not select an answer within 1 minute, it will be recorded as “No Answer” and earn you 0 points. Once the tournament is over, subjects with the **top 2 scores** in a group will be the winners, independent of what type they are. Thus, you are competing with **ALL TYPES** of subjects in your group of 6. In case of a tie, the computer will randomly break the tie.

Payoff from Part 1: The 2 winners i.e. the top 2 scorers will get a payoff of \$15.

The other 4 members in a group will get a payoff of \$5.

Following completion of Part 1, there will be two more parts in this experiment that will be explained in detail after the completion of Part 1. Instructions to these parts will be provided later.

Instructions for Ex-Ante Treatment

PART 2

You will retain your type assignment from the previous part. Thus, if you were Type A in the previous part 1, you will be Type A in part 2 as well. You will also have the same group as before.

NOTE: The difference in the conditions between the two types is not the same as earlier and will be explained here.

Stages

The experiment consists of two stages as before, a practice round and a tournament.

Stage 1: Practice Round

In this stage, if you are in **Type A**, you will take a practice round. Now, if you are in **Type D** you will be **given an option** to take the practice round. **Thus, if you are in Type D, you will decide whether to take the practice round. However, the number of questions from the practice round that will be repeated in the next round will be LESS for Type D than for Type A.** Thus if you are in Type D you will have **LESS** repeat questions in the tournament stage than Type A. As before, the questions will come with feedback.

Stage 2: Tournament

Once the practice round is over, the tournament will be played out exactly as in Part 1.

Once the tournament is over, subjects with the **top 2 scores** in a group will be the winners, **independent of what type they are.** Thus, you are competing with **ALL TYPES** of subjects in your group of 6. In case of a tie, the computer will randomly break the tie.

Payoff from Part 2:

The 2 winners i.e. the top 2 scorers in a group will get a payoff of \$15.

The other 4 members in a group will get a payoff of \$5.

Instructions for Ex-Post Treatment

PART 2

You will retain your type assignment from the previous part. Thus, if you were Type A in the previous part 1, you will be Type A in part 2 as well. You will also have the same group as before.

Type A will have an advantage over Type D in this part as earlier. However, the way winners are chosen will vary in this part. This will be explained later in detail.

Stages

The experiment consists of two stages as before, a practice round and a tournament.

Stage 1: Practice Round

In this stage, if you are in **Type A**, you will take a practice round. As before, the practice round will have questions, some of which will be repeated in the next round and the questions will come with feedback, visible **only to Type A subjects**.

If you are in Type D you will not have the opportunity to take the practice round.

Stage 2: Tournament

Once the practice round is over, proceed as earlier, with 1 point per correct answer. However, now subjects with the **top score from each type**, will be winners. Thus, there will be **1 winner of each type**, and 2 winners in total. Note, you are therefore competing only with subjects of the **SAME TYPE** as you. In case of a tie, the computer will randomly break the tie.

Payoff from Part 2:

The winners i.e. . the top scorer of each type in a group will get a payoff of \$15.

The other 4 members in a group will get a payoff of \$5.

Instructions for Ability Round

PART 3

In this part, your task is to correctly answer as many questions as you can in 5 minutes.

Please note, each correct answer earns you \$0.50.

You lose \$0.25 per incorrect answer.

You neither gain nor lose anything if you choose to skip a question.

Payoff from Part 3 = (No. of correct answers) *\$0.50 – (No. of incorrect answers) *\$0.25

Final Earnings:

You will be paid for Part 3 and either Part 1 or Part 2 of the experiment. The Part 1 or Part 2 chosen to be paid will be decided by a coin toss. If the coin toss yields **Head, Part 1** will be chosen for payment. If the coin toss yields **Tail, Part 2** will be chosen for payment.

For the part chosen:

If you are a winner, your final payoff is: \$15 + (Your payoff from Part 3) + \$5 (show up fee)
If you lose, your final payoff is: \$5 + (Your payoff from Part 3) + \$5 (show up fee)

Figure 10: Screenshot from Experiment

QUESTION 3

If 25% of the given number is 12.5, then what is 60% of the same number?

- 50
- 60
- 10
- 30

Please answer the above question and click submit to proceed to the next question.

Submit your answer

Your performance history

Question	Result	Point
Question 1	Your answer was Incorrect	0
Question 2	Your answer was Correct	1

6.2 Appendix B

Table 19: Newly Married Event Study Estimates

	Full Sample	Female	Male
Medicaid x Year 2011	0.0001 (0.0016)	-0.0012 (0.0015)	0.0013 (0.0023)
Medicaid x Year 2012	0.0028 (0.0018)	0.0001 (0.0022)	0.0054** (0.0022)
Medicaid x Year 2013	0.0009 (0.0016)	-0.0010 (0.0018)	0.0029 (0.0021)
Medicaid x Year 2015	-0.0015 (0.0013)	-0.0043*** (0.0014)	0.0013 (0.0018)
Medicaid x Year 2016	0.0015 (0.0010)	0.0043** (0.0015)	0.0009 (0.0013)
Medicaid x Year 2017	-0.0019 (0.0015)	-0.0031** (0.0016)	-0.0006 (0.0020)
p-values for joint significance	0.24	0.66	0.06

Source: American Community Survey

Notes: The sample is restricted to non-elderly adults above 26 years of age with a high school degree or lower who got married or stayed unmarried in the calendar year previous to the survey year. The base year for the event study is 2014 survey year i.e. 2013 calendar year.

*** Significant at 1 per cent level. ** significant at 2 per cent level. * significant at 10 per cent level.

Table 20: Newly Divorced Event Study Estimates

	Full Sample	Female	Male
Medicaid x Year 2011	-0.0005 (0.0008)	-0.0001 (0.0015)	-0.0009 (0.0014)
Medicaid x Year 2012	0.0012 (0.0011)	0.0019 (0.0017)	0.0005 (0.0013)
Medicaid x Year 2013	-0.0015 (0.0010)	-0.0018 (0.0013)	-0.0012 (0.0012)
Medicaid x Year 2015	0.0013 (0.0010)	0.0019 (0.0016)	0.0007 (0.0014)
Medicaid x Year 2016	0.0005 (0.0010)	0.0006 (0.0015)	0.0004 (0.0013)
Medicaid x Year 2017	0.0011 (0.0011)	0.0026 (0.0012)	-0.0001 (0.0016)
p-values for joint significance	0.04	0.07	0.40

Source: American Community Survey

Notes: The sample is restricted to non-elderly adults above 26 years of age with a high school degree or lower who got divorced in the past 12 months and all those who remained married, for each calendar year.

