CERGE - EI
Center for Economic Research and Graduate Education
–
Economics Institute

Essays in Experimental Economics

Theodor Kouro

Dissertation

Prague 2026

Dissertation Committee:
Ole Jann (CERGE-EI, Chair)
Michal Bauer (CERGE-EI)
Štěpán Jurajda (CERGE-EI)
Jan Zápal (CERGE-EI)

Referees:
Sebastian Kube (University of Bonn)
Suanna Oh (Paris School of Economics)

Contents

Acknowledgements

I am deeply grateful to my supervisor, Ole Jann, for walking me through the world of academia and helping me understand my value in it, often in difficult times. I am very grateful to my advisor, Michal Bauer, for inspiring me to run field experiments, challenging me intellectually, and pushing my creativity to other levels. I thank Nikolas Mittag, Byeongju Jeong, and the rest of my dissertation committee members: Štěpán Jurajda, Jan Zápal, and Andreas Menzel for their time and effort throughout my research. I am grateful to my wonderful co-authors, Enkelejda Havari, Rebecca Dizon-Ross, and Esmeralda Shehaj Zhabjaku, for their patience during my journey of learning how to bring big field partners on board and run field experiments that change the world; it would not have been the same without them. I thank the members of the Academic Skills Center for carefully editing this thesis. I thank John List for inviting me to his team at UChicago and for his honest feedback. I thank Elidona Prenga, Vasjana Tanka, and Mateo Jovani for their great research assistance. I thank all the partners in the field and the Ministry of Education for supporting my field work in Albania. I thank Marta Cota for her great support throughout our common PhD journey and beyond, and all my friends at CERGE-EI. Last but not least, I thank my parents, my sister, and my brother for their invaluable support. I dedicate this work to the memory of my grandmother. This output was supported by: Charles University, GAUK project No 322521, PRIMUS Research Programme of Charles University, PRIMUS/20/HUM/019, the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 870245, and the National Agency for Scientific Research and Innovation in Albania.

Prague, Czech Republic Theodor Kouro

February, 2026

Abstract

This thesis explores behavioral and rational mechanisms underlying people's prosociality, effort provision, and human capital decisions. Chapter 1 examines whether and why charitable giving increases when donors have more choice about how their donations are used, specifically the option to allocate their gift to three projects as they prefer. Partnering with a charity and many firms, I show that allowing donors to direct their gifts increases giving, primarily because donors can better target causes they value. Chapter 2 examines the mechanics of autonomy in a large-scale natural field experiment with volunteers on effort provision. Allowing volunteers to choose their tasks rather than being randomly assigned increases effort by 33% and improves effort quality. I provide causal evidence that this effect is driven by ability-task matching rather than intrinsic preferences for choice, task attachment, or increased commitment, suggesting that autonomy can be a cost-effective alternative to performance-based monetary incentives. Chapter 3 documents parental misperceptions at the time of secondary school track choice. Combining administrative data with a large-scale survey of parents in Albania, I show that parents overestimate their own child's academic prospects and underestimate those of similarly performing peers, which helps explain high enrollment in academic tracks among students unlikely to meet college eligibility thresholds. In contrast, beliefs about vocational education returns are broadly accurate, highlighting the role of information frictions in educational choices.

Introduction

These three chapters explore behavioral and rational mechanisms underlying people's prosociality, effort provision, and human capital decisions. In the first two chapters, I focus on the role of autonomy/agency in prosocial behavior, focusing on charitable giving and volunteering efforts, and emphasizing the behavioral mechanisms that drive agents to increase their giving and effort. In particular, these two chapters try to distinguish between two possible explanations of the effect of agency/autonomy on money and time (effort) donation. Do people give more or exert more effort because they value the intrinsic value of agency/autonomy, or do they use it as a self-selection mechanism? The experimental designs allow me to separate these two explanations, and conclude that people primarily use agency and autonomy to self-select into preferences (e.g., donors selecting projects they like more) and abilities (e.g., volunteers selecting tasks they are good at). In the third chapter, I show that Albanian parents systematically overestimate their own child's academic prospects while underestimating those of their peers, potentially contributing to high enrollment in academic tracks among students unlikely to meet college eligibility thresholds.

Chapter one examines whether charitable giving increases if donors have more choice about how their donations are used. In a field experiment, employees of large Albanian companies were asked to donate to projects administered by Down Syndrome Albania. Treatments varied in whether participants were allowed (or forced) to choose between different projects, and in the amount of information they were given. Giving donors a choice substantially increased giving; information did not. Our setting allows us to consider various mechanisms that could underlie this behavior. We conclude that allocation choice mainly increases donations because donors can target projects they like.

Chapter two explores the mechanics of autonomy in a large-scale natural field experiment of effort provision with volunteers, and in collaboration with a UNICEF-funded NGO. I show that allowing volunteers to choose tasks rather than randomly assigning them increases effort by 33% and significantly improves effort quality. While this autonomy effect is unlikely to persist, the results suggest that it is cost-effective to motivate effort through autonomy rather than ex-post, competitive, and performance-contingent monetary rewards. Lastly, I provide causal evidence that the increase in effort occurs because volunteers match their abilities with the tasks and not because they value choice, care more about certain tasks, or feel a greater commitment to perform well when the choice is theirs.

Chapter three documents substantial parental misperceptions about academic trajectories and college eligibility at the time of secondary school track choice. Using administrative data combined with a large-scale survey of parents in Albania, we show that parents systematically overestimate their own child's future academic performance and chances of meeting college eligibility thresholds, while being more pessimistic about similarly-performing peers. These misperceptions coexist with high enrollment in academic high schools, even among students who are unlikely to be college-eligible. In contrast, parents' beliefs about vocational education returns and stigma appear broadly accurate. The findings highlight the potential role of information frictions in shaping secondary school track choices.

1 Allocation Choice in Charitable Giving: A Natural Field Experiment

1.1 Introduction

Despite the lack of data on global charitable giving, in the US charitable donations worth around $200-$400 billion have been made in the last decade (Giving USA, 2020). These donations play a role in both channeling funds to the poorest and weakest and providing a valuable service to those who want to do good but have no personal connection to those in need (Ariely et al., 2009). Although the experimental literature on charitable giving has introduced a gamut of key considerations for day-to-day fundraising practices (see Jasper and Samek (2014), for a review), only a few studies explore the impact of various donation choices on prosocial behavior. In particular, the experimental literature on offering donors a choice about how their donation will be used is limited. This aspect of fundraising is important because it supports the idea that community engagement in decision-making may help improve economic outcomes. For example, a fundraiser who seeks to maximize charitable donations might collect more funds if she forgoes a part of her flexibility in allocation decisions. On the other hand, it is natural for donors to like some projects more than others, and satisfying these preferences may increase charitable donations.

The most common donation choices are those between different recipients (Cryder et al., 2013; Aretz and Kube, 2013) and different charities that may or may not vary in the type of their recipients (Eckel et al., 2017; Heist and Cnaan, 2018). These choice treatments make it almost impossible to distinguish whether any positive treatment effect is attributed to preferences for having a choice or for the options associated with the choice. For example, consider donor A, who has the choice to donate to charity (recipient) X or charity (recipient) Y, and donor B, who has no choice but is asked to donate to a randomly chosen charity (recipient). Suppose donor A gives more than donor B. In that case, she does so because either she (1) likes to have a choice regardless of taking it, (2) she identifies herself more with recipient X or Y, (3) has stronger preferences for charity X or Y, (depending on the characteristics of the charities including revenues, overhead costs, and reputation, among others) or (4) she values the openness of the fundraiser which increases trust in general. It is a compelling task to disentangle these four explanations.

In this paper, we build upon the concept of empowering donors with choice and engaging them in the fundraisers' decision making by introducing a novel donation strategy: a choice to allocate gifts to three different projects that benefit the same recipients of a charity, i.e., helping the same individuals in three different ways. Providing donors with the opportunity to make a decision on how their donation will be used may not only increase giving, but may also help donors to gain more meaning from their donation experience (Whillans, 2016). Hence, we question whether offering donors an allocation choice would increase giving. Since all donations go to the same charity and the same type of recipients, we can test, in a causal way, different mechanisms that could explain the positive effect of the allocation choice. Specifically, do donors give more because they like to have a choice or because they have stronger preferences for particular projects over other projects? The

former is consistent with increased agency, and the latter implies that donors are incentivized to increase giving because each dollar of their gift will go to the project they like more, relative to donors who do not have a choice, for which only a third of the dollar goes to support the preferred project.

We further explore the possibility that donors may like to engage in allocation decisions, but they may not feel sufficiently informed about the impact of the projects. This motivates the question of whether giving increases if we offer donors a chance to make more-informed allocation decisions. To answer these questions, we designed and implemented a natural field experiment of donations to children with Down Syndrome in Albania. The experimental design consisted of three treatments: (1) a choice to allocate donations between therapy, humanitarian aid, and entertainment; (2) a forced-allocation decision; and (3) a forced allocation decision with a chance to acquire more information about the aid-effectiveness of each of these projects.

Previewing the results of our study, offering donors an allocation choice increased donations by 80.9%, and 74.55% took the allocation option. However, when we took away the choice and forced donors to allocate, mean donations doubled, suggesting that the large effect of the allocation choice was not causally driven by donors liking to have a choice. We also found that the mean donations in each of the three projects were different, and most donors chose to allocate their gifts unequally. Hence, the effect is causally driven by the fact that through allocation, donors can donate more to the projects they like more. Since we found no effect on the likelihood of donating, it is unlikely that trust plays a role. Lastly, introducing an information link increased donations relative to the control group by 52.3%, which is lower than the shift in the forced allocation treatment. This result suggests that allocation and information on aid effectiveness do not work hand in hand. To shed more light on why the provision of more information in addition to allocating donations lead to a decrease in donations, and to obtain feedback on the fundraising campaign, we conducted an incentivized follow-up survey with participants and non-participants of the main experiment. The results of the follow-up survey are consistent with our findings.

This paper is closely related to four natural field experiments that empowered donors with more agency. Eckel et al. (2017) asked the alumni of a US university to choose to donate to a general fund or a restricted fund that benefits the students of the program from which they graduated. Conditional on taking the choice, donors could allocate a part or all of their contribution to the restricted fund. They found significant effects of the choice on donations, and only a few donors availed themselves of the fund choice. While it is tempting to attribute this treatment effect to increased agency or to donors liking to have a choice, such an interpretation would not be causal since taking up the choice is entirely endogenous. Our experimental design is different in that it isolates the choice between two charities or different recipients, allowing donors to help the same individuals in three different ways. This makes it possible to test whether the positive effect on generosity is attributable to donors liking to have a choice.

Heist and Cnaan (2018) asked participants in the treatment group to choose one out of ten charities to donate to and assigned one randomly chosen charity to participants in the control group. They found a positive effect of having a choice of charity on giving, although

it is impossible to disentangle preferences for choice from preferences for charities. Kessler et al. (2019) gave a sense of choice to the alumni of a university in the US over how donated funds would be used. The treated alumni could choose one out of four projects based on what they thought was more important for students. They found significant treatment effects only among rich and powerful donors. While in the study by Kessler et al. (2019) donors might have been uncertain whether their choice would affect the allocation of funds, in this study, we empower donors with agency rather than a sense of agency.

Design-wise, we build on the study by Aretz and Kube (2013), who asked donors whether they wanted to choose their object of benevolence. Conditional on taking up the choice, they could choose one to five recipients, but the amount would be split equally among those chosen unless they picked a single recipient who would receive the entire gift. Our treatment is different since it allows donors to allocate their gifts among three projects as they prefer. Moreover, and to the best of our knowledge, this is the first study that evaluates the impact of a forced choice on donations. This was suggested, but not addressed by Aretz and Kube (2013).

This study contributes to the broad literature on increasing charitable giving by examining allocation choice as an unexplored fundraising strategy that increases giving. Specifically, we add to the growing literature on how donated funds will be used (Gneezy et al., 2014; Kessler et al., 2019). Further, our experimental design allows the separation of several mechanisms that may underlie the effect of the allocation choice. We provide causal evidence that donors do not increase giving because they like to have a choice but because they have stronger preferences for some projects than others. To the best of our knowledge, this is the first study to show that donors increase giving even when the fundraiser forces them to make allocation decisions rather than allowing them to choose.

The findings of this study also relate to the literature on directed giving by showing that the vast majority of donors avail themselves of the allocation choice, in contrast to the results of Eckel et al. (2017) and Aretz and Kube (2013), which found that only a few donors chose to direct their gifts. This suggests that the choice take-up rate may depend on the framing and strength of the treatment. Moreover, our findings relate to the underresearched literature on the role of information about aid effectiveness in giving. Karlan and Wood (2017) find that information about aid effectiveness has a positive effect on the contributions of large prior donors, while it harms the contributions of small prior donors, with no effect overall. On the other hand, Metzger and Günther (2019) find no impact of information about aid impact on average donations. While this study does not focus on the causal impact of information about aid effectiveness on giving, it adds to this literature by concluding that information on aid effectiveness does not work hand in hand with allocating donations.

This study is also related to the literature on tax compliance, in which allowing citizens to allocate a certain percentage of their tax payment to government spending categories has been shown to increase tax compliance (Lamberton, 2013; Lamberton et al., 2014). Further, it relates to the volunteering literature, in which Mertins and Walter (2021) found that volunteers produced more output when they had a chance to vote on how the money that was raised would be spent. Lastly, we add to the growing literature on detecting misperceptions

about others (Bursztyn and Yang, 2022). Similar to the recent study by Drouvelis and Marx (2022), we find that people are overly optimistic about the charitable donations of others.

1.2 Experimental Design

1.2.1 Setting

In this project, we partnered with the Down Syndrome Albania Foundation (DSA), a member of Down Syndrome International and the European Down Syndrome Association, among other reputable international NGOs. Dedicated to promoting and supporting the right of acceptance, inclusivity, and integration of children with Down Syndrome in Albanian society, the DSA is a fundraising organization that aims to facilitate the therapeutic treatment, entertainment, and humanitarian needs of children with Down Syndrome and other intellectual disabilities. The setting consists of donations to children with Down Syndrome (DS) in Albania.

We used the list of companies that regularly support the DSA and their employees were our potential donors. We prepared a detailed invitation letter that was sent to the HR departments of these firms along with detailed instructions about the implementation of the project. Overall, we recruited 22 out of the 38 firms that we approached, with a total of over 5000 employees (excluding production/field workers). While the board of each corporation was fully informed about the scientific purpose of this project, the employees were not. Randomization happened at the individual level, and in most cases, we performed it on the HR managers' computers so that we would have no access to the e-mail distribution list of these firms. In the remaining cases, the HR departments willingly performed the randomization on their own using written and video instructions that we sent them in advance, and documented the process afterwards.

In the next step, the HR departments sent four randomized group e-mails,[1] which contained an introduction to the cause and an online survey link, which differed according to the four experimental conditions. The introduction to the e-mail made it clear that their firm was collaborating with the DSA to raise awareness and funds to support children with Down Syndrome in Albania. We employed a web developer to design the survey links through a Google server, which was built into the website of the DSA, via a sub-domain. This choice make the survey look more professional, it helped to eliminate doubts that the survey link and donation process were scams.

It is worth noting that if an employee clicked the survey link, she could observe the treatment section after filling in the initial survey questions. The first ten questions of each survey link were identical and intended to collect information on gender, age, civil status, parental status, education, job position, and past donation behavior. Through the rest of the questions, we asked employees whether they would like to donate and, if so, what the size of their gift would be. The survey was fully anonymous, and its questions were not sensitive. The ethical approval of the experiment was also sent to the HR departments, along with the invitation to participate in the project.

[1]The translated version of the e-mail can be found at Appendix A.6 in Figures 35 and 36

Further, to give donors more flexibility in giving, the donation links remained active for roughly two weeks. During this period, the HR departments sent 2-3 kind reminder emails, according to the response rate. While the timing of the first e-mail was random, the kind reminder e-mails were programmed during low workload days. In the following sub-section, we describe the four experimental conditions.

1.2.2 Experimental conditions

The experimental design consisted of four experimental conditions and measured two outcomes: the likelihood to donate (the extensive margin) and the size of the gift (the intensive margin). Participants were randomly assigned to one of the following experimental conditions. Before making any decision, they were invited to read three short sentences about the three projects that benefit children with Down Syndrome: therapy, entertainment, and humanitarian aid. They then had to decide whether they wished to donate and, conditional on agreeing to donate, were asked to write down the amount of their gift and proceed to the bank transfer section of the website. These decisions were expected to vary across experimental conditions (see also Figure 1):

C Participants assigned to the control group (**C**) made two decisions: whether to donate and how much to donate. They were not offered an allocation choice or asked to complete an allocation task. However, they knew their donation would fund all three non-mutually-exclusive projects.

T1 Participants assigned to the first treatment (**T1**) were asked to choose whether they preferred to allocate their contribution to three projects themselves or delegate this decision to the NGO. Conditional on taking the choice, they had to allocate their donation to one or a combination of the following projects: therapy, humanitarian aid, and entertainment. If they did not take the allocation option, they were asked to indicate their donation amount. We label this treatment as *allocation choice*.

T2 The second treatment (**T2**) was the *forced-allocation*, in which, conditional on agreeing to donate, participants had to allocate their contribution to one, or a combination of the three projects. Their only way out was not to donate.

T3 The third treatment group (**T3**) was identical to **T2** up to the addition of a link that participants could click if they wished to read detailed information about the charity projects and how these projects impact children with Down Syndrome — clicking the link and reading the information means that they can make more-informed allocation decisions. We label this treatment as *forced-allocation and link* or simply *link*.

To ensure that the treatments affected the donation decisions at both the extensive and intensive margins, the treatment information and both decisions appeared at the same time on the screen.

Note that the objective of the last treatment, T3, is to address a complementary policy question: whether donors value allocative control more when it is accompanied by an opportunity to require additional information.

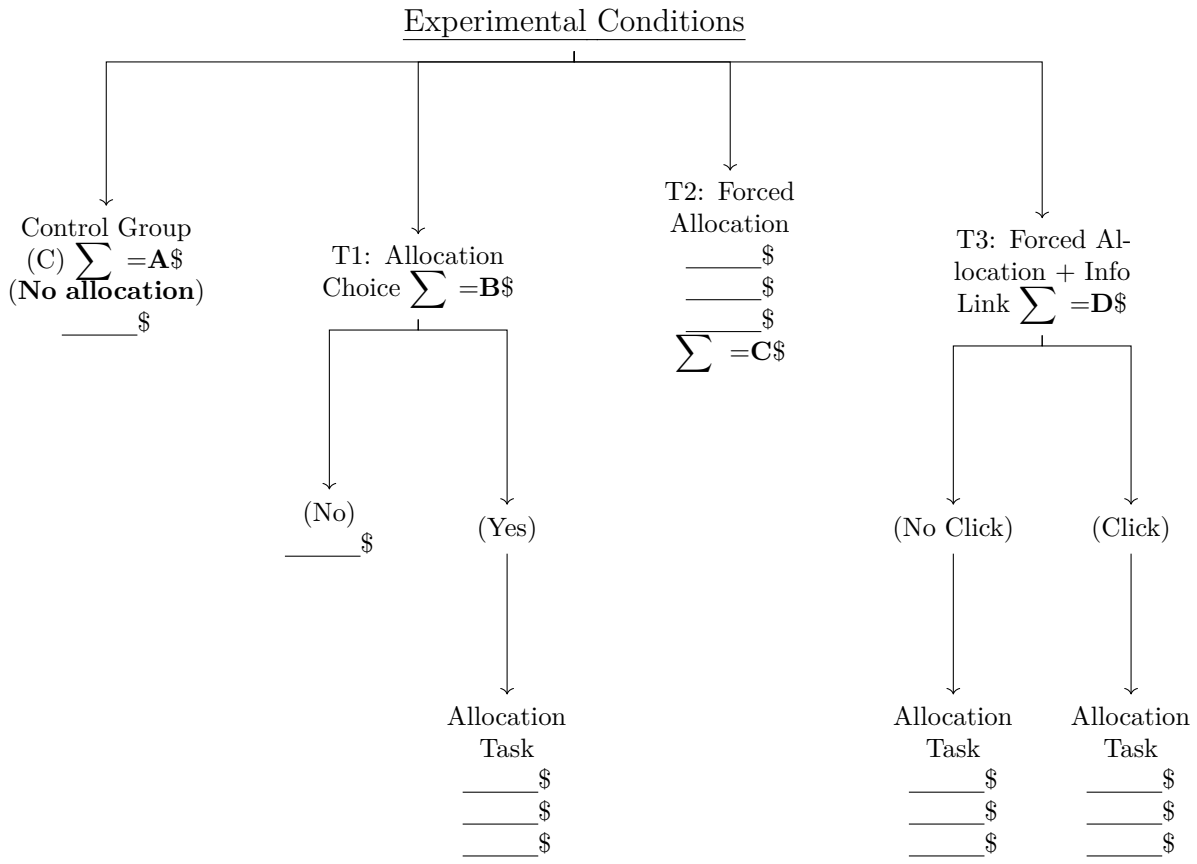


Figure 1: Experimental design

Note: The summation symbol represents the total donated amount. In the "*no allocation*" condition, each underline accompanied by a dollar symbol corresponds to the total donated amount. By contrast, in the "*allocation task*" condition, the total donated amount is the sum of the three donation amounts allocated across projects; each underline represents the amount donated to a specific project.

1.3 Mechanisms

Experimentally, it is a compelling task to distinguish whether an increase in giving from donors who have allocation choice is attributable to preferences for having a choice or preferences for the options associated with it. For example, some donors may care about how they wish to donate, in the sense that they may like to distribute donations to several projects as they prefer, while others may want to delegate this task to the fundraiser. The fact that donors have the freedom to choose how they wish to donate may lead to higher giving. We define this potential mechanism as *preference for having a choice*, which is consistent with increased agency.

Alternatively, donors may have *preferences for the projects* they can donate to. Through the allocation option, donors can clearly state these preferences and are thus likely to increase

their giving. For example, a donor who has allocation choice and likes one project more than the other two can donate the whole dollar to that project, and thus she is likely to give more. By contrast, a donor who does not have the allocation choice expects that only a third of that dollar will go to that project. In this section, we describe how we plan to disentangle these two mechanisms.

We designed a control group and two treatment groups: allocation choice and forced-allocation. Under the former condition, participants may choose whether to allocate their gift to three projects. Under the latter condition, they must allocate their donations or choose not to donate as their only way-out option. The preference for having a choice mechanism would be at work if those forced to allocate their gifts respond negatively by donating less than those who do not have any allocation option and less than those who can choose to allocate donations. Alternatively, if forcing people to allocate increases giving relative to the control group but does not necessarily increase donations relative to the allocation choice group, the preference for projects mechanism would be driving the effect of the allocation choice. It is worth noting that taking the allocation choice is entirely endogenous and cannot be used as a tool for disentangling mechanisms. The share of donors who take the allocation option suggests whether donors like to allocate. The second mechanism is valid if we show that donors have stronger preferences for some projects and choose to give more to those projects.

1.3.1 Alternative mechanisms

Here we discuss a few other non-behavioral mechanisms that may mediate the effect of allocation choice. For example, people may donate more when they have a choice to allocate their gifts because the charity may seem more transparent/open to them. Further, allocating donations may reduce the uncertainty over a possible mismanagement of the charity's funds by the fundraiser. While donors have preferences for impactful giving (Cryder et al., 2013; Aknin et al., 2012), mismanagement is a problem since it might reduce the impact of donations. It has been shown in laboratory experiments that donors give less when there is a greater risk that their donation will have less impact (Krawczyk and Le Lec, 2010; Brock et al., 2013; Exley, 2016; Garcia et al., 2020). To isolate mechanisms related to trust or transparency, which are well-explored in the literature, donors were asked to donate to an NGO that receives annual gifts from these companies. Further, they knew that the fundraising campaign was an institutional collaboration between the charity and their employer. It is unlikely that donors perceived the engagement by the fundraiser as a signal of her incompetency in using the donated funds. These and other mechanisms related to trust would be more relevant in contexts where the charity has a bad reputation for wasting funds. We provide additional evidence from a follow-up survey that mechanisms related to trust are unlikely to play a role.

1.4 Results

1.4.1 Sample characteristics and randomization check

This section presents the results of this study. First, we describe a few characteristics of the overall sample. Only 1042 employees responded to the e-mails, a roughly 20% response rate. Table 4 in Appendix A.1 presents the response rate by treatment. The forced-allocation group had a lower participation rate of 18.64%, which is statistically different from that of the control (p-value = 0.026) and link (p-value = 0.065[2]) groups, but statistically indistinguishable from the participation rate of the allocation choice group (p-value = 0.160). Since none of the participants could observe the treatment without answering the initial survey questions, the small-in-size differential attrition is random.

Sample characteristics are displayed in Table 5 in Appendix A.1. The mean age of participants is 36, and 58% of them are females. Regarding civil status, 54.4% are married, 35.7% are single, 7.2% cohabit, and only 2.7% are divorced. Parents make up 55.6% of the sample. 69% of the employees have a master's degree, 9.5% have a doctoral degree, 18.3% are graduates, and only 2.5% are undergraduates. Concerning job positions, 63.2% are specialists, 28.9% are managers, and 7.87% are executives. While 40.3% were sufficiently aware of people with intellectual disabilities (PID), 57.5% had previously donated to support them. 60.2% were well aware of the Down Syndrome Albania Foundation, but only 12.4% had previously donated to the NGO. Table 6 in Appendix A.1 displays the results of the OLS regression of the likelihood to donate on the baseline characteristics. Older and more educated employees are more likely to donate, and gender does not matter in their decision to give. Those employed in high-role jobs are also more likely to donate. While having donated in the past to PID increases the likelihood of donating, past donors to the DSA are equally as likely to donate as new donors. Interestingly, those who feel sufficiently aware of PID are less likely to donate, and being aware of the NGO does not affect the decision to donate.

We present our randomization check in Table 7 in Appendix A.2. Given multiple treatments, we ran three least square regressions on three dummies that take the value 0 if an employee is assigned to the control group and 1 if the employee is assigned to one of the other three treatments. We find that the share of those who decided not to specify their gender is smaller in the choice group than in the control group. However, in testing for mean-equality among both groups, we fail to reject the mean-equality hypothesis (Table 8 in Appendix A.2). Similarly, the share of master's degree holders is slightly lower in the choice group than the control group but not statistically different. The share of divorced employees is statistically smaller in the forced-allocation and forced-allocation and link groups (confirmed by Table 9 and Table 10 in Appendix A.2). Nevertheless, the p-values of the joint-orthogonality tests suggest that randomization was implemented successfully and that any differences among the control and treatment groups are random.

[2]The cut-off p-value for statistical significance throughout this work is 0.05.

1.4.2 Main results

We now turn to the treatment effects on giving, displayed in Figure 2 and Table 1. First, offering donors a choice to allocate their contributions among three projects increases donations relative to the control group by 80.9% (the control mean is 810.9 Lek ($7.15) and the allocation choice mean is 1466.9 Lek ($12.93)). Second, forcing donors to allocate doubles giving relative to the control group. Third, allowing donors to make more informed allocation decisions (link group) increases giving by 52.3%, and 33.6% clicked and confirmed that they read the information on the aid-effectiveness of these projects. These results align with the distribution of the gift size by treatment, shown in Figure 24 in Appendix A.3: relative to the control group, there are fewer small and more large donations by treated donors.

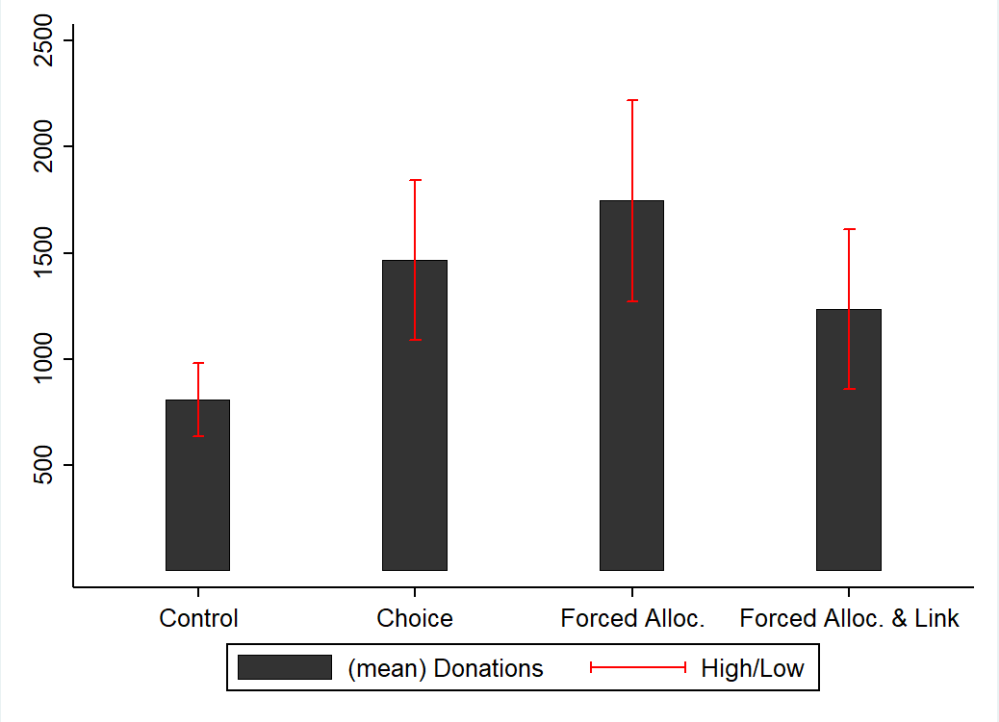


Figure 2: Donation means by treatment group

Moreover, all three effects are robust to controlling for baseline characteristics, past donation behavior, and firm fixed effects. We replicated these results using winsorized donations (Table 11 in Appendix A.3) to check whether our results are robust to outliers. While the effects of allocation choice and forced-allocation are stable when we perform winsorization at the 1%, 2%, and 5% levels, the effect of the link treatment vanishes with the 5% winsorization. The treatment effects in Table 1 are also robust to multiple hypothesis testing, and using clustered standard errors at the firm level does not harm our results (Table 12 in Appendix A.3).