*** Significant at 1 per cent level. ** significant at 2 per cent level. * significant at 10 per cent level.

Table 21: Newly Married Difference-in-Differences Estimates for Low Income Sample

	Full Sample	Female	Male
Medicaid x Post	-0.0014 (0.0012)	-0.0015* (0.0008)	-0.0016 (0.0020)
Pre-treatment mean	0.0217	0.0183	0.0267
% Change	-6.45%	-8.19%	-5.99%
N	1,093,923	660,884	433,039

Source: American Community Survey

Notes: The sample is restricted to non-elderly adults above 26 years of age with income below 138% of FPL who got married or stayed unmarried in the calendar year previous to the survey year.

*** Significant at 1 per cent level. ** significant at 2 per cent level. * significant at 10 per cent level.

Table 22: Logit Model Estimates for Newly Married

	Full Sample	Female	Male
Medicaid x Post	-0.0022** (0.0011)	-0.0032*** (0.0008)	-0.0013 (0.0015)
N	1,628,480	774,843	853,637

Source: American Community Survey

Notes: The sample is restricted to non-elderly adults above 26 years of age with a high school degree or lower who got married or stayed unmarried in the calendar year previous to the survey year.

Regression includes demographic and unemployment controls.

*** Significant at 1 per cent level. ** significant at 2 per cent level. * significant at 10 per cent level.

Table 23: Difference-in-Differences Estimates with Alternate Definition of Newly Married

	Full Sample	Female	Male
Medicaid x Post	-0.0019* (0.0010)	-0.0018** (0.0008)	-0.0018 (0.0013)
Pre-treatment mean	0.0196	0.0168	0.0222
% Change	-9.69%	-10.71%	-8.11%
N	1,630,366	775,539	854,827

Source: American Community Survey

Notes: The sample is restricted to non-elderly adults above 26 years of age with income below 138% of FPL who got married in the last 12 months or stayed unmarried in the survey year.

Regression includes demographic and unemployment controls.

*** Significant at 1 per cent level. ** significant at 2 per cent level. * significant at 10 per cent level.

Table 24: Newly Married Difference-in-Differences Estimates by Dropping Early Expansion States

	Full Sample			Female			Male		
	Dropping Full Expansion Early Expanders	Dropping Partial Expansion Early Expanders	Dropping All Early Expanders	Dropping Full Expansion Early Expanders	Dropping Partial Expansion Early Expanders	Dropping All Early Expanders	Dropping Full Expansion Early Expanders	Dropping Partial Expansion Early Expanders	Dropping All Early Expanders
Medicaid Expansion x Post	-0.0023** (0.0011)	-0.0020 (0.0013)	-0.0020 (0.0014)	-0.0035*** (0.0008)	-0.0030*** (0.0009)	0.0030*** (0.0010)	-0.0014 (0.0016)	-0.0013 (0.0019)	-0.0013 (0.0021)
Pre-treatment mean	0.0350	0.0334	0.0337	0.0306	0.0288	0.0295	0.0390	0.0376	0.0373
% change	-6.57%	-5.99%	-5.94%	-11.44%	10.42%	-10.17%	-3.59%	-3.46%	-3.49%
N	1,484,787	1,032,469	888,776	704,520	495,058	424,735	780,267	537,411	464,041

Source: American Community Survey.

Notes: *** Significant at 1 per cent level. ** Significant at 2 per cent level. * Significant at 10 per cent level.

6.3 Appendix C

Table 25: Effect of Information Treatment on Making a Donation to a Qualifying Charity

	T2 vs. T3	T2 vs. T3	T4 vs. T5	T4 vs. T5
1(Information Treatment)	0.0358 (0.0457)	0.0362 (0.0451)	0.0273 (0.0564)	0.0257 (0.0567)
1(Female)		0.0984** (0.0448)		0.0100 (0.0583)
1(Below Age 35)		0.1458** (0.0650)		-0.0216 (0.0771)
1(Above Age 65)		0.0815 (0.0554)		-0.0231 (0.0656)
1(Annual Income below \$50,000)		-0.0194 (0.0527)		-0.0193 (0.0690)
1(Annual Income Above \$100,000)		-0.0131 (0.0580)		0.0219 (0.0715)
Fixed Effects:				
List Fixed Effects	Yes	Yes	Yes	Yes
Wave Fixed Effects	Yes	Yes	Yes	Yes
R^2	0.008	0.042	0.011	0.013
N	307	307	291	291

Notes: The dependent variable is the aggregate donation (in \$) to a qualifying charity. Probit models. Average marginal effects reported. Robust standard errors are reported in parentheses.

*** denotes significance at the 1 percent level

** denotes significance at the 5 percent level

* denotes significance at the 10 percent level

Table 26: Effect of Information Treatment on Making a Donation to a Non-Qualifying Charity

	T2 vs. T3	T2 vs. T3	T4 vs. T5	T4 vs. T5
1(Information Treatment)	-0.0659 (0.0513)	-0.0727 (0.0509)	-0.2196*** (0.0605)	-0.2222*** (0.0612)
1(Female)		0.1260** (0.0519)		-0.0133 (0.0623)
1(Below Age 35)		0.0777 (0.0712)		0.1043 (0.0836)
1(Above Age 65)		-0.0598 (0.0616)		-0.1279* (0.0710)
1(Annual Income below \$50,000)		-0.0559 (0.0613)		0.0171 (0.0747)
1(Annual Income Above \$100,000)		-0.0533 (0.0648)		0.0224 (0.0747)
Fixed Effects:				
List Fixed Effects	Yes	Yes	Yes	Yes
Wave Fixed Effects	Yes	Yes	Yes	Yes
R^2	0.013	0.043	0.056	0.073
N	307	307	291	291

Notes: The dependent variable is the aggregate donation (in \$) to a non-qualifying charity. Probit models. Average marginal effects reported. Robust standard errors are reported in parentheses.

*** denotes significance at the 1 percent level

** denotes significance at the 5 percent level

* denotes significance at the 10 percent level

Note : The number of “useful incompletes” in our sample that we include for this analysis is 347. Thus the total sample size is 1251 (904+347). These are people who made the donation decision but dropped out before completing the experiment. We do not have demographic information for these people. Thus in column 2 with demographic controls we merely see the number of completes that is the sample for our regressions earlier. This serves as a quick check to see whether including the incompletes in column 1 leads to results different from Column 2 which comprises of completes only. Estimates from linear regressions are reported.

Tables 27-32 includes estimates including incompletes.

Table 27: Effect of Information Treatment on Donation

	Pooled	Pooled with Demographic Controls
1(Information Treatment)	-3.5697 (2.9481)	-6.7569* (3.5570)
1(Indicator for Treatment with Multiple Recipients)	3.4374 (2.4201)	0.4358 (2.9171)
Information Treatment × Multiple Recipients Indicator	3.6869 (3.5960)	6.5963 (4.3751)
1(Female)		6.8362*** (2.1068)
1(Below Age 35)		-1.2378 (2.7190)
1(Above Age 65)		5.7966** (2.5532)
1(Annual Income below \$50,000)		-9.4282*** (2.4667)
1(Annual Income Above \$100,000)		-3.4799 (2.5900)
Fixed Effects:		
Wave Fixed Effects	Yes	Yes
R^2	0.015	0.044
N	1,251	904

Notes: The dependent variable is the aggregate amount donated by an individual in each wave, pooled over treatments. Robust standard errors are reported in parentheses.

*** denotes significance at the 1 percent level

** denotes significance at the 5 percent level

* denotes significance at the 10 percent level

Table 28: Effect of Information Treatment on Participation (Likelihood of being a Donor)

	Pooled	Pooled with Demographic Controls
1(Information Treatment)	-0.0659* (0.0348)	-0.1031** (0.0450)
1(Indicator for Treatment with Multiple Recipients)	-0.0179 (0.0271)	-0.0428 (0.0351)
Information Treatment \times Multiple Recipients Indicator	0.0640 (0.0422)	0.0968* (0.0549)
1(Female)		0.0889*** (0.0261)
1(Below Age 35)		0.0514 (0.0323)
1(Above Age 65)		0.0178 (0.0319)
1(Annual Income below \$50,000)		-0.0748** (0.0307)
1(Annual Income Above \$100,000)		-0.0468 (0.0322)
Fixed Effects:		
Wave Fixed Effects	Yes	Yes
R^2	0.011	0.029
N	1,251	904

Notes: The dependent variable is a binary indicator which takes the value 1 when a positive amount is allocated for donation, and 0 otherwise. Robust standard errors are reported in parentheses.

*** denotes significance at the 1 percent level

** denotes significance at the 5 percent level

* denotes significance at the 10 percent level

Table 29: Effect of Information Treatment on Fraction of Donation to Qualifying Charities

	Pooled	Pooled with Demographic Controls
1(Information Treatment)	-0.0326 (0.0487)	-0.0707 (0.0574)
1(Indicator for Treatment with Multiple Recipients)	-0.1018*** (0.0370)	-0.1246*** (0.0435)
Information Treatment \times Multiple Recipients Indicator	0.1200** (0.0544)	0.1660** (0.0645)
1(Female)		0.0388 (0.0281)
1(Below Age 35)		0.0119 (0.0357)
1(Above Age 65)		0.0635* (0.0343)
1(Annual Income below \$50,000)		-0.0494 (0.0327)
1(Annual Income Above \$100,000)		-0.0504 (0.0344)
Fixed Effects:		
Wave Fixed Effects	Yes	Yes
R^2	0.011	0.021
N	1,251	904

Notes: The fraction is defined as donation to qualifying charities over the total donation of a subject. This is closer to a measure of conditional giving to qualifying charities. For subjects who did not donate at all, we code the fraction as zero, while keeping them in the sample.

Robust standard errors are reported in parentheses.

*** denotes significance at the 1 percent level

** denotes significance at the 5 percent level

* denotes significance at the 10 percent level.

Table 30: Effect of Information Treatment on Selecting a Qualifying/ Non-Qualifying Charity

	Qualifying Charity	Qualifying Charity with Demographic Controls	Non-qualifying Charity	Non-qualifying Charity with Demographic Controls
1(Information Treatment)	0.0474 (0.0400)	0.0596 (0.0450)	-0.2030*** (0.0464)	-0.2279*** (0.0561)
1(Female)		-0.0439 (0.0467)		-0.0808 (0.0563)
1(Below Age 35)		-0.0611 (0.0673)		0.1337* (0.0741)
1(Above Age 65)		-0.0180 (0.0524)		-0.0639 (0.0638)
1(Annual Income below \$50,000)		-0.0090 (0.0596)		0.0350 (0.0664)
1(Annual Income Above \$100,000)		0.0567 (0.0547)		0.0127 (0.0689)
Fixed Effects:				
List Fixed Effects	Yes	Yes	Yes	Yes
Wave Fixed Effects	Yes	Yes	Yes	Yes
R^2	0.014	0.037	0.098	0.139
N	411	291	411	291

Notes: This table includes T4 and T5 treatments only since these treatments allow the choice of more than one charity of similar type.

The dependent variable is a binary indicator which takes the value 1 when a qualifying(non-qualifying) charity is chosen, and 0 otherwise.

Robust standard errors are reported in parentheses.