We further explore the treatment effect on the likelihood to donate, i.e., the extensive margin (Figure 3 & Table 2). The overall donating rate among those who participated

Table 1: Treatment effects on giving

Treatment	(1) Donation	(2) Donation	(3) Donation	(4) Donation
Choice	656.564***	682.073***	664.291***	734.148***
	(210.386)	(224.478)	(213.316)	(226.895)
Forced Allocation	934.770***	957.512***	938.081***	999.018***
	(255.054)	(255.846)	(247.176)	(252.017)
Forced Alloc. & Link	424.674**	512.078**	427.483**	490.501**
	(210.111)	(226.496)	(201.617)	(210.343)
Control mean	810.369	810.369	810.369	810.369
Firm FE	No	Yes	No	Yes
Covariates	No	No	Yes	Yes
Observations	1,042	1,042	1,042	1,042
Choice = Forced: p-value	0.364	0.368	0.370	0.381
Forced = Link: p-value	**0.096**	0.159	0.088	0.099

Note: The first column of this table shows the OLS results of donations on each of the treatment dummies, where the base category is the control group. The remaining columns replicate these results controlling for firm fixed effects (column 2), controls (column 3), and both (column 4). The currency is in Albanian Lek and 1 Lek = 0.0088 $. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

voluntarily in the experiment is 43.7%. The link group has the lowest share of donors, and the forced-allocation group has the highest.

Table 2: Treatment effects on the likelihood to donate

Treatment	(1) Likelihood	(2) Likelihood	(3) Likelihood	(4) Likelihood
Choice	-0.021	-0.022	-0.024	-0.022
	(0.043)	(0.043)	(0.041)	(0.042)
Forced Alloc.	0.051	0.049	0.064	0.062
	(0.044)	(0.043)	(0.043)	(0.042)
Forced Alloc. & Link	-0.046	-0.015	-0.043	-0.021
	(0.042)	(0.043)	(0.041)	(0.042)
Control Mean	0.442	0.442	0.442	0.442
Firm FE	No	Yes	No	Yes
Covariates	No	No	Yes	Yes
Observations	1,042	1,042	1,042	1,042

Note: The first column of this table shows the OLS results of the likelihood to donate on each of the treatment dummies, where the base category is the control group. The rest of the columns replicate these results, controlling for firm fixed effects (column 2), controls (column 3), and both (column 4). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

The results displayed in Table 2 show that none of the treatments has a causal effect on

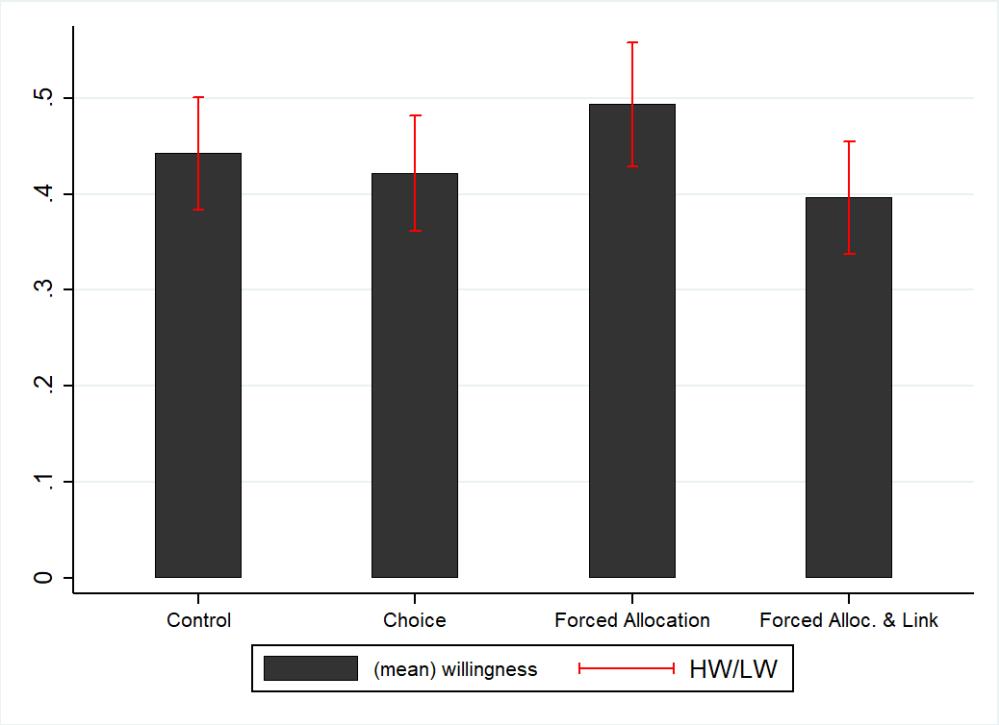


Figure 3: Mean likelihood to donate by treatment group

the likelihood of donating. This result is typical in the experimental literature on charitable giving, suggesting that in most natural fundraising contexts, donors have made up their minds to donate before they face the treatment, meaning that the treatment manipulation does not affect their behavior on the extensive margin (Aretz and Kube, 2013; Eckel et al., 2017; Kessler et al., 2019). Given that there is no treatment effect on the extensive margin, we disregard the treatment effect on donations, conditional on giving (Table 13 in Appendix A.4).

1.4.3 Understanding the effect of the allocation choice

In the previous section, we established a strong effect of allowing donors to allocate gifts among three projects. In this section, we intend to understand what drives this effect. Treatment comparisons suggest that offering donors a choice to allocate their gifts did not lead to more giving than forcing them to allocate. Moreover, 74% of the donors in the choice group availed themselves of the allocation choice. In addition, those who were forced to allocate did not massively choose the way-out option, i.e., not to donate, since there were no differences in the likelihood of donating among the control, choice, and forced-allocation groups. Therefore, donors did not increase their giving because they like to have a choice before donating, i.e., to allocate or not. Conversely, taking away that choice and forcing them to allocate did not harm giving.

Table 3 presents the average donation in each of the three projects. We focus on the choice and forced-allocation columns because donors in the link group might have formed pref-

erences from receiving more information on the aid effectiveness of these projects. Donors in each treatment group had stronger preferences for therapy and donated less for humanitarian aid and entertainment. Moreover, the majority allocated their gift unequally among these projects. Therefore, the effect of the allocation choice is driven by the fact that through allocation, participants donated more to their preferred projects. In other words, if they preferred therapy to the other two projects, each dollar of their donation went to support therapy, which incentivized them to increase the size of their gift relative to the control group, in which only a third of donation would on expectation, go to therapy. Nevertheless, while the results show that the intrinsic value for choice is not at work, the experimental design still cannot cleanly determine whether the allocation choice effect is driven by a preference for projects or by a preference for allocative control.

Table 3: Mean donation to each project by treatment

Projects	Choice	Forced Allocation	Link
Therapy	1499.54	1704.37	1462.43
Humanitarian Aid	1210.74	895.66	901.73
Entertainment	1072.94	935.78	752.30
% of unequal allocations	56.1%	64.3%	61.7%
Obs.	82	115	107

Note: Columns 2-3 indicate the mean donations in each project conditional on giving. The penultimate row shows the percentage of unequal allocation decisions among the three projects. An unequal allocation is any unequal split of a dollar among the three projects. The currency is in Albanian Lek.

The treatment effect of the allocation choice is unlikely to be channeled through an increase in the trust level, because if the treatment increased the average trust level we would also find significant differences at the extensive margin, i.e., the likelihood to donate (Table 2). We further exclude the possibility of anchoring effects because the projects were listed in alphabetical order and were not numbered. In addition, the description of the cause and the projects was unbiased towards a particular project, meaning that the recipients of their gifts would benefit equally from all three projects. While we have argued that trust is not likely to drive the treatment effect, it is more appropriate to ask directly some of the participants of the main experiment and other employees from companies that decided not to participate whether they believe trust plays a role. We designed an online incentivized survey that aims to obtain feedback about the results of this study and the fundraising campaign in general. We present the results of the follow-up survey in subsection 1.5.

1.4.4 Understanding the Role of Information on Aid-Impact

In Table 1, we showed that donors in the link treatment donated on average 52.3% more than donors in the control group, a shift in the mean donation that is lower relative to the forced-allocation treatment. There is also a difference in mean donations among the forced and link groups at the 10% level, suggesting that information on aid effectiveness and allocating donations do not work hand-in-hand. In Table 14 and Figure 25 in Appendix A.4,

we test for mean equality of donations among several sub-treatment groups, i.e., information clickers/non-clickers, those who took the option to allocate, and those who did not. These comparisons show a negative correlation between receiving more information on aid effectiveness and giving. In particular, those who clicked the link donated on average 1073 Lek less than those forced to allocate (p-value = 0.053). Information clickers and non-clickers donated on average more than donors in the control group, but the difference in donations is larger and more significant among non-clickers. The results of the least-squares prediction of clicking the information link (Table 15) suggest that employees with a master's degree and those who felt sufficiently aware of PID were less likely to click the link. This means that more-informed/aware donors did not find it worth receiving more information about how these projects help children with Down Syndrome. The R^2 of the estimated regression suggests that the observables likely explain 17.3% of the variation in clicking the link.

1.5 Follow-up Survey

For a better understanding of our results, we conducted a follow-up feedback survey with participants and non-participants of the main experiment. The motivation behind this survey was to understand why the information link harmed donations relative to the group that was forced to allocate donations. Moreover, through this survey, we addressed the possibility that the allocation choice increased donors' trust in the NGO as a possible alternative explanation for the increased giving. The screenshots of the online survey can be found in Appendix A.5.

It was difficult for some firms to resend a survey link to their employees, but fortunately, some did. We also approached firms that did not participate in the main experiment. We did so because the characteristics of their employees, including wage, age, education, and occupation, among others, were expected to be similar to those who participated in the main experiment. The sample included verified marketing experts and academics in the marketing field in order to receive more professional feedback. We assumed that experts and academics are more knowledgeable about predicting donors' behavior.

We provided a weak incentive to fill in the online survey by revealing the study's main results right after participants completed the survey. We shared the survey in a marketing expert group on Facebook, which contained 35 verified experts, and asked the HR departments of new firms and some of the firms from the first stage to share it with their employees. Lastly, the NGO sent an e-mail to all those who shared their e-mail address in the main experiment. Overall, we received 200 feedback responses. As shown in Table 16, 28% of the respondents were marketing experts, and 52% had taught marketing as a course in universities or professional training. While 53% were aware of the fundraising campaign, only 33.5% participated as donors.

A possible explanation of the negative effect of clicking the link is that the information diluted the donors' excitement to donate to their preferred projects, for example if a donor preferred the entertainment or humanitarian projects over therapy, but perceived therapy to be the most needed project after reading the information in the link. In such cases, the donors' expectations regarding the aid-effectiveness of the projects in the link group were

mismatched with the actual effectiveness of the projects, inducing them to give less. We disregard the possibility that the information was framed such that it would harm donations, but we cannot omit the possibility that donors did not understand the information in the link box simply because it might have been difficult. Further, we question whether donors felt a greater cognitive burden by first reading and processing the information and then making allocation decisions. Accordingly, we asked respondents whether the donation process in the link group might have been tiring. The order of these questions was randomized.

To identify the respondents' perceptions of the role of trust associated with the allocation choice, we directed them in the follow-up survey to choose why those who had a choice to allocate donations donated less or more than those who did not. We designed the survey following the methodology of DellaVigna and Pope (2018), which consisted of asking behavioral experts to evaluate the behavior of participants in an online effort provision experiment. Experts were asked to guess the average effort level in different treatment groups based on the average effort level in the baseline treatments. Similarly, we asked employees from the firms that participated in the main experiment, and from firms who did not, to guess the average donation in the choice, forced-allocation, and link treatments in random order. Respondents stated their confidence level for each guess on a scale from 1 to 10. Before asking respondents to guess the average donations in each treatment group we provided the actual texts from the main experiment to bring them as close as possible to the decision-making environment of those who took part in the main experiment. The texts included the initial invitation, the treatment description, the control group's average donation, and the information in the link box.

Figure 4 presents the actual and guessed mean donations for each treatment. First, we observe a general overestimation of prosocial behavior in terms of mean donations. The overestimation is more considerable and almost identical for the allocation choice and link treatments and smaller for the forced-allocation treatment. The average confidence level for each guess falls between 5.5-6 out of 10. Second, the guessed mean donations in the allocation choice and forced-allocation treatments are statistically indistinguishable. This result is in line with the comparison of actual mean donations among these treatment groups.

Third, the guessed mean donation in the forced-allocation treatment is statistically larger than that in the link group. Figures 32 - 34, display the guessed mean donations for each treatment by the type of evaluator, i.e., expert, academic, aware of the campaign, and donor status, along with the actual mean donation from the main experiment. There is an overestimation of mostly similar sizes by all types of evaluators for all treatment conditions. Regardless of the treatment, actual donors tended to guess closer to the actual mean donations, and academics tended to overestimate more than other respondents.

We then turned to understanding the possible channel through which the information link affected donations. We asked respondents to evaluate the information in the link box before guessing the average donation in this treatment. From Figure 5, we observe that 74.5% thought that the information was in line with donors' expectations about the aid-effectiveness of the projects, suggesting that the drop in the treatment effect is not likely attributable to the fact that the information would surprise donors negatively. This evaluation is consistent with their guesses: although they believed the information was well-matched with donors'

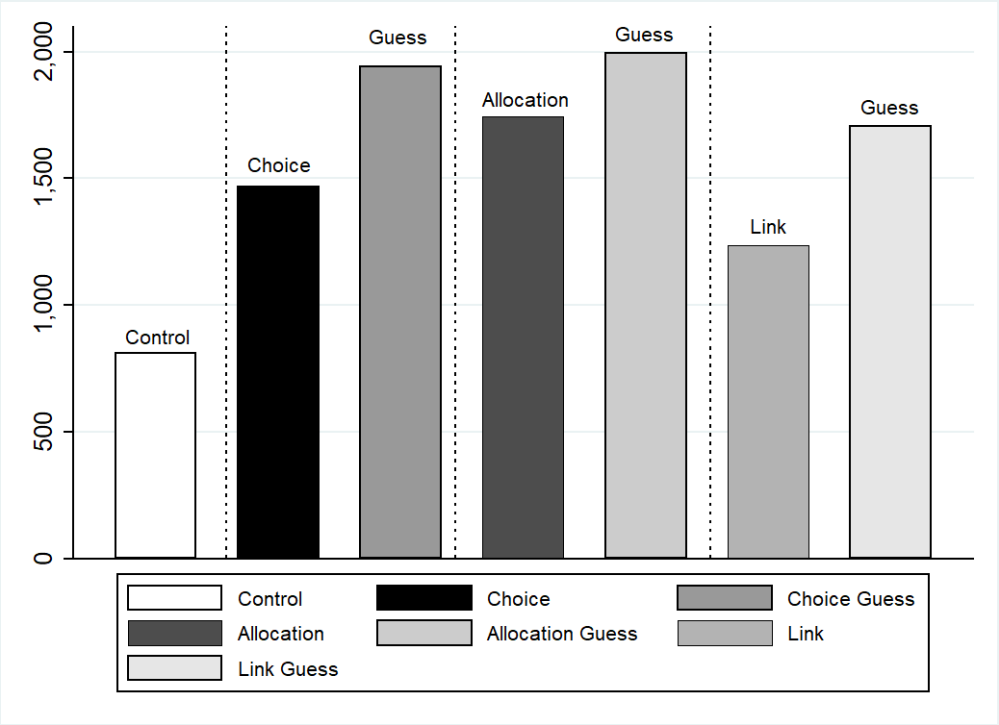


Figure 4: Actual and guessed mean donations for each treatment

Notes: Each bar labeled in the chart area as "Choice", "Allocation" and "Link" shows the actual mean donations for each treatment group. Each bar next to the treatment bars labeled as "Guess" represents the guessed mean donation for each treatment group. The short-dashed lines separate the actual and guessed mean donations by treatment from each other. For convenience the labels "Choice", "Allocation" and "Link" stand for allocation choice, forced-allocation and forced-allocation and link, respectively.

expectations, there are still differences between the guessed mean donations in the forced-allocation and link treatments.