*** denotes significance at the 1 percent level

** denotes significance at the 5 percent level

* denotes significance at the 10 percent level

Table 31: Effect of Information Treatment on Selecting a Specific Number of Qualifying Charities

	Chose Zero	Chose Zero	Chose One	Chose One	Chose Two	Chose Two
1(Information Treatment)	-0.0474 (0.0400)	-0.0596 (0.0450)	-0.1555*** (0.0472)	-0.1683*** (0.0562)	0.2030*** (0.0464)	0.2279*** (0.0561)
1(Female)		0.0439 (0.0467)		-0.1246** (0.0584)		0.0808 (0.0563)
1(Below Age 35)		0.0611 (0.0673)		0.0726 (0.0797)		-0.1337* (0.0741)
1(Above Age 65)		0.0180 (0.0524)		-0.0819 (0.0634)		0.0639 (0.0638)
1(Annual Income below \$50,000)		0.0090 (0.0596)		0.0260 (0.0683)		-0.0350 (0.0664)
1(Annual Income Above \$100,000)		-0.0567 (0.0547)		0.0694 (0.0710)		-0.0127 (0.0689)
Fixed Effects:						
List Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Wave Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.014	0.037	0.050	0.079	0.098	0.139
N	411	291	411	291	411	291

Notes: This table includes T4 and T5 treatments only since these treatments allow the choice of more than one charity of similar type.

The dependent variable is an indicator which takes values between 0 and 2 indicating the number of qualifying charity/charities being selected.

Robust standard errors are reported in parentheses.

*** denotes significance at the 1 percent level

** denotes significance at the 5 percent level

* denotes significance at the 10 percent level

Table 32: Effect of Information Treatment on Fraction of Donation to Qualifying Charities

	T2 vs. T3	T2 vs. T3	T4 vs. T5	T4 vs. T5
1(Information Treatment)	0.0657** (0.0255)	0.0793** (0.0313)	0.1082*** (0.0410)	0.1164** (0.0501)
1(Female)		0.0437 (0.0312)		0.0422 (0.0510)
1(Below Age 35)		0.0815** (0.0369)		-0.0560 (0.0645)
1(Above Age 65)		0.1022** (0.0432)		0.0242 (0.0597)
1(Annual Income below \$50,000)		-0.0656* (0.0362)		-0.0431 (0.0604)
1(Annual Income Above \$100,000)		0.0002 (0.0413)		0.0003 (0.0630)
Fixed Effects:				
List Fixed Effects	Yes	Yes	Yes	Yes
Wave Fixed Effects	Yes	Yes	Yes	Yes
R^2	0.027	0.071	0.038	0.054
N	417	307	411	291

Notes: This table includes T4 and T5 treatments only since these treatments allow the choice of more than one charity of similar type.

The dependent variable is an indicator which takes values between 0 and 2 indicating the number of qualifying charity/charities being selected.

Robust standard errors are reported in parentheses.

*** denotes significance at the 1 percent level

** denotes significance at the 5 percent level

* denotes significance at the 10 percent level

Figure 11: Information about the CTC

Since 1998, Arizona offers a one-to-one state income tax credit for donations to qualifying charitable organizations. A one-to-one income tax credit means that your state income tax owed will be reduced by the entire amount you choose to donate to a qualified charity up to a threshold described below. In other words, your donations simply redirect tax dollars owed to the state from the general tax fund to a qualified charity of your choice. For example, if you choose to donate \$X to a charity that is qualified for the state income tax credit, you can receive a refund of \$X when you file your state income tax return. Donations to qualified charities are "free" for you if you claim the tax credit.

The credit program is called Contributions to Qualifying Charitable Organizations and allows taxpayers to claim a credit of up to \$400 (\$800) for an individual (joint) filer for contributions to qualified 501(c)(3) organizations serving particular causes certified by the Arizona Department of Revenue. Qualifying organizations are defined as follows:

An individual income tax credit is available for contributions to Qualifying Charitable Organizations that provide assistance to residents of Arizona who receive Temporary Assistance of Needy Families (TANF) benefits, are low income residents of Arizona, or are children who have a chronic illness or physical disability. In addition, a Qualifying Charitable Organization may be considered a Qualifying Foster Care Charitable Organization if it meets additional criteria in serving qualified individuals.

A Qualifying Charitable Organization is a charity that meets ALL of the following provisions:

- *Is exempt from federal income taxes under Section a 501(c)(3) or is a designated community action agency that receives community services block grant program monies pursuant to 42 United States Code Section 9801.*
- *Provide services that meet immediate basic needs.*
- *Serves Arizona residents who receive temporary assistance for needy families (TANF) benefits, are low income residents whose household income is less than 150% of the federal poverty level, or are chronically ill or physically disabled children.*
- *Spends at least 50% of its budget on qualified services to qualified Arizona residents.*
- *Affirm that it will continue spending at least 50% of its budget on qualified services to qualified Arizona residents.*


A charity must apply for and meet all requirements of the law to be considered as a Qualifying Charitable Organization. Approved charities' names are listed on the Department of Revenue's website.

You can obtain information about how to file the tax credit at www.azdor.gov or by typing "Arizona charitable giving tax credit" in the search engine of your choice.

Please look at the lists carefully and choose your preferred charitable organizations. It is entirely up to you how you determine your preferred charities; you can choose any charity you would like, regardless of how much you know about the charity or if you have donated to it in the past. The lists consist of both qualified organizations (denoted with [qualified for the tax credit]) and other charities that are not qualified for the income tax credit. If you give to a qualified organization, your contributions are eligible for the tax credit. Your chosen charity will be recorded by the survey software and will be the potential recipient in the allocation task of stage two.

Notes: Subjects in treatments with information about the CTC see this information page before choosing a charity for the allocation task.

Figure 12: Example Receipt for Donation made to a Qualifying Charity

 ANDREW YOUNG SCHOOL OF POLICY STUDIES	EXPERIMENTAL ECONOMICS CENTER 404-413-0194 TEL 404-413-0195 FAX www.andrewyoungschool.org	14 Marietta Street NW Suite 454 Atlanta, Georgia 30303 Mail: Department of Economics P.O. Box 3992 Atlanta, Georgia 30302-3992
---	---	--

Date: 3/31/2018

Dear Participant:

Thank you for your contribution. You contributed a total of \$30:

\$30 to Southwest Behavioral Health Services [qualified for the tax credit]

100% of your contribution has been directed to the charity designated above. This charity is a I.R.C. § 501(c)(3) organization and a Qualifying Charitable Organization (QCO) on the Arizona Department of Revenue's QCO list for 2017.