In a randomized order, we asked participants to evaluate whether the information might have been difficult to understand, and 85% thought this was not the case. Finally, we asked respondents whether they believed it might have been tiring for donors to read the information and then make allocation decisions. 69.5% found the donation process in the link treatment tiring, and their guesses were much closer to actual donations in the forced-allocation and link treatments than those who thought otherwise. This set of results suggests that it is not the information content that was harmful to donations but rather the cognitive burden induced by processing information and making allocation decisions. Therefore, fundraisers may find it more beneficial to offer donors information about aid effectiveness separately from asking them to make allocation decisions.

Lastly, we describe the respondents' beliefs regarding the mechanism driving the effect of the allocation choice on donations (Figure 6). If the respondents' guesses of the mean donation in the choice group fell below the average donation in the control group, they were asked to justify their guesses by selecting the alternative "donors do not like to have a choice" or writing down another explanation. If their guess was above the average

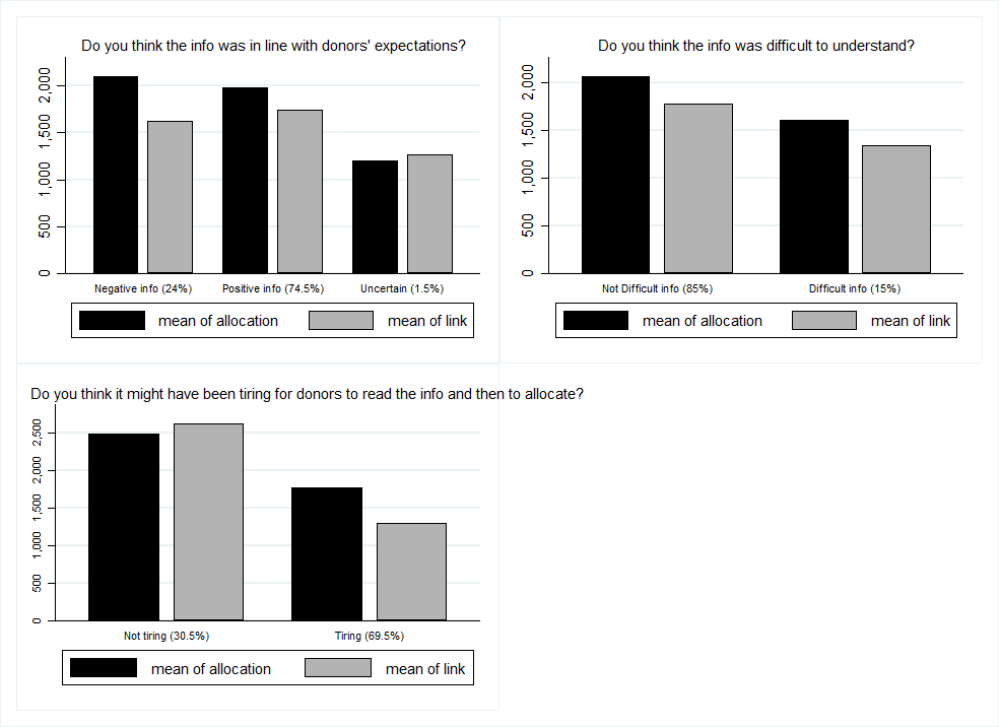


Figure 5: Understanding the backfire of the information link

Notes: Each bar chart shows the guessed mean donations in the forced-allocation and forced-allocation and link treatments, labeled as "allocation" and "link", respectively, for each option of the questions displayed in the title of the charts. The percentage figures represent the share of respondents who chose those options in the survey.

donation in the control group, they could choose one of the three explanations, i.e., "donors like choice," "donors could donate more to the projects they liked more," "donors trusted the NGO more," or write any other explanation. To cancel anchoring effects, we randomized the order of these potential explanations.

First, only 8.5% believed that the allocation choice would harm giving, justifying this with the notion that donors do not like to have a choice. However, the majority, (91.5%), guessed correctly that the average mean donation in the choice group would be greater than the mean donation in the control group. Only 6% of this majority believed that the effect of the allocation choice was driven by the notion that donors like to have a choice, and 14% stated that trust is the mechanism at work. These beliefs are consistent with their guesses: the guessed average donations in the choice treatment were above those in the forced-allocation treatment. Importantly, 70.5% of all participants believed the preference story that the allocation choice allows donors to donate more to the projects they like more. We further allowed donors to state any other reason they believe the allocation choice would lead to more giving. However, only 1% chose to do so, because their reason differed from the options provided.

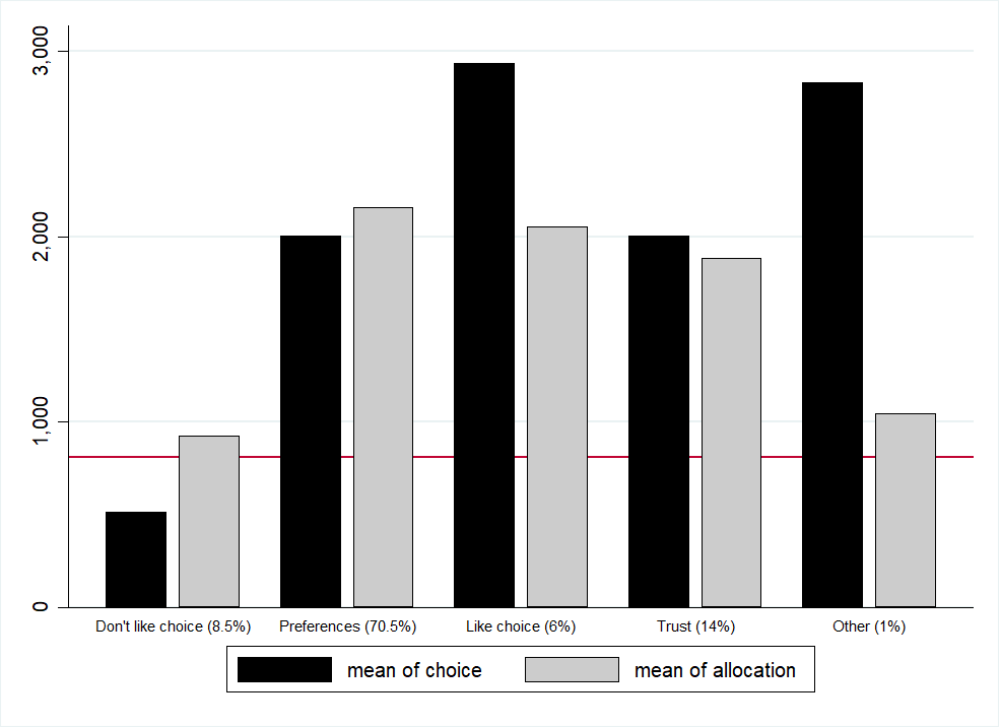


Figure 6: Respondents' beliefs about the mechanism driving the effect of the allocation choice

Notes: Each bar chart shows the guessed mean donations in the allocation choice and forced-allocation treatments, labeled as "choice" and "allocation", respectively, for each option of the question "why do you think donors donated less/more?". The percentage figures represent the share of respondents who chose the displayed options. The red line represents the mean donation in the control group.

1.6 Conclusion

In this paper, we explore a novel fundraising strategy: a choice to allocate donations to three projects that benefit the same type of recipients. We also explore the effect of forcing donors to allocate and allowing them to make more-informed allocation decisions. Through a natural field experiment of donations to children with Down Syndrome, we find that offering donors an allocation choice impacts donations significantly. The vast majority of those offered a choice took the allocation option. Moreover, forcing donors to allocate doubled donations.

Combining these results, we reject the hypothesis that donors increase their donations because they like to have a choice. We find that by allocating donations, donors can give more to the projects they prefer, thus increasing average donations relative to those who cannot allocate. We further provide evidence that other mechanisms related to trust are unlikely to channel this effect. For instance, we do not find differences in the likelihood of donating between the control and choice groups. If the allocation choice increased the general trust in the NGO, donors would respond positively by being more likely to donate.

Further, given that the fundraising campaign consisted of an institutional collaboration between the fundraiser and the donors' employers, it is unlikely that donors perceive

the engagement by the fundraiser as a signal of her incompetency to use the donated funds. Therefore, trust-related mechanisms would be more relevant when the charity has a reputation for wasting/mismanaging funds. Lastly, through a follow-up survey, we provide evidence that over 70% of respondents believed the effect is consistent with a preference story.

Allowing donors to make more-informed allocation decisions increased giving relative to a control group. However, their giving decreased relative to donors who were forced to allocate without the information link option. Hence, the allocation of donations and information provision on aid effectiveness do not work hand in hand in this setting. A the follow-up survey suggests it is likely that the link and allocation treatment induced greater cognitive burden on donors, making the donation process tiring.

This study contributes to the broad literature on increasing prosocial behavior in the domain of charitable giving by introducing a novel strategy to increase donations: a choice to allocate gifts to several projects. More specifically, it contributes to the literature on how donated funds will be used and the literature relating donors' choices to charitable donations. Unlike other experimental studies in this literature, our experimental design allows for a separation of two mechanisms underlying the effect of allocation choice on donations: preference for having a choice and preferences for the options associated with the choice. To the best of our knowledge, this is the first study that tests the impact of a forced-choice and provides causal evidence that donors react positively even when forced to make allocation decisions. Moreover, this study contrasts an established result in the literature of directed giving that suggests most donors do not avail themselves of choice, whereas we find evidence that most donors did so. Thus suggests that the choice-take-up rate depends on the framing and strength of the treatment.

Further, this study adds to the literature relating information about aid-effectiveness to charitable giving by showing that information on aid-effectiveness and allocating donations do not work hand in hand. Lastly, we add to the growing literature on detecting misperceptions about others (Bursztyn and Yang, 2022). Similar to Drouvelis and Marx (2022), we find that donors are overly optimistic about the charitable donations of others.

One limitation of this study is its inability to show whether the established treatment effects would persist over time. Measuring the persistence of the effects would be possible if we partnered with an NGO with an extensive database of donors who contribute periodically and their characteristics, including past donation behavior. Another limitation is the low response rate of roughly 20%, leading to a smaller sample size than similar natural field experiments. This makes treatment comparisons underpowered, although this is not the main focus of this study. Lastly, it is likely that there are negative or even positive spillover effects because corporate employees usually share offices and might have discussed their donation conditions, i.e., having versus not having an allocation choice. This might explain why the treatment effects are very large. Therefore, it would be interesting to replicate these results more traditionally, i.e., with donors in a country with a higher giving index and in partnership with a fundraiser with a sufficiently large donor database, without the need to recruit donors through HR departments.

2 Let Me Choose What I'm Best at: A Natural Field Experiment with Volunteers

2.1 Introduction

What motivates people to exert effort? While money remains the primary work motivation for paid workers (see Levitt and Neckermann, 2014 for a review), several studies have found that monetary incentives can crowd out intrinsic motivation (Gneezy and Rustichini, 2000; Ryan and Deci, 2000; Bénabou and Tirole, 2003; Conrads et al., 2016). Economists have explored various interventions to motivate effort when financial resources are limited or absent. These interventions include strategies focused on recognition (Kosfeld and Neckermann, 2011; Ashraf et al., 2014; Conrads et al., 2016), adding meaning to mundane tasks (Chandler and Kapelner, 2013), and increasing workers' autonomy (Mertins and Walter, 2021)(Bloom et al., 2015), among others. The role of autonomy, a cornerstone of intrinsic motivation, has received relatively less attention in the literature. Autonomy can take numerous forms: prior studies have elicited autonomy by, for example, allowing employees to set their own work goals (Georg and Kube, 2012), to work from home (Bloom et al., 2015), to choose where the money collected by volunteer efforts will be spent (Mertins and Walter, 2021), or to choose among *different tasks* (Krügel and Meemann, 2024; Freyer et al., 2025). Different tasks differ in terms of skills, difficulty, intrinsic motivation, preferences, and ability, among other factors. While allowing workers to choose among entirely different tasks is a broad manifestation of autonomy, it makes it almost impossible to understand the mechanisms underlying the effect of autonomy on effort. In this study, I elicit autonomy by allowing people to independently choose one of three versions of the *same task*. This approach recognizes the diversity of individual preferences, interests, and abilities, fosters ownership and control, holds skills and difficulty constant, and broadly preserves variation in intrinsic motivation. By focusing on task choice within the same task, this design maintains a broad notion of autonomy that extends beyond specific contexts and involves decision-making at a fundamental level, allowing me to test the mechanisms through which autonomy affects effort.

Autonomy can be particularly pertinent in volunteering (Bidee et al., 2013). While we understand motivations to donate one's time [3], little is known about how to motivate volunteers to increase effort and effort quality. This is particularly important because there is evidence showing that people display stronger preferences for donating time rather than money (Brown et al., 2019; Wu and Eftekhar, 2024), even when it is costlier to do so.

This study focuses on motivating volunteers intrinsically through autonomy. I partnered with a UNICEF-funded NGO in Albania and implemented a controlled, large-scale

[3]For example, warm glow (Andreoni, 1990; Brown et al., 2019), image (Ariely et al., 2009) and career (Al-Ubaydli and Lee, 2011) concerns, and the six motives of Clary et al. (1998): 1. Values (humanitarian and philanthropic), 2. Understanding (knowledge, skills, and abilities), 3. Enhancement (e.g., gaining self-esteem and experiencing psychological growth), 4. Social (strengthen social connections and making new friends), 5. Career (use volunteering for career experiences), and 6. Protective (self-defense mechanism against negative feelings).

natural field experiment involving approximately 4,400 high school students to test whether allowing volunteers to choose tasks incentivizes them to exert more effort and perform better compared to volunteers who are randomly assigned tasks. The task consisted of writing awareness-raising messages against bullying, depression, and social exclusion of people with disabilities. Treated volunteers could choose which topic/task they wanted to work on. The first advantage of this approach is that it keeps the skills needed to perform the task constant, making effort and its quality across the three task versions properly comparable. Second, because the task is creative, cognitive, and meaningful, it should appeal more to volunteers than repetitive, simple, and potentially dull tasks requiring fewer cognitive skills (Charness et al., 2018). Third, this setup can easily extend to other settings. For example, report writers in consulting firms often self-select the specific domain/field they will write about/report on. This is also relevant in the context of sports and arts, where, e.g., football players may choose their field position while still playing football, or actors choose roles in a play while still acting in the theater. Fourth, despite the creative nature of the task, I measure effort objectively by using the total number of topic-relevant messages that volunteers write. Of course, the measures of effort quality in this paper are rather subjective, though we tried to define the properties of a "good message".

Further, I question whether empowering volunteers with the autonomy to choose tasks makes them more likely to participate in future volunteering events - a question that has not been answered before. This issue is particularly interesting because most of today's volunteers are episodic (Hustinx et al., 2008; Mertins and Walter, 2021). It is also interesting to examine whether the effect of autonomy is persistent. To answer this open question, I asked volunteers at the end of the experiment to write awareness-raising messages about a different cause (they were not given a choice). If the effect of autonomy persists, one would expect volunteers to exert more effort even when they cannot choose tasks.

Another important aspect of this paper is that it benchmarks the choice effect with the effect of incentivizing the best volunteers through ex-post and performance-contingent monetary rewards, which do not necessarily intend to crowd out effort. To benchmark the autonomy effect, I designed a financial reward treatment where participants were randomly matched with tasks; the volunteer with the best performance received a large reward of 20 Euros. This treatment combines the effects of both competition and monetary incentives. While it is not the aim of this paper to disentangle the two effects, in the context of volunteering, it is more natural to reward the best helpers, e.g., volunteer of the month awards, rather than treating them as episodic paid workers.

Lastly, this paper sheds light on the mechanisms underlying the effect of choosing tasks. Volunteers may exert more effort and write "better" messages because they choose the cause they care about or are more interested in, match their abilities with the task, value choice, or feel a heightened commitment to perform well when the choice is theirs, among other explanations. This study focuses on the ability-matching mechanism and discusses the other potential underlying channels. To test the effect of ability-matching, I asked all participants to rate their ability to raise awareness of each cause and then asked them to write awareness-raising messages about the cause corresponding to their highest-rated ability.

I find that allowing volunteers to choose tasks increases the quantity of output by 33-

37%, with a significant quality improvement. While those who could choose tasks reported a greater likelihood of engaging in future activities, there was no effect on signing up for future volunteering calls. This is a typical inconsistency problem of stated and revealed preferences. Next, I find that the effect of autonomy on effort is not persistent because, when volunteers are not given a choice, their effort declines to the baseline level.

I find no evidence suggesting that monetary incentives that target the best volunteers crowd out effort. While competitive monetary rewards work, I show that it is cost-effective to increase volunteers' autonomy rather than to pay them for two main reasons. First, the effort quality in the monetary reward group is not different from that in the control group, but it is significantly lower than in the choice treatment. Second, effort drops to the baseline level when I remove the monetary incentive in the other task. This means that, once one offers to pay volunteers for performing tasks well, they should continue doing so.

Finally, I provide causal evidence that ability matching is a mechanism at work, because, similarly to choosing tasks, matching tasks to abilities improves both effort and quality. Moreover, correlational evidence suggests that most volunteers (75%) in the choice condition selected tasks based on their perceived abilities rather than interests/preferences, which I proxy by the awareness and knowledge levels. This claim is further supported by the absence of a causal impact on enjoyment: if people choose tasks they like, they should enjoy performing them more. I also find no evidence suggesting that helpers exert more effort simply because they like choice or feel a heightened commitment to do well because the choice was theirs. These results are consistent with the competence component of the self-determination theory (Ryan and Deci, 2000), which states that people have a psychological need to feel capable and effective and to take actions that align with their needs. They are inconsistent with the "autonomy" component, which posits that people need to have power and control over their actions.

This paper contributes to the broad literature on effort motivation (Levitt and Neckermann, 2014; DellaVigna and Pope, 2018) by showing that more autonomy at the task level increases effort and its quality. While the positive impact of autonomy on effort has been shown before (e.g., Bloom et al., 2015; Mertins and Walter, 2021; Krügel and Meemann, 2024; Freyer et al., 2025), I elicit autonomy in a broad sense and in *a natural rather than an online experimental setting*, which can be extended to other contexts aside from the volunteering one. Because participants in my experiment chose tasks only after committing to volunteer, the setting mirrors situations in which managers seek to motivate workers who have already agreed to work for pay. Moreover, this study takes a further step in understanding the role of autonomy on effort. First, it considers the persistence of the autonomy effect, an aspect that I show is unlikely to be true. Second, it shows that people use autonomy to choose tasks they are best at.

More narrowly, this paper contributes to the limited experimental literature on ways to motivate volunteers. Existing studies suggest that, in the absence of financial incentives, volunteers expend more effort when they receive feedback on their performance and have a chance to vote on where money raised through their volunteering work should be spent (Mertins and Walter, 2021), but do not donate more time when the call for volunteers is

tailored to their primary motivation for volunteering (Al-Ubaydli and Lee, 2011).[4] This paper also informs nonprofits on the benefits of letting volunteers to choose tasks as opposed to rewarding top volunteers, taking a further step towards understanding the mechanisms through which volunteers become more prosocial when nonprofits empower them with autonomy.

The results also relate to the directed giving literature (Li et al., 2015; Eckel et al., 2017; Kessler et al., 2019), which shows that giving increases when donors can choose where their gifts go. However, the notion of directed giving has not been tested before in the context of time and effort donation. To the best of my knowledge, this is the first study showing that choosing how to donate time and effort makes people more prosocial. A common criticism of the directed giving literature is that charities lose all or part of their flexibility to manage donated funds. Similarly, nonprofits that allow volunteers to choose tasks may face situations where some tasks do not reach their completion rate. In this study, I show that managers can reassign the least chosen tasks to unpaid helpers after they have completed their preferred tasks, and this approach does not, at the very least, diminish their overall effort. This result can easily be extrapolated to paid worker settings.

The results are also consistent with the impact philanthropy model, which predicts that donors make choices and decisions with the aim to personally "make a difference" through their giving (Duncan, 2004). A few experimental studies have tested this model in the domain of charitable donations and found that donors respond positively to detailed information about the charity's intervention effectiveness (Cryder et al., 2013), prefer to give more to smaller charities (Borgloh et al., 2013), and avoid overhead costs (Gneezy et al., 2014) because they perceive a greater impact. While this prediction has been tested for financial donations, it has not been explored in the context of time and effort donations. The behavior of volunteers in this study mirrors that of impact philanthropists: by choosing tasks they feel most able to perform, volunteers perceive a greater ability to make an impact, which leads to increased effort.

Lastly, this paper adds to the discussion of crowding-out intrinsic motivation, which has been shown to occur for small, but not for large, monetary rewards (Gneezy and Rustichini, 2000; Conrads et al., 2016). Monetary incentives may backfire for voluntary helpers precisely because they choose to engage in these activities willingly. Whether *ex-post* and competitive monetary rewards that target only top-performing helpers (aka *the winner takes it all*) crowd out intrinsic motivation remains largely unexplored. Volunteers may perceive unconditional or *ex ante* monetary compensation as offensive (Conrads et al., 2016), yet they may increase their effort when rewards are conditional on exceptional performance — a distinction that this paper highlights. Ex post rewards can also boost intrinsic motivation and self-esteem, as individuals learn that they are talented and are rewarded for their exceptional contribution (Bénabou and Tirole, 2003). This means they have no reason to believe that their behavior is controlled by the nonprofit and consequently do a better job. In line with this theory, this study tests whether volunteers motivated by competitive, performance-contingent rewards indeed exert more effort without experiencing motivational

[4] In the study of Al-Ubaydli and Lee (2011), only volunteers who use volunteering for career concerns supplied more hours.

crowding out.

2.2 Conceptual Framework & Mechanisms

Human resources often hire volunteers to help with specific tasks to maximize output. Suppose a non-profit hires episodic volunteers to raise awareness of three different social causes using three corresponding writing tasks. The production function of writing awareness-raising messages depends on the following inputs: knowledge and awareness levels, perceived ability, and writing skills. Knowledge and awareness are necessary but not sufficient conditions of caring about or being interested[5] in a social cause. Hence, if a volunteer feels more aware or knowledgeable about the cause in task x than task y, she is more likely to choose task x and to exert more effort. Non-profits can easily assess a volunteer's knowledge/awareness of a specific cause via online surveys or quizzes prior to hiring helpers. On the other hand, the ability to raise awareness is unobserved to the non-profit, and primarily involves personal identification and experience that are assumed to vary across different tasks. If a volunteer feels more able to raise awareness in task x than cause y, she is more likely to choose that task and to do a better job. Given that abilities are unobservable and that it is costly for human resources to learn about a worker's ability, allowing volunteers to match their own abilities with the tasks may be an optimal solution. Finally, writing skills are also observable and involve a set of tools that make a person's writing output clear, grammatically correct, consistent, and organized. Because the tasks are identical up to the social cause, the set of writing skills required to perform each task should be task-invariant.

Volunteers choose to exert effort in tasks that align with their preferences/interests or their abilities in order to maximize their overall utility, which is generally comprised of an individual and a social utility component. If they care only about their individual utility, they may choose tasks based on preferences. However, if they care about the society in general, they may choose to allocate their effort in tasks they feel they the best at. While their effort allocation choice may not be trivial, it is reasonable to assume that volunteers who donate time and effort to help others have no incentive to waste their potential (ability) to do good.

The ability-matching mechanism also aligns with the competence component of the self-determination theory (Ryan and Deci, 2000), which posits that individuals have a psychological need to feel capable and effective. On the other hand, a preference for the options associated with the choice mechanism is related to the autonomy component: when individuals can choose tasks that align with their interests and values, they are more likely to be intrinsically motivated and experience greater enjoyment (Ryan and Deci, 2000). By contrast, while ability-matching is related to feeling capable and effective, having preferences for the options associated with the choice is related to experiencing greater enjoyment. The ability-matching mechanism is also connected to the impact philanthropist model, in which an impact philanthropist wants to personally "make a difference" (Duncan, 2004). In mirroring an impact philanthropist by choosing tasks they feel most able to perform, volunteers perceive a greater ability to make an impact, which may lead to increased effort.

[5]The term care in this context is equivalent to having preferences for social causes.

It is possible that volunteers do not choose tasks that align with their preferences and abilities but still reciprocate positively to having a choice, and therefore behave more prosocially. This mechanism is also consistent with the autonomy component of the self-determination theory, which postulates that individuals have a psychological need for control and choice. In behavioral economics, this mechanism is usually referred to as a preference for having a choice.

Alternatively, volunteers may not care about preferences for choice or ability-matching but may feel a greater responsibility to do well because they were given a choice. This result would be consistent with the job characteristics model (Hackman and Oldham, 1976). The model postulates that regardless of the task, when people choose, they may feel a heightened commitment to do well because the choice was theirs. This sense of ownership and accountability can lead to increased effort and focus on performing at their best. Despite focusing on the ability-matching mechanism, this paper addresses other potential channels via causal and descriptive tests.

2.3 Experimental Design

2.3.1 Field setting

I partnered with the Observatory for the Rights of Children and Youth in Albania - a UNICEF-founded NGO whose goal is to serve as an observer at the community level, ensuring practical application of children's rights by actively assisting and endorsing civil society efforts[6]. The Observatory manages UNICEF's U-Report platform[7], which was promoted during the experiment. Our partnership consisted of recruiting high school students to help the Observatory with the U-Report initiative and to raise awareness of selected causes that UNICEF, society, and youth, in general, are likely to care about. To ensure that, I went through several meetings with school principals and U-Report teachers to choose three social causes: bullying, depression, and social inclusion of people with disabilities.

We followed two steps to set up the experiment. First, the Observatory gained permission from the Ministry of Education and Sports in 2021 to promote the platform among high school students. Second, I was granted permission from the ministry to conduct a study in four high schools[8] in Tirana (around 5,000 students), and received special approval from the Information and Data Protection Commissioner, allowing me to implement the study without students' or parents' consent for participation in the study, conditional on my protecting each student's anonymity [9].

[6]Find more about the NGO by clicking here.

[7]A global platform working to change the lives of young people around the world through social messaging. It is designed by UNICEF to empower young people to voice their opinions on various issues and concerns.

[8]All four high schools are part of the U-Report initiative, and each has a U-Report teacher who is paid from the Observatory

[9]The project proposal was extensively reviewed by psychologists at all institutions, including psychologists at all four high schools.

2.3.2 Baseline data

This study combines administrative data from schools with baseline information from a supplementary survey. In February 2023, we distributed around 4,700 surveys in all schools on the same day and time, during a non-mandatory "education hour", which students are encouraged to attend because mentors discuss social problems that impact the school community.[10] We obtained permission to use this class to survey students and to conduct the main experiment.

We designed the survey (Figures 43–44 in Appendix B.5) to learn more information on how much students know about the selected social causes, how able and aware they feel to raise awareness of them, and how easy/difficult they think it is to raise awareness. In this setting, the ability to raise awareness of a social cause is an umbrella term that includes writing and communication skills, knowledge, awareness, personal experience, and identification (attachment). Because the tasks are identical up to the cause, writing and communication skills are task-invariant. On the other hand, knowledge, perceived awareness, personal experience, and identification are task-specific inputs. Because I measure knowledge and perceived awareness through the supplementary survey, the term ability primarily refers to one's personal experience and identification with a social cause. Personal experience and identification (or emotional attachment) should vary across individuals and tasks. Table 38 in Appendix B.5 shows that the ability to raise awareness varies statistically across all three social causes.

We measure perceived awareness and ability through a scale of 0 to 10, and perceived ease/difficulty in raising awareness through a dummy variable. To objectively measure students' knowledge of these causes, they completed a quiz about each cause that consisted of five multiple-choice questions. These questions were similar for each cause, and students were asked to mark only the correct alternatives. The average number of multiple options was 6, and roughly 50% were correct answers. In addition, we asked students whether they had volunteered for a local or international non-profit. To ensure anonymity, each survey was uniquely coded [11].

Before distributing the surveys, the teachers read the instructions out loud to students. They emphasized that filling in the survey was entirely voluntary and that they were free not to do so. Those who did not wish to participate did not receive a survey and were allowed to study for the next class or enjoy free time. While students were not paid to complete the survey, they were promised a shorter class if they paid careful attention to the survey. Meanwhile, mentor teachers were instructed to prevent any communication among students. To ensure that students had enough time to answer all survey questions, we asked ten high school students from non-sampled private schools to complete the survey and report their time, finding that on average, it should take less than 30 minutes.

Then, I received administrative data on students' gender (sex), age, cohort, and their

[10] We chose the last week of February strategically because it was several weeks apart from term exams.

[11] The U-report representatives trained teachers in advance on distributing the printed surveys. In addition, teachers received written instructions (printed and digital) and were reminded through stickers posted on the first-to-be-distributed survey. These steps ensure that each present student in the classroom receives their own coded survey. For strict anonymity reasons, only teachers could decipher the code.

most recent official grades in Albanian language, literature, math, and civic education. Math is generally a good indicator for screening students' academic performance. The grade in the Albanian language controls for students' writing skills, and the grade in literature proxies the ability of students to write in an impactful way using figurative language. Lastly, the grade in civic education screens students' overall knowledge of topics including community coexistence, law and human rights, health and nutrition, and environment, among others. Because the experimental task is creative, writing skills and knowledge should affect students' effort. This information complements the survey information tailored to the social causes. I linked the survey data to the administrative data through the unique codes that we assigned to each student.