This letter serves as a receipt of the payment of your contributions to your selected charity, Southwest Behavioral Health Services [qualified for the tax credit], by the research team at Georgia State University. You have the option to print this page or save a digital copy for your tax records.

Principal Investigator's Signature

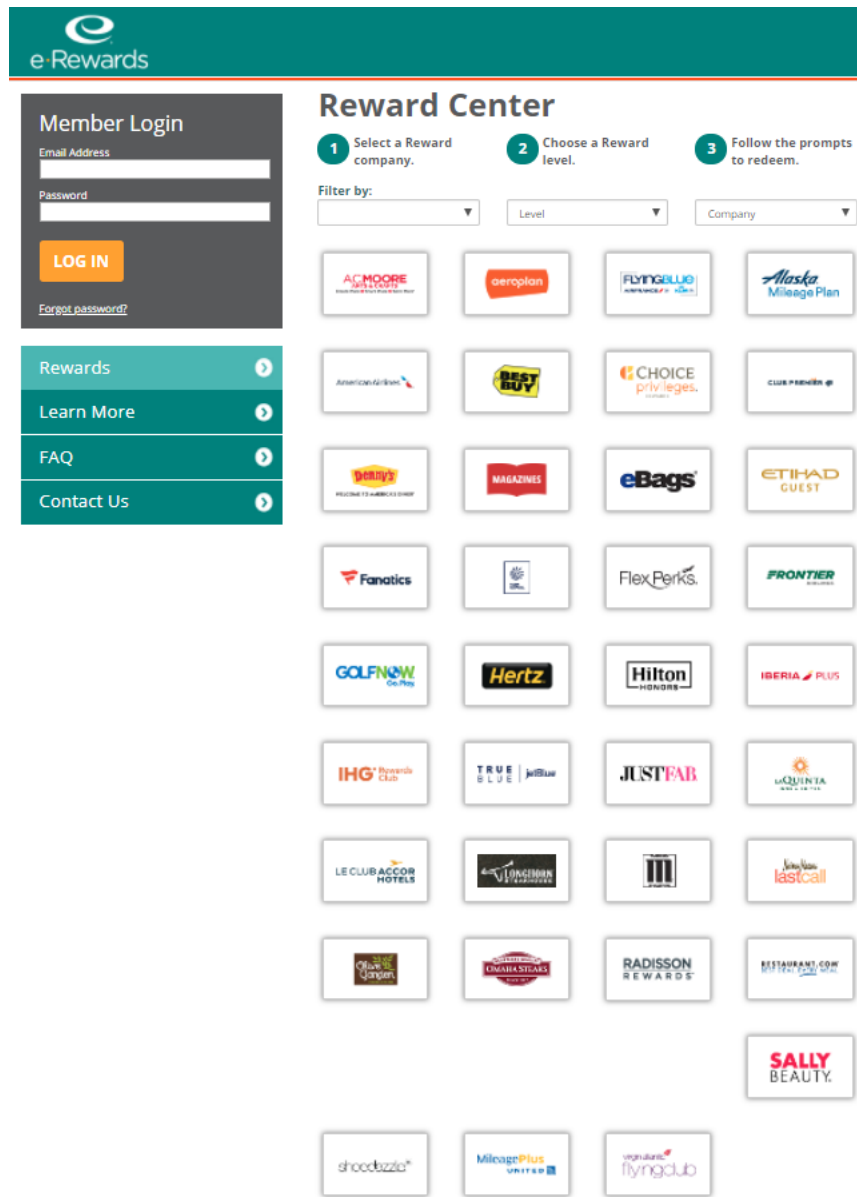
[Click Here to Print or Download to PDF](#)

Donor Information:

Please enter your name and address in this box if you want to print or save this page for your records. To proceed to the next stage of the survey, please delete the content of this box and enter 0.

Notes: Example receipt for donations made to a qualifying charity. The text box at the bottom of the page allows subjects to enter their personal information. Before proceeding to the next page of the survey, subjects are forced to delete any personally identifiable data. A click on the button opens a standard print window that provides functionalities for saving a PDF locally or printing the receipt directly. The receipt summarizes the donation amount, the recipient charity or charities, information about the donor, and information about the research team.

Figure 13: e-Rewards Currency Portal



Notes: Subjects received all show-up fees and payments based on their allocation decision in e-Rewards currency. e-Rewards can be used to purchase goods from online vendors.

Table 33: Charities in the Experiment, Part 1

Charity	Qualifies for the CTC	Mission Statement
Arizona Community Foundation	No	Lead, serve and collaborate to mobilize enduring philanthropy for a better Arizona.
Arizona Youth Partnership	No	Arizona Youth Partnership's mission is to partner with communities to cultivate healthy foundations for youth and promote strong families.
Arizona's Children Association	Yes	The mission of Arizona's Children Association is protecting children, empowering youth, and strengthening families.
Association for Supportive Child Care	No	Our mission is to enhance the quality of care for children in Arizona.
Childhelp	No	Childhelp exists to meet the physical, emotional, educational, and spiritual needs of abused, neglected and at-risk children. We focus our efforts on advocacy, prevention, intervention, treatment and community outreach.
Community Food Bank	Yes	We change lives in the communities we serve by feeding the hungry today, and building a healthy, hunger-free tomorrow.
Feed My Hungry Children	No	Feed My Hungry Children helps stand in the gap to provide the things that needy, hurting people may need to survive and become self-sufficient. Feed My Hungry Children's humanitarian projects are committed to helping children and their families around the world.
Food for the Hungry	No	Together we follow God's call responding to human suffering and graduating communities from extreme poverty.
Hospice of the Valley	No	Comfort and dignity as life nears its end.
Make-A-Wish Foundation	No	We grant the wishes of children with life-threatening medical conditions to enrich the human experience with hope, strength, and joy.
NARBHA	No	Managing integrated health care with a conscience. To be recognized as the innovative leader in managing superior behavioral health care.
Phoenix Rescue Mission	Yes	Providing Christ-centered, life-transforming solutions to persons facing hunger and homelessness.