2.3.3 Treatments, randomization, and outcomes

The experiment (*the volunteering event*) was implemented in April 2023 during the "education hour". This was a less stressful period for students because they had finished term exams, which made it easier to ask them to volunteer. There is a 5% attrition rate, which is not a concern because students were unaware that the volunteering event would occur on that day and did not observe the treatments before deciding not to show up in the class. Teachers did not read the instructions out loud because they were printed on the experimental sheets. However, they emphasized that this was a voluntary contribution to the Observatory and supervised the process to avoid communication among students [12]. To incentivize attention, students knew they had 45 minutes to complete the task and could take more time from the subsequent academic class. Overall, 4390 students filled in the supplementary survey and participated in the volunteering event. Students who only completed the survey or only volunteered to raise awareness are not part of the experimental sample.

Treatments

The main experimental task in this study is to write awareness-raising messages against bullying, depression, or social exclusion of people with disabilities. Students were randomly allocated into the four following conditions (see Figure 7):

Control: Students in classrooms assigned to the control group were told: The Observatory for the rights of Children and Youth is raising awareness of depression, disabilities and bullying. You can voluntarily help us raise awareness against (**a randomly chosen cause out of the three**) through writing awareness-raising messages. We will filter and post the best ones on the U-report platform, developed by UNICEF.

Choice: Students in classrooms assigned to the choice group were told: The Observatory for the rights of Children and Youth is raising awareness of depression, disabilities

[12]To rule out any influence of teachers on the decision to volunteer, the representative person at the Ministry of Education and Sports formally instructed the principals of the schools in advance to let mentor teachers know that this is an entirely voluntary event.

Further, teachers received the instructions for the volunteering event a day before the experiment and the envelopes containing the experimental sheets minutes before the 'educational hour' started.

and bullying. You can voluntarily help us raise awareness against the cause of your choice (☐ **social exclusion of people with disabilities** ☐ **depression** ☐ **bullying**) through writing awareness-raising messages. We will filter and post the best ones on the U-report platform, developed by UNICEF.

Ability-matching: Students in classrooms assigned to the ability-matching group were told: The Observatory for the rights of Children and Youth is raising awareness of depression, disabilities, and bullying. You can voluntarily help us raise awareness against (**the cause they felt more able to raise awareness in the survey stage**) through writing awareness-raising messages. We will filter and post the best ones on the U-report platform, developed by UNICEF. Note that they were not told that they were matched with tasks based on their perceived ability.

Monetary reward: Students in classrooms assigned to the monetary reward group were told: The Observatory for the rights of Children and Youth is raising awareness of depression, disabilities, and bullying. You can voluntarily help us raise awareness against (**a randomly chosen cause out of the three**) through writing awareness-raising messages. We will filter and post the best ones on the U-report platform, developed by UNICEF. The student in your classroom with most selected messages will receive 20 Euros.

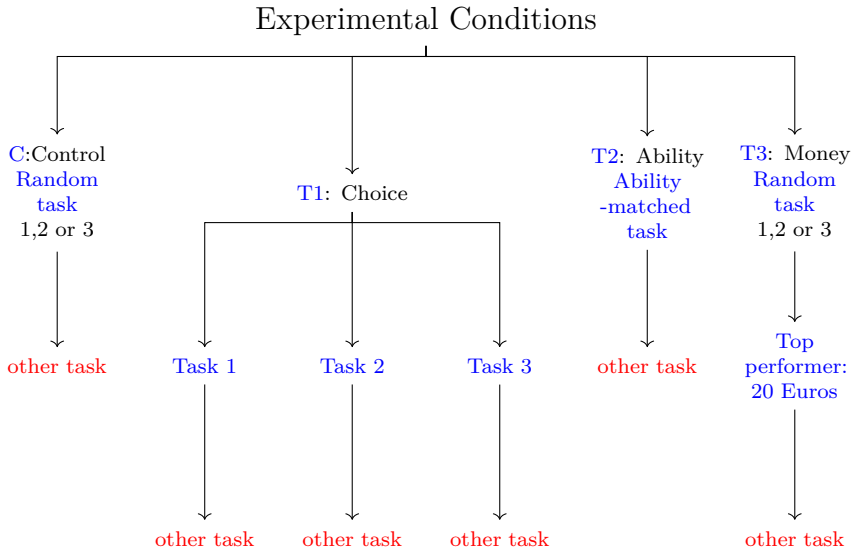


Figure 7: Experimental design

Note: The other task is writing awareness-raising messages about environmental pollution, and volunteers were asked to perform it after completing the main task. See the detailed experimental sheets in Appendix C.

The monetary reward (aka *the winner takes it all*) is a common practice of many NGOs worldwide in the form of "volunteer of the month/quarter/year". This treatment was designed to incentivize both the quantity and quality of effort. The winner was the

student who wrote the most messages for publication (*good messages*). For example, the ideal or most obvious winner would write the maximum possible number of messages, and all would be deemed publishable by the Observatory. The full description of the treatments is included in Appendix B.3 (Figures 38–41). I refer to the ability-matching and monetary reward treatments as simply "ability" and "money" for convenience.

In order to make the task more appealing, all students were told that, according to sociologists and psychologists, their input makes a real impact in society, and they were briefly informed about the U-report platform and its importance. They were asked to scan the QR code at the back of the sheet to register on the platform. Lastly, they were asked to write each message in capital letters to enhance readability and were also told to avoid short, mechanical messages, e.g., "*Stop cause X!*".

Upon completion of the first task, students were asked to rate how much they enjoyed the process, state their enjoyment level of the process, whether they would be willing to volunteer in future, and to sign up via email for future volunteer calls. Finally, they were asked to spend a few minutes writing awareness-raising messages about environmental pollution. At the end, teachers collected the experimental sheets and delivered them to U-report representatives, who monitored and documented the process through pictures and notes. Students signed a consent that the Observatory could use the pictures in reports, social media, and other outputs.

Randomization

Randomization occurred at the classroom level to rule out any spill-over effects arising from students sitting next to each other and learning about each other's treatments. I randomized 148 classrooms in four schools to four treatment arms[13] of unequal cluster sizes using the *randtreat command*. The monetary reward treatment has only 25 classrooms because the number of students whom I could pay was restricted by funding amounts. Further, I stratified randomization at the school level because schools may have different morale, prosocial principles, or volunteering spirit.

Tables 34–37 in Appendix B.4 present a series of regressions of baseline characteristics on treatment indicators for four comparison groups: (i) control & choice, (ii) control & ability, (iii) control & money, and (iv) choice & money. This set of tables shows that randomization was adequate. There are cases in which some variables appear individually imbalanced at the 5% level (e.g., Table 37), but in all five comparisons, the p-value of joint orthogonality is greater than 0.10.

Instead of concentrating solely on whether differences are statistically significant, it is more informative to focus on the magnitude of the differences. I estimate the normalized differences following Imbens and Rubin (2015). This approach calculates the difference in means between treatment and control groups and then divides it by the square root of half the sum of the variances of both groups, therefore presenting a size-based measure that is not influenced by the scale of the data. Imbens and Rubin (2015) argue that differences of 0.25 or less generally indicate a favorable balance between the groups. As shown in Tables

[13]See the randomization diagram in Figure 42 in Appendix B.4

34–37, the estimated normalized difference exceeds 0.25 in only one case (Table 37, variable *age*).

Outcomes

The main outcome of this study is the total number of relevant messages, i.e., the number of messages written for the cause assigned to the students or chosen by those in the choice treatment. This outcome variable measures effort. Off-topic messages are not counted in the analysis [14]. Students could also write awareness-raising messages about more than one cause. In such cases, I counted only the messages relevant to the cause they were assigned or chose. Lastly, I do not observe the chosen task for students in the choice group who did not write any messages [15].

I apply two measures of effort quality. The first is the share of good messages - the ratio of messages that have the potential to be posted on the platform (aka *"good" messages*) to the total number of relevant messages. A message has the potential to be posted on the platform if it is grammatically well written, has a clear meaning, and may impact the reader's thinking, touch the reader's heart, or both. Generic messages and uncited definitions are not considered "good" messages. The share of good messages is a standardized measure, allowing me to properly compare the quality of effort across different individuals. However, the number of good messages does not adjust for differences in total effort. A student who writes many messages (including low-quality ones) could appear to perform better in terms of raw output, but this does not reflect how well their overall effort translated into quality. For example, one student might write 20 messages, 10 of which are good, while another might write 5, all of which are good—without adjusting for total effort, the effort quality of these two students would be hard to compare.

The second quality measure assigns a grade from 1 to 10 based on the overall work, encompassing all messages, regardless of whether they are labeled "good". The first measure is fairly more objective because it does not rate the impact of messages. The second measure is subjective because it is similar to grading essays. To structure this measure, a student's performance is considered poor if their grade falls between 1 and 4, average if it falls between 5 and 7, and very good if it falls between 8 and 10.

To rate the quality of the volunteers' output, I hired a long-term volunteer at the Observatory. She received instructions on evaluating the messages and was unaware of the research hypotheses. Each experimental sheet had the treatment solicitation paragraph covered with a sticker, and the evaluator was asked to write the scores on the sticker.

Secondary outcomes include (i) willingness to volunteer in the future, measured by a dummy variable which takes the value 1 if the student wishes to volunteer in the future and 0 otherwise, and whether they sign up for future volunteer calls, meaning that they are willing to be contacted via email by the Observatory in the future, (ii) enjoyment level, and (iii) number of relevant awareness-raising messages produced against environmental pollution.

[14] Off-topic messages include those written for the other causes whenever students switched tasks. 11.09 % in the control group, 2.68 % in the ability group, and 8.31 % in the money group did so.

[15] This occurred in only 9 cases.

2.4 Main Treatment Effects

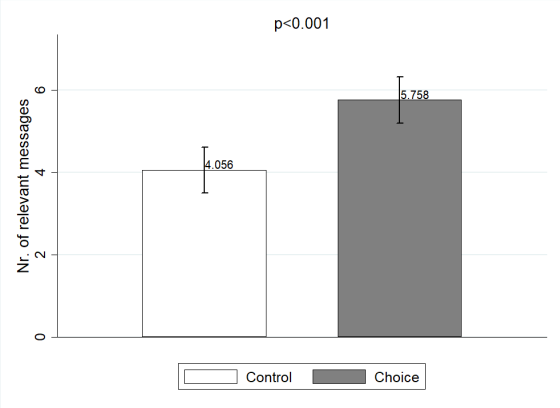
First, I describe the effect of choice on volunteers' effort, effort quality, and willingness to volunteer in the future. Then, I discuss the effect of the ex-post performance-contingent monetary reward on effort and its quality. Lastly, I benchmark the effect of choice with the effect of rewarding volunteers monetarily. Throughout the empirical analysis, I present a set of necessary robustness checks.

2.4.1 Does autonomy enhance effort and effort quality?

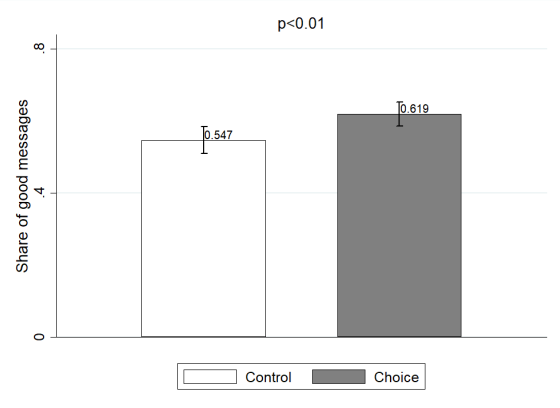
This section presents the effects of choice on the quantity and quality of the volunteers' effort. First, Figure 37 (a-c) compares the mean effort and its quality among the control and treatment groups. There is a strong effect of choosing the task: these volunteers wrote 42% more messages on average than volunteers in the control group (control mean = 4.056). This is the raw choice treatment effect, unconditional on strata and task fixed effects, and on baseline characteristics.

In Table 18 in Appendix B.1, I estimate the ATE of choice on effort, conditioning on strata, and task fixed effects solely (column 1) and including baseline characteristics (column 2). The effect remains significant in both specifications (p <0.001), though it drops to 37% and 33%, respectively. This drop occurs because strata, task-fixed effects, and controls—while balanced across treatment groups—are still important predictors of the outcome. By accounting for these controls, the estimation isolates the treatment effect more accurately, resulting in a smaller but more reliable estimate. For almost all estimates in this paper, standard errors became smaller with the inclusion of controls, thus increasing precision.

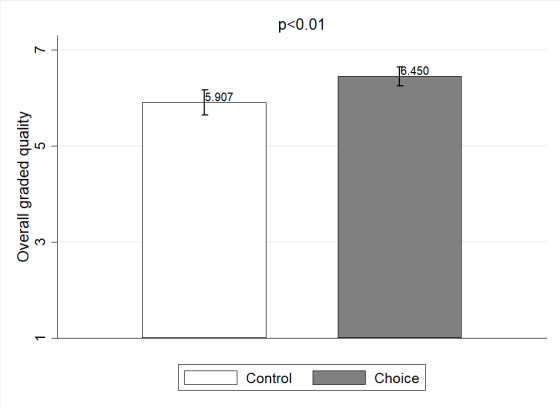
The choice treatment also has a significant effect on effort quality, measured by the share of "good" messages and the overall graded quality. These two measures are only observable for those who wrote at least one relevant message; therefore, the effect of choosing tasks on both measures is not entirely causal.



(a) Nr. of relevant messages



(b) Share of good messages



(c) Overall graded quality

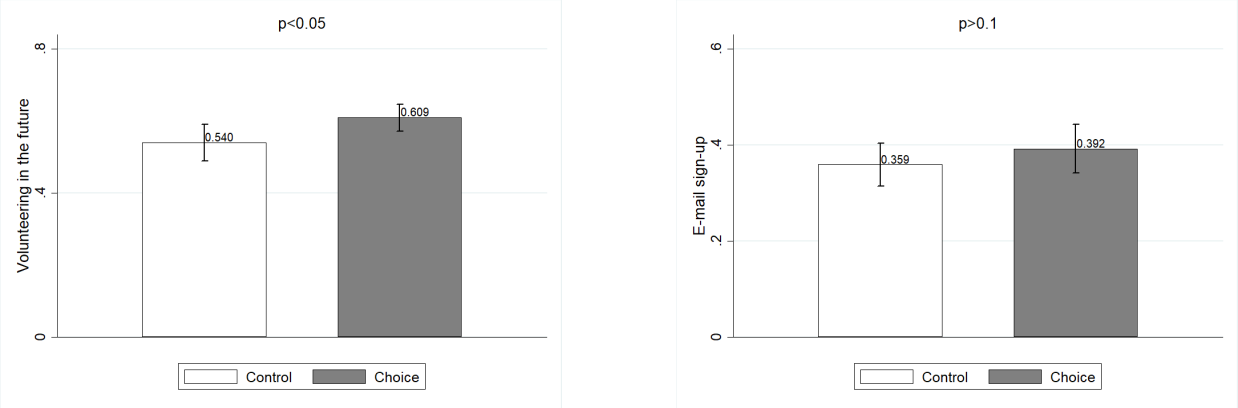
Figure 8: Choice vs. Control: differences in effort and quality

Notes: The confidence intervals are estimated using clustered standard errors. The p-value displayed in each figure is estimated from a regression of the outcome on the choice dummy, where standard errors are clustered at the classroom level.

The share of good messages and the overall graded quality are 13% (p <0.01) and 9% (p <0.01) higher for volunteers who could choose tasks than those who were matched with random ones (control means are 54.7% & 5.907 out of 10, respectively). These effects are robust to controlling for strata, task, and cohort fixed effects, and baseline characteristics (columns 4 & 6 in Table 18), although they drop to 10% and 7%. Again, including control variables leads to more precise and reliable estimates.

2.4.2 Likelihood to volunteer in the future

Next, I explore two secondary aspects of allowing volunteers to choose a task: (i) the likelihood that volunteers will engage in future activities with the non-profit and (ii) their willingness to sign up for these future events. These are important aspects in current times when volunteers are becoming episodic, i.e., switching from long to short-term commitments. Figure 9 (a-b) displays the raw means of these outcomes for the control and treatment arms.



(a) Willingness to volunteer in the future

(b) Email sing-up

Figure 9: Choice vs. Control: differences in the willingness to volunteer in the future

Notes: The confidence intervals are estimated using clustered standard errors. The p-value displayed in each figure is estimated from a regression of the outcome on the choice dummy, where standard errors are clustered at the classroom level.

Those who chose tasks are 6.9 % points more likely to volunteer in the future than those in the control group (control mean = 54%, (p <0.05)). However, they are equally likely to sign up via email for future events (p=0.331). Both estimates are robust to conditioning on strata fixed effects and baseline characteristics (Table 19 in Appendix B.1). These results align with the common inconsistency between stated and revealed preferences. While it might have been easy for them to circle "*yes*" on the future volunteering option, it is costlier to sign up via email. This inconsistency may not be due to the cost of an action or access to an email address, which students have because they regularly receive assignments from teachers. For example, students could scan the U-report QR code at the end of the page without signing up via email, but this data is inaccessible due to data protection policies. Alternatively, this inconsistency between stated and revealed preferences aligns with the pattern that volunteers often transition from long-term or regular commitments to shorter-term or even one-off volunteer engagements (Hustinx et al., 2008).

2.4.3 What is a better effort motivator: autonomy or monetary rewards?

While money is the standard effort motivator, some studies have shown that financial incentives can crowd out intrinsic motivation (Deci, 1971; Gneezy and Rustichini, 2000; Conrads et al., 2016). However, whether monetary rewards undermine intrinsic motivation may depend on how these rewards are delivered, as discretionary or ex-post rewards can bolster self-esteem and intrinsic motivation by signaling appreciation rather than control (Bénabou and Tirole, 2003).

In this study, the best volunteer in the classroom received a large reward of 20 Euros, around four times the average hourly wage in the country. The best volunteer was the one who produced the highest number of "good" messages. This is a standard practice that NGOs and the corporate sector employ, e.g., the volunteer/employee of the month/year, though the

reward may not necessarily be financial (Neckermann et al., 2014; Frey and Gallus, 2017). Some NGOs use gift certificates rather than money.

The comparison of the mean effort and effort quality between the control and money groups is displayed in Figure 10 (a-d). The sizeable monetary reward increased the total number of relevant messages by 28% (p <0.05). Table 20 in Appendix B.1 presents the adjusted average treatment effects, of 26% and 23% [16]. This result suggests that nonprofits can effectively motivate voluntary helpers through competitive and performance-contingent monetary rewards, as there is no evidence of a crowding-out effect on effort. This result is in line with the theoretical prediction of Bénabou and Tirole (2003), suggesting that non-promised rewards, such as performance-contingent ones, should not crowd out intrinsic motivation because behavior is not controlled by money.

Moreover, 46.9% of the volunteers in the money treatment pre-agreed that they would like to be considered for the reward, and only 7.5% stated otherwise. The rest did not specify [17]. Because the reward is highly competitive, the chances of winning it should increase with the number of relevant written messages, and this is one potential reason why there is no evidence of a crowding-out effect.

Alternatively, given that the reward goes to the best performer in the classroom, this involves an expected learning effect. That is, top performers expect to learn that they are talented and better than their peers, thus boosting their intrinsic motivation to exert effort in volunteering. It is, however, not the aim of this study to dig into the mechanisms that drive the effect of a monetary reward on effort, but rather to benchmark its effect with that of allowing volunteers to choose tasks.

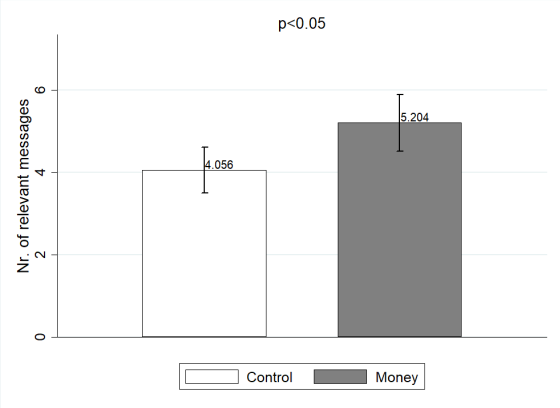
The monetary reward does not affect the quality of effort, although it was designed to affect both effort and quality. The share of good messages is almost identical for both groups (p=0.993), and the overall graded quality is insignificantly higher for the money group (p=0.384). Combining the effects on the number of relevant messages, share of good messages, and overall graded quality, it seems that volunteers in the monetary reward group pushed harder to win 20 Euros, but likely sacrificed the quality of their effort. Overall, it is unsurprising that competitive and monetary rewards do not enhance quality of effort in an entirely creative task. Charness and Grieco (2019) report similar findings for paid workers: tournament financial incentives, i.e., competitive rewards, do not impact creativity.

Figure 10 (d) suggests that money is a strong incentive for volunteers to exert more effort because, in its absence, the effect disappears. That is, monetarily incentivized volunteers exert statistically as much effort as volunteers in the control group in the other (*unpaid*) task (control mean = 3.334 & money mean = 3.106). These results are robust to controlling for fixed effects and baseline characteristics (Table 20 in Appendix B.1).

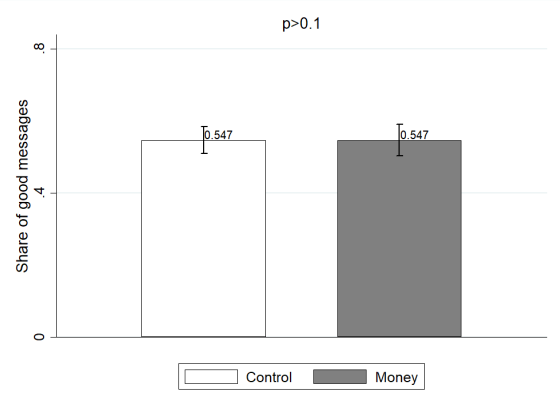
In a broader sense, these results suggest that, once volunteers are paid, they will

[16]Including fixed effects and controls increases the precision of the average treatment effects, leading to more reliable estimates.

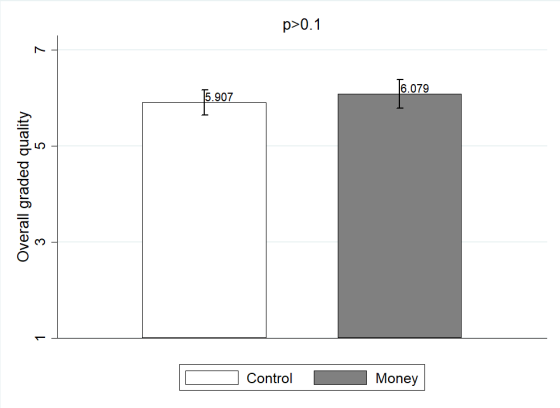
[17]These statistics are not out of the ordinary because students were told they could decide upon the announcement of the winners after consulting with their parents. Given that most high school students are minors, it is ethical to allow them to get permission from their parents before receiving money from third parties.



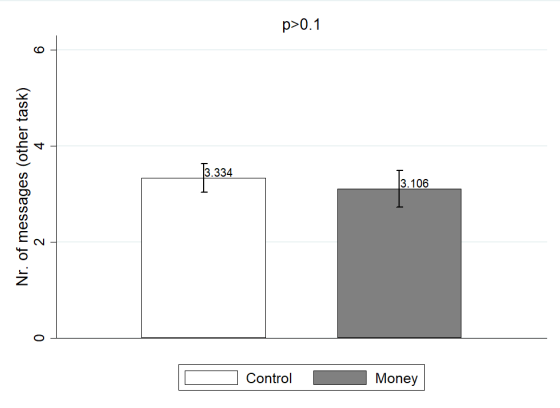
(a) Nr. of relevant messages



(b) Share of good messages



(c) Overall graded quality



(d) Nr. of relevant messages (other)

Figure 10: Money vs. Control: differences in effort and quality

Notes: The confidence intervals are estimated using clustered standard errors. The p-value displayed in each figure is estimated from a regression of the outcome on the choice dummy, where standard errors are clustered at the classroom level.

require payment each time tasks are assigned to them. Similarly, in his seminal paper, Deci (1971) concluded that participants who were paid to solve puzzles showed less interest in continuing the task once the payments stopped than participants who were not given external rewards. The conceptual difference between his findings and mine is that he shows that external rewards undermine intrinsic motivation when experienced as controlling one's behavior. In the present study, the behavior of volunteers in the monetary reward condition was not controlled by the NGO, because the winners received the reward ex-post. Hence, there is no evidence of crowding out intrinsic motivation.

Lastly, this section benchmarks the autonomy effect of choosing tasks with a large monetary reward. Figure 11 (a-d) first shows that there are no differences in mean effort and effort quality between the choice and money treatments. First, there are no differences between the choice and money treatments regarding the number of relevant messages (p=0.218). Second, the choice treatment outperforms the money treatment in terms of the quality of effort.

The share of good messages is 13% higher for volunteers who could choose tasks than those who were monetarily incentivized (p <0.05). Relative to the money treatment, choosing the task also increases the overall graded quality by 6% (p <0.05). Finally, the number of messages in the "other" task is 15.8% to 17.6% higher in the choice treatment. These results are robust to controlling for fixed effects and baseline characteristics (see Table 21 in Appendix B.1), except for the difference in overall graded quality, which becomes significant at the 10% level (p=0.052) after conditioning on baseline characteristics.

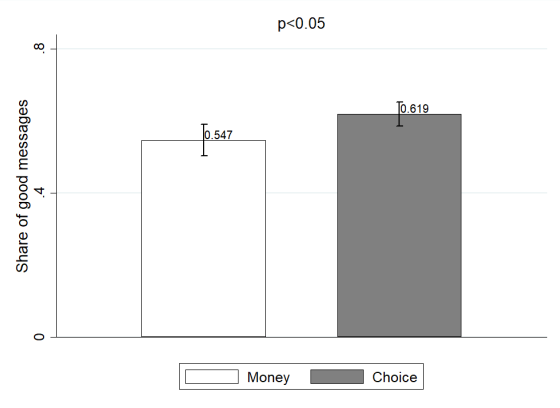
When benchmarking the choice effect with the monetary reward effect, one may claim that the proper comparison is relative to the control group, i.e., a difference in differences (DiD) approach. If one treatment impacts effort and quality more, this DiD estimate will capture that relative impact. By comparing each treatment against the control group and then comparing these differences, I reduce the risk of bias due to any baseline differences that the simple OLS regression might not capture. The DiD estimates are reported in Table 22 in Appendix B.1. The adjusted estimates remain similar in size and statistical significance for effort in both tasks and share of good messages. However, the difference in the overall graded quality becomes insignificant.

While money is a strong motivator, I show that it is cost-effective for nonprofits to empower their volunteers with autonomy because they gain in terms of the quality of the produced output at the cost of losing some of their flexibility to prioritize tasks or to reach specific completion rates for these tasks. If nonprofits allow volunteers to choose tasks, some tasks may not reach their completion target, which would not be an issue if volunteers were exogenously assigned tasks and paid for their good work. However, nonprofits can reassign the least chosen tasks to volunteers after they complete the task they choose, and consistently, with Figure 11 (d), they are likely to provide at least much effort as monetarily incentivized volunteers. Suppose that the other task, raising awareness of environmental pollution, was the fourth option, and no one chose it. Figure 11 (d) suggests that it is possible to ask volunteers who previously had the autonomy to choose tasks to perform the least chosen one because they will exert more effort than those motivated by monetary rewards. Human resources departments and management of firms may find these results beneficial as they effectively dictate efficient ways to use financial resources and human capital.

2.5 Mechanisms

In this section, I try to uncover the mechanisms underlying the effect of choosing tasks on effort and effort quality. Several channels may explain the autonomy effect. First, volunteers may exert more effort and enhance the quality of their effort because they may like having a choice, even if they do not specifically like the task they chose, which is consistent with a pure agency effect. Second, they can match their abilities with their chosen tasks and increase effort and quality. Third, they may sort into tasks they like most and do a better job. Fourth, conditional on their choice, they may feel a heightened commitment (or responsibility) to give their best effort because the choice was theirs. Specifically, I provide causal evidence that ability-matching is a mechanism at work and provide descriptive and causal evidence on the presence of the other potential channels.

(a) Nr. of relevant messages



(b) Share of good messages

(c) Overall graded quality

(d) Nr. of relevant messages (other)

Figure 11: Choice vs. Money: differences in effort and quality

Notes: The confidence intervals are estimated using clustered standard errors. The p-value displayed in each figure is estimated from a regression of the outcome on the choice dummy, where standard errors are clustered at the classroom level.

2.5.1 Does ability-task matching induce effort?

Here, I test whether exogenously matching tasks to perceived ability increases effort and quality. The motivation behind the ability-task matching mechanism is dual. First, it is costly for nonprofits, firms, and managers to learn their workers'/helpers' abilities, and thus, allowing them to work on tasks they feel they are best at may increase effort and quality. Second, while engaging in voluntary activities, volunteers should care more about maximizing social utility or their impact rather than their individual utility. Consequently, they are expected to select tasks they are best at rather than tasks they find interesting or care more about.

In the present setting, the ability to raise awareness of a social cause is an umbrella term that includes writing and communication skills, knowledge, awareness, personal experience, and identification (attachment). Because the tasks are identical up to the cause, writing and communication skills are task-invariant. I use the official grades from language

and literature to control for one's writing and communication skills. In contrast, knowledge, perceived awareness, personal experience, and identification are task-specific inputs. Because I measure knowledge and perceived awareness through the supplementary survey, the term ability primarily refers to one's personal experience and identification with a social cause. To show that perceived ability, awareness, and quizzed knowledge are task specific, I conduct several pair-wise comparisons in Table 38 in Appendix B.5. Indeed, there are significant differences in all of these inputs by task.