Table 34: Charities in the Experiment, Part 2

Charity	Qualifies for the CTC	Mission Statement
Pima Council on Aging	Yes	Our mission is to provide dignity and respect for aging, and to advocate for independence in the lives of Pima County’s older adults and their families, now and for generations to come.
Pima Prevention Partnership	No	Building partnerships with young people, families, and communities to improve their quality of life.
Southwest Behavioral Health Services	Yes	We inspire people to feel better and reach their potential. Through helping people discover their strengths, we improve our communities.
Southwest Human Development	Yes	Southwest Human Development strengthens the foundation Arizona’s children need for a great start in life.
St. Mary’s Food Bank	Yes	St. Mary’s Food Bank serves to alleviate hunger through the gathering and distribution of food while encouraging self-sufficiency, collaboration, advocacy and education.
Teen Lifeline	Yes	To provide a safe, confidential, and crucial crisis service where teens help teens make healthy decisions together.
Tucson Urban League	Yes	The mission of the Tucson Urban League is to advance economic and social prosperity for African Americans and other underserved Tucson area residents by creating access to opportunity through advocacy, community partnerships, and programs and services.
United Food Bank	Yes	The mission of United Food Bank is to provide access to nutritious food for those who are without - by servicing as a community bridge between those who want to help and those who are in need.

Table 35: Charity Lists by Treatment

Subgroup 1	Subgroup 2
<i>Pair 1: Baseline and Treatment 1</i>	
Arizona Community Foundation	Arizona's Children Association
Hospice of the Valley	Arizona Youth Partnership
Make-A-Wish Foundation	Association for Supportive Child Care
NARBHA	Childhelp
Phoenix Rescue Mission	Community Food Bank
Pima Council on Aging	Feed My Hungry Children
Pima Prevention Partnership	Food for the Hungry
Southwest Behavioral Health Services	Teen Lifeline
Southwest Human Development	Tucson Urban League
St. Mary's Food Bank	United Food Bank
<i>Pair 2: Treatment 2 and Treatment 3</i>	
<i>List 1</i>	
Phoenix Rescue Mission	Arizona Youth Partnership
Pima Council on Aging	Association for Supportive Child Care
Southwest Behavioral Health Services	Childhelp
Southwest Human Development	Feed My Hungry Children
St. Mary's Food Bank	Food for the Hungry
<i>List 2</i>	
Arizona Community Foundation	Arizona's Children Association
Hospice of the Valley	Community Food Bank
Make-A-Wish Foundation	Teen Lifeline
NARBHA	Tucson Urban League
Pima Prevention Partnership	United Food Bank
<i>Pair 3: Treatment 4 and Treatment 5</i>	
<i>List 1</i>	
Hospice of the Valley	Arizona's Children Association
Pima Prevention Partnership	Arizona Youth Partnership
Southwest Behavioral Health Services	Childhelp
Southwest Human Development	Community Food Bank
St. Mary's Food Bank	Feed My Hungry Children
<i>List 2</i>	
Arizona Community Foundation	Association for Supportive Child Care
Make-A-Wish Foundation	Food for the Hungry
NARBHA	Teen Lifeline
Phoenix Rescue Mission	Tucson Urban League
Pima Council on Aging	United Food Bank

Notes: Within each treatment, subjects were subsequently assigned to one of two subgroups that differed in terms of the charity list(s) from which subjects could choose the recipient(s) in the allocation task. Each treatment pair contained the same lists of charities.

References

- Abraham, J., & Royalty, A. (2005). Does having two earners in the household matter for understanding how well employer-based health insurance works? *Medical Care Research and Review*, 62 (2), 167–186.
- Abramowitz, J. (2016). Saying, “I Don’t” The Effect of the Affordable Care Act Young Adult Provision on Marriage. *The Journal of Human Resources*, 51.4.0914-6643R2.
- Andreoni, J., & Payne, A. A. (2003). Do government grants to private charities crowd out giving or fund-raising? *American Economic Review*, 93(3): 792-812.
- Andreoni, J., & Payne, A. A. (2011a). Is crowding out due entirely to fundraising? Evidence from a panel of charities. *Journal of Public Economics*, 95, 334-343.
- Andreoni, J., Payne, A.A, & Smith, S. (2014). Do grants to charities crowd out other income? Evidence from the UK. *Journal of Public Economics*, 114, 75-86.
- Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics: An empiricist’s companion*. Princeton, NJ: Princeton University Press.
- Arcidiacono, P., Lovenheim, M., & Zhu, M. (2015). Affirmative action in undergraduate education. *Annu. Rev. Econ.*, 7(1), 487-518.
- Auten, G. E., Seig, H. & Clotfelter, C. (2002). Charitable giving, income, and taxes: An analysis of panel data. *American Economic Review*, 92(1): 371–82.
- Bagde, S., Epple, D., & Taylor, L. J. (2011). Dismantling the legacy of Caste: Affirmative Action in Higher Education. *mimeo*.
- Balafoutas, L. & Sutter, M. (2012). Affirmative action policies promote women and do not harm efficiency in the laboratory. *Science*, 579-582.
- Barkowski, S., & McLaughlin, J. S. (2020). In sickness and in health: Interaction effects of state and federal health insurance coverage mandates on marriage of young adults. *Journal of Human Resources.*, 55(2).
- Beaman, L., Chattopadhyay, R., Duflo, E., Pande, R. & Topalova, P. (2009). Powerful women: Does exposure reduce bias? *The Quarterly Journal of Economics*, 124(4), 1497-1540.
- Becker, G. (1973). A Theory of Marriage: Part I. *The Journal of Political Economy*, 813-846.
- Bracha, A., Cohen, A. & Conell-Price, L. (2019). The heterogeneous effect of affirmative action on performance. *Journal of Economic Behavior and Organization*, 158, 173-218.

- Buchmueller, T. C., & Valletta, R. G. (1999). The Effect of health insurance on married female labor supply. *The Journal of Human Resources*, 34(1):42–70.
- Burnstein, N. R. (2007). Economic influences on marriage and divorce. *Journal of Policy Analysis and Management*, 26, 387–429.
- Card, D., & Giuliano, L. (2016). Universal screening increases the representation of low-income and minority students in gifted education. *Proceedings of the National Academy of Sciences*, 113(48), 13678-13683.
- Casamiglia, C., Franke, J., & Rey-Biel, P. (2013). The incentive effects of affirmative action in a real-effort tournament. *Journal of Public Economics*, 15-31.
- Cason, T. N., & Zubrickas, R. (2018). Crowdfunding for public goods with refund bonuses: An empirical and theoretical investigation. *Working Paper* https://krannert.purdue.edu/faculty/cason/papers/cont_time_crowdfunding.pdf
- Cassan, G. (2019). Affirmative action, education and gender: Evidence from India. *Journal of Development Economics*, 136, 51-70.
- Chen, T. (2018), Health insurance coverage and marriage behavior: Is there evidence of marriage-lock? *Working Paper*, https://tianxuchen.weebly.com/uploads/2/2/8/1/22817854/marriage_lock_tianxu_chen_12282018.pdf
- Chen, T. (2019), Health Insurance and Marriage Behavior: Will Marriage Lock Hold Under Healthcare Reform? *Working Paper*, https://tianxuchen.weebly.com/uploads/2/2/8/1/22817854/chen_ma_reform_04152019.pdf
- Clotfelter, C. (1980). Tax incentives and charitable giving: evidence from a panel of taxpayers. *Empirical Economics*, 13(3), 319-40.
- Coate, S., & Loury, G. (1993). Will affirmative-action policies eliminate negative stereotypes? *The American Economic Review*, 83(5), 1220-1240.
- Corazzini, L., Cotton, C., & Valbonesi, P. (2015). Donor coordination in project funding: Evidence from a threshold public goods experiment. *Journal of Public Economics*, 128, 16-29.
- Courtemanche, C., Marton, J., Ukert, B., Yelowitz, A., & Zapata, D. (2017). Early impacts of the Affordable Care Act on health insurance coverage in Medicaid expansion and non-expansion states. *Journal of Policy Analysis and Management*, 36(1), 178-210.
- Cunha, F., & Heckman, J. J. (2007). The Evolution of Inequality, Heterogeneity and Uncertainty in Labor Earnings in the US Economy. *NBER working papers* 13526.
- Dechenaux, E., Kovenock, D., & Sheremeta, R. M. (2015). A survey of experimental research on contests, all-pay Auctions and tournaments. *Experimental Economics*, 18(4), 609-669.

- De Zwart, F. (2015). Unintended but not unanticipated consequences. *Theory and Society*, 44(3), 283-297.
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550.
- Duquette, N. J. (2016). Do tax incentives affect charitable contributions? Evidence from public charities' reported revenues. *Journal of Public Economics*, 137, 51-69.
- Dynarski, S. (2017). Simple Way to Help low-income students: Make everyone take SAT or ACT. *New York Times*, Retrieved from <https://www.nytimes.com/2017/07/14/upshot/how-universal-college-admission-tests-help-low-income-students.html>
- Eckel, C. C., Grossman, P. J., & Johnston, R. M. (2005). An experimental test of the crowding out hypothesis. *Journal of Public Economics*, 89(8), 1543-1560.
- Ek, C. (2017). Some causes are more equal than others? The effect of similarity on substitution in charitable giving. *Journal of Economic Behavior and Organization*, 136, 45-62.
- Epstein, R. A. (1995). *Forbidden grounds: The case against employment discrimination laws*. Harvard University Press.
- Feldstein, M. & Clotfelter, C. (1976). Tax incentives and charitable contributions in the united states: A microeconomic analysis. *Journal of Public Economics*, Vol. 5(1-2), 1-26.
- Fack, G., & Landais, C. (2010). Are tax incentives for charitable giving efficient? Evidence from France. *American Economic Journal: Economic Policy*, 2(2), 117-41.
- Filiz-Ozbay, E., & Uler, N. (2019). Demand for giving to multiple charities: An experimental study. *Journal of the European Economic Association*, 17(3), 725-753.
- Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 13(4):399-411.
- Fryer Jr, R. G., & Loury, G. C. (2013). Valuing diversity. *Journal of Political Economy*, 121(4), 747-774.
- Gibson-Davis, C. M., Edin, K., & McLanahan, S. (2005). High hopes but even higher expectations: The retreat from marriage among low-income Couples. *Journal of Marriage and Family*, 67:1301–12.
- Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in competitive environments: Gender differences. *The Quarterly Journal of Economics*, 118(3), 1049-1074.

- Gruber, J. (2011). The impacts of the Affordable Care Act: How reasonable are the projections? *National Bureau of Economic Research, Working paper no. 17168*.
- Harrison, G. W., Harstad, R. M., & Rutström, E. E. (2004). Experimental methods and elicitation of values. *Experimental Economics*, 7(2), 123-140.
- Heckman, J. J., LaLonde, R. J., & Smith, J. A. (1999). The economics and econometrics of active labor market programs. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3A, Chapter 31, pp. 1865–2097. New York: North-Holland
- Hyman, J. (2017). ACT for all: The effect of mandatory college entrance exams on postsecondary attainment and choice. *Education Finance and Policy*, 12(3), 281-311.
- Heim, B., Lurie, I., & Simon, K. (2018). The impact of the Affordable Care Act young adult provision on childbearing: evidence from tax data. *Demography*, 55(4), 1233-1243.
- Hoffman, S., & Duncan, G. (1995). The effect of incomes, wages, and AFDC benefits on marital disruption. *The Journal of Human Resources*, 30(1), 19-41.
- Holzer, H. J., & Neumark, D. (2000). What does affirmative action do? *ILR Review*, 53(2), 240–271.22
- Hu, W.Y. (2003). Marriage and economic incentives : Evidence from a welfare experiment. *The Journal of Human Resources*, 942-963.
- Kaestner, R., Garrett, B., Chen, J., Gangopadhyaya, A., & Fleming, C. (2017). Effects of ACA Medicaid expansions on health insurance coverage and labor supply. *Journal of Policy Analysis and Management*, 36(3), 608-642.
- Karaca-Mandic, P. , Norton, E.C. , Dowd, B. (2012). Interaction terms in nonlinear models. *Health Services Research*, 47 (1), 255–274.
- Kidd, M. P., Carlin, P. S. & Pot, J. (2008), Experimenting with affirmative action: The Coate and Loury Model. *Economic Record*, 84, 322-337.
- Kingma, B. R. (1989). An accurate measurement of the crowd-out effect, income effect, and price effect for charitable contributions. *Journal of Political Economy*, 97(5), 1197-1207.
- Lavelle, B., & Smock, P. J. (2012). Divorce and women’s risk of health insurance loss. *Journal of Health and Social Behavior*, 53, 413–431.
- Leonard, J. (1990). The impact of affirmative action regulation and equal employment law on black employment. *The Journal of Economic Perspectives*, 4(4), 47-63.
- Lillard, L.A. & Panis, C.W.A. (1996). Marital status and mortality: The role of health . *Demography* 33: 313.

- List, J. A. (2004). Young, selfish and male: Field evidence of social preferences. *The Economic Journal*, 114(492), 121-149.
- List, J. A. (2011). The market for charitable giving. *Journal of Economic Perspectives*, 25(2): 157-80.
- List, J. A. & Price, M. K. (2012). Charitable giving around the world: Thoughts on how to expand the pie. *CESifo Economic Studies*, 58(1), 1-30.
- List, J. A., Sadoff, S., & Wagner, M. (2011). So you want to run an experiment, now what? Some simple rules of thumb for optimal experimental design. *Experimental Economics*, 14(4), 439.
- Loury, G. C. (1997). How to mend affirmative action. *Public Interest*, (127), 33-44.
- Meer, J. (2014). Effects of the price of charitable giving: Evidence from an online crowdfunding platform. *Journal of Economic Behavior and Organization*, 103, 113-124.
- Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much?. *The Quarterly Journal of Economics*, 122(3), 1067-1101.
- Niederle, M., & Vesterlund, L. (2010). Explaining the gender gap in math test scores: The role of competition. *Journal of Economic Perspectives*, 24(2), 129-44.
- Niederle, M., Segal, C., & Vesterlund, L. (2013). How costly is diversity? Affirmative action in light of gender differences in competitiveness. *Management Science*, 1-16.
- Null, C. (2011). Warm glow, information, and inefficient charitable giving. *Journal of Public Economics*, 95, 455-465.
- Payne, A. A. (1998). Does the government crowd-out private donations? New evidence from a sample of non-profit firms. *Journal of Public Economics*, 69(3), 323-345.
- Peters, E., Simon, K., & Taber, J. (2014). Marital Disruption and Health Insurance. *Demography*, 51:1397-1421.
- Puhani, P.A. (2012). The treatment effect, the cross difference, and the interaction term in nonlinear "Difference-in-Differences" models. *Economics Letters*, 115 (1), 85-87.
- Randolph, W. C. (1995). Dynamic income, progressive taxes, and the timing of charitable contributions. *Journal of Political Economy*, 103(4), 709-38.
- Robles, V. C. F., & Krishna, K. (2012). Affirmative Action in Higher Education in India: Targeting, Catch Up, and Mismatch at IIT-Delhi. *National Bureau of Economic Research Working Paper*, (17727), 1-44.
- Rodgers, W.M., & Spriggs, W.E. (1996). What does the AFQT really measure? Race, wages, schooling, and the AFQT score. *Review of Black Political Economy*, 24, 13-47

- Schotter, A., & Weigelt, K. (1992). Asymmetric tournaments, equal opportunity laws, and affirmative action: Some experimental results. *The Quarterly Journal of Economics*, 107(2), 511-539.
- Sohn, H. (2015). Health Insurance and Risk of Divorce: Does Having Your Own Insurance Matter? *Journal of Marriage and Family*, 77, 982–995.
- Simon, K. , Soni, A. & Cawley, J. (2017), The Impact of health insurance on preventive care and health behaviors: Evidence from the first two Years of the ACA Medicaid expansions. *Journal of Policy Analysis and Management*, 36, 390-417.
- Sutter, M., & Rützler, D. (2010). Gender differences in competition emerge early in life. *IZA Discussion Paper No. 5015*.
- Wellington, A. J., & Cobb-Clark, D. A. (2000). The labor supply effect of universal health coverage: What can I learn from individuals with spousal coverage? *Research in Labor Economics*, 19, 315–44.
- Yelowitz, A. (1998). Will extending Medicaid to two-parent families encourage marriage? *The Journal of Human Resources*, 33 833-865.
- Zimmer, D. (2007). Asymmetric effects of marital separation on health insurance among men and women. *Contemporary Economic Policy*, 25, 92–106.

Vita

Chandrayee Chatterjee was born in Kolkata, India. She received her Bachelor's degree with Honors in Economics from Presidency College, University of Calcutta in 2010. She went on to complete her Master's degree in Economics from the Centre for Economic Studies and Planning, Jawaharlal Nehru University, New Delhi. From 2012 to 2014, Chandrayee worked at the Hongkong and Shanghai Banking Corporation in Kolkata in strategy analysis and portfolio management.

Chandrayee joined the Department of Economics in the Andrew Young School of Policy Studies of Georgia State University in 2014. Her research uses experimental methods and behavioral economics to address unintended consequences of public policies. She has received the Andrew Young School Dissertation Fellowship as well as a fellowship from the Center for Economic Analysis of Risk in the Robinson College of Business of Georgia State University.

In August 2020, Chandrayee will be joining the Department of Economics and Finance at Southern Utah University as a faculty.