Figure 12 (a-c) compares the raw means of effort and effort quality between the control and ability-matched groups. Matching tasks to volunteers' abilities significantly increased the number of relevant messages by 28% (p <0.05), compared to a control group mean of 4.056. After adjusting for strata fixed effects, task fixed effects, and baseline characteristics, the estimated ATEs are 22.5% (p <0.05) and 21.5% (p <0.05) (Table 23 in Appendix B.1), making them more precise and reliable than the raw means shown in Figure 12.

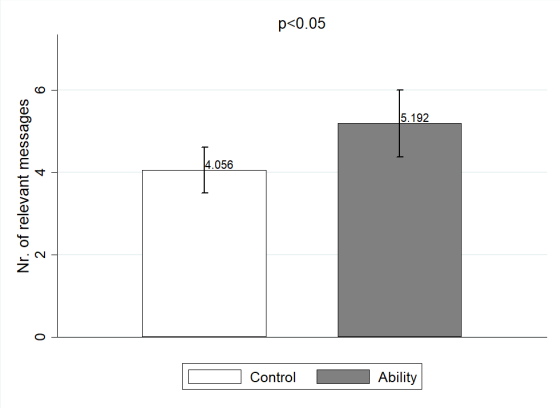
The effect on the quality of effort (Figure 12 b & c) is also positive and significant for both effort quality measures. The share of good messages and the overall graded quality are 11.7% (p <0.01) and 7.4% (p <0.01) higher in the ability treatment. These results are also robust to controlling for strata, task, cohort fixed effects, and baseline characteristics (Table 23 in Appendix B.1). The fully adjusted average treatment effects on the share of good messages and overall graded quality are 12.6% (p <0.01) and 7.7% (p <0.01).

These results provide causal evidence that matching tasks to abilities significantly enhances effort and effort quality, thus being one potential mechanism underlying the effect of choosing tasks. People naturally perform better when they engage in activities they feel they are good at. This aligns with the competence component of the self-determination theory, according to which people have a psychological need to feel competent; when they feel capable and effective in their activities, they are motivated to invest more effort and to enhance the quality of their work.

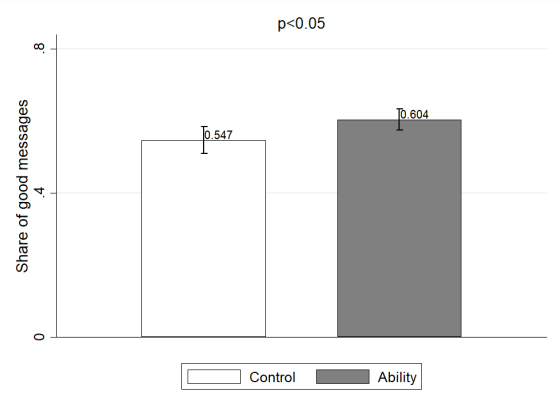
To further investigate the ability-matching mechanism, I utilized the random task assignments within the control group. The randomization meant that some volunteers were, by chance, matched with tasks that aligned with their highest perceived ability, while others were not. I created two indicators for ability matching. The first indicator equals one if a volunteer was matched with their highest perceived ability, excluding cases where all abilities were equally rated.

I exclude those cases because when a volunteer feels equally able to raise awareness of bullying, depression, or social exclusion, abilities should not matter. The second indicator equals one if volunteers were matched with their highest-rated ability, excluding all cases where they rated their abilities equally for two or more causes. I name the first indicator "ability-matched" and the second "exact ability-matched." The purpose of constructing these indicators is to compare the effort between volunteers in the ability treatment with those in the control group who were matched by chance with the task they felt most able to perform. Then, I replicate this analysis within the control group to compare differences in effort among ability-matched and unmatched volunteers. The results are shown in Table 24 in Appendix B.1.

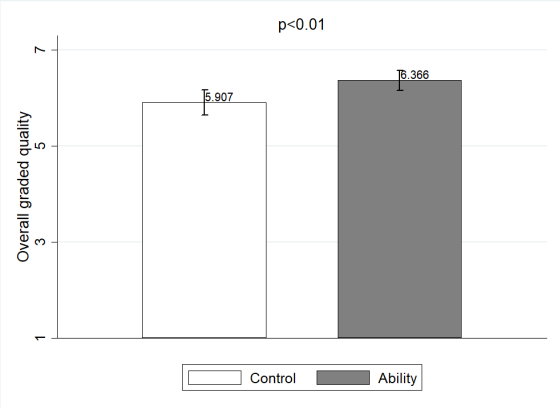
First, there is no significant difference in the number of relevant messages between



(a) Nr. of relevant messages



(b) Share of good messages



(c) Overall graded quality

Figure 12: Ability vs. Control: differences in effort and quality

Notes: The confidence intervals are estimated using clustered standard errors. The p-value displayed in each figure is estimated from a regression of the outcome on the choice dummy, where standard errors are clustered at the classroom level.

ability-matched volunteers in the ability and control groups (p=0.580; column 2, Panel A). Similarly, there is no significant difference in effort among exactly matched volunteers in both groups (p=0.868; column 3, Panel A). As ability-matching becomes more precise from the first indicator to the second, the difference in effort among these groups becomes smaller. Moreover, among volunteers in the control group, those who were matched with the highest ability tasks wrote, on average, 20.9% (p <0.01) more messages than those who were not (column 4, Panel B). Similarly, those who were exactly matched with their abilities wrote 22.7% (p <0.01) more messages than those assigned to tasks randomly or to tasks that somewhat matched their abilities (column 5, Panel B). As ability-matching becomes more precise, the difference in output between matched and unmatched volunteers increases. These estimates are similar in size to the causal effect of ability matching (column 1). In all regressions, I control for fixed effects and baseline characteristics.

2.5.2 Understanding the choice of tasks

In this section, I provide correlational evidence supporting the ability-matching mechanism by analyzing how volunteers in the choice treatment selected tasks. Table 25 in Appendix B.1 presents the results of three regressions of task choice dummies ($task_i = 1$, $i = bullying, disability, depression$) on ability, awareness, knowledge, and task ease measures. The results suggest that volunteers are significantly more likely to choose a task when their ability to perform it is higher. This pattern is consistent across all tasks, suggesting task selection is mainly driven by volunteer's perceptions of their ability (Panel A). Additionally, there is a negative and significant correlation between the ability to perform other tasks (e.g., to write about depression or social exclusion of people with disabilities) and the likelihood of choosing the focal task (e.g., bullying). This further reinforces the notion that individuals specialize in tasks where they feel most competent and refrain from spreading their efforts across areas where they feel they may not perform as well.

On the contrary, no significant and consistent pattern suggests that being more aware or knowledgeable of one cause made volunteers more likely to choose the task associated with that cause (Panel B & C). Regarding the perceived ease of raising awareness of a specific cause, column 1 (Panel D) results suggest that those who feel that raising awareness of bullying is easier are less likely to choose bullying. However, the coefficients of the ease dummies on the other two causes are also negative and significant, which does not help to build a story of task choice based on perceived ease.

Overall, around 75 % of volunteers in the choice treatment performed tasks corresponding to their highest-rated ability. This includes cases in which they believed they were equally able to perform all tasks. When asked to choose only one task, volunteers in the choice treatment can still differentiate in abilities and choose the one they feel slightly more able to perform despite rating their perceived abilities in all three tasks equally. The composition of task choice for volunteers in the choice treatment arm is as follows: 52.2% chose bullying, 29.3% chose depression, and 19.5% chose social exclusion of people with disabilities.

2.5.3 Alternative mechanisms: preference for choice & tasks, and heightened commitment

In this section, I discuss other potential mechanisms. One possibility is that having the opportunity to choose increases prosocial behavior, as suggested by Kamdar et al. (2015) and Eckel et al. (2017) because people appreciate having choice/agency. In this study, the idea that individuals may have a preference for choice, regardless of the specific options available, could explain why volunteers increase effort and quality. To test this mechanism, I compare the mean effort in a task presented after volunteers completed the main task. This other task was identical for both choice and control groups; no one had an option to choose (Figure 13 a & Table 26). The results show no significant difference in effort ($p = 0.189$).

This result suggests that when volunteers who could previously choose tasks no longer have that choice, they do not exert more or less effort than those who never had a choice. If the preference for choice mechanism were at work in this longer timeframe, one would expect a significant positive or negative effect for the following reasons. If volunteers appre-

ciate having a choice, they might reciprocate positively by exerting more effort even when that choice is removed. Conversely, they might react negatively by exerting less effort simply because they prefer to have choice. However, the "no difference in effort" result is inconsistent with either of these explanations. Therefore, there is no evidence to suggest that the preference for choice is driving the positive effect of choosing tasks in this study. Fatigue could in principle affect effort in this other task. However, reported enjoyment from the first task is high and comparable across treatment arms (see 13), suggesting that participants did not experience differential depletion. Given random assignment and identical task structure across conditions, differential fatigue is unlikely to drive the results. This result also suggests that the autonomy effect of choosing tasks is unlikely to persist.

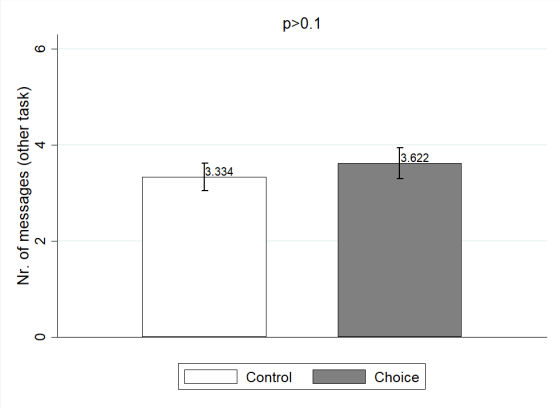
Another potential mechanism is preferences for one or more of the choices. Volunteers could choose tasks that align more closely with their preferences/interests. They were not asked to rank preferences directly[18], but were asked to state the degree to which they were aware of each cause and were quizzed about their knowledge. Knowledge and awareness aspects should be necessary, but not sufficient to care about, prefer, or be more interested in a particular social cause. The fact that there is no significant correlation of task choice with awareness or knowledge suggests that the preference for the options associated with the choice is not likely to be a mechanism at work. Moreover, if volunteers chose tasks they liked the most, they should have reported a greater enjoyment than those who were matched with tasks randomly. Figure 13 (b) suggests that this is not the case: there is no difference in the mean enjoyment level across volunteers in the control and choice groups[19]. Overall, the results are inconsistent with the autonomy component of the self-determination theory (Ryan and Deci, 2000), which posits that when people are empowered with the freedom of choice on how they engage in certain activities, they are likely to find these activities more enjoyable and personally satisfying because it aligns with their preferences.

In summary, this paper provides causal and correlational evidence that having autonomy to choose a task increases effort and quality mainly because people match tasks to their abilities. This implies that when people donate their time to help others, they do not want to waste their potential to contribute. This represents a rational decision for the following reason: if people willingly donate their time to help others, they do not want to waste their potential to do good when they can. Alternatively, if they would allocate their effort in tasks they like, they may end up in situations where their personal utility is maximized at the cost of wasting their potential to do good. This is fully in line with the impact philanthropy model (Duncan, 2004), which posits that people make choices and decisions to maximize their impact or to personally make a difference.

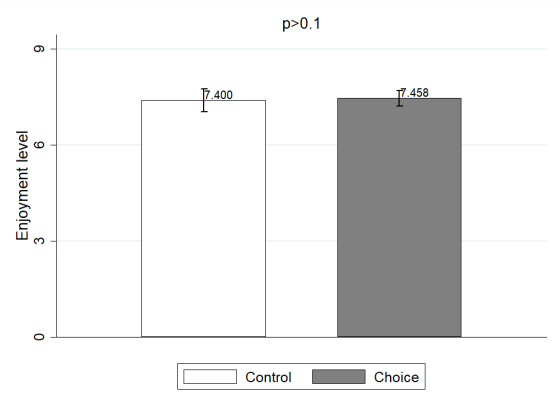
Ability-matching may not be the only mechanism underlying the effect of choosing tasks on effort, even though 75% of participants in the choice treatment selected tasks that they signalled best reflected their abilities. One potential alternative mechanism is a heightened sense of responsibility to do a better job after making a choice (Hackman and Oldham, 1976). When individuals can choose their tasks, they may experience greater

[18]Simply asking people how much do you like or care about cause X may be cheap talk. For instance, someone may answer they like X over Y while their knowledge or exposure to X & Y may be limited.

[19]Table 9 in Appendix A confirms that this result is robust to controlling for strata fixed effects, cohort fixed effects and baseline characteristics.



(a) Nr. of relevant messages (other)



(b) Enjoyment (0-10)

Figure 13: The effect of choosing tasks on effort in the other task and enjoyment level

Notes: The confidence intervals are estimated using clustered standard errors. The p-value displayed in each figure is estimated from a regression of the outcome on the choice dummy, where standard errors are clustered at the classroom level.

ownership and accountability, leading to increased effort. If this mechanism were the primary driver of the results, we would expect volunteers in the choice group to make random task selections, committing to increased effort simply because they had the freedom to choose. However, this is not what I observe in the study; most participants chose the bullying task, followed by depression and social exclusion, matching the overall ranking of abilities in the whole sample.

The heightened commitment mechanism may enhance the effect of ability-matching. Those who self-selected tasks based on their abilities may commit to exerting more effort than those who were exogenously given tasks based on their highest-rated abilities. To test this, I compare the number of relevant messages produced by participants in the choice group with those in the ability group. Table 27 in Appendix B.1 (column 1) suggests no significant difference in effort between the two groups.

To address the potential issue that not all participants in the choice group based their decisions on abilities, I conduct further comparisons, focusing only on those matched based on abilities (column 2): the difference in effort remains insignificant. Furthermore, when testing this hypothesis across different degrees of ability matching—such as exact matching (column 4) or excluding the cases where all three tasks were equally rated based on abilities (column 3)—there are no differences in effort. These results suggest that the heightened commitment or responsibility does not significantly impact effort beyond ability-matching alone.

2.5.4 Robustness checks

In prior sections, I have presented various tests showing that the treatments' effects are robust to conditioning on strata, task, and cohort fixed effects, as well as baseline characteristics. Here, I conduct additional robustness checks.

Choice effect by task: To test whether any specific task drives the effect of choosing tasks on effort, I run three OLS regressions of the number of relevant messages on the choice treatment conditioning on strata and cohort fixed effects as well as baseline characteristics. The estimated treatment effects by task are reported in Table 28 in Appendix B.2. The results do not suggest that one specific task drives the overall effect. Even though the effect sizes are slightly different, they are all in the same direction (positive) and statistically significant.

Spillover effects: It might have happened that those who performed the tasks on a following day learned about the treatments from their friends and altered their behavior accordingly. To test for any spillover effects on effort, I replicate the main regressions, excluding those who did not perform the tasks on the same day as their peers. The results are reported in Table 29 in Appendix B.2. The estimates are significant at the same level and almost identical to those from the regressions estimated using the full sample. Therefore, the main treatment effects are unlikely to be driven by any spillover effects.

Robust to outliers: Next, I test whether the treatment effects on effort are subject to outliers. I winsorized the number of relevant messages at the 1% and 5% levels and regressed it on each treatment indicator as shown in Table 30 in Appendix B.2. Although the average treatment effects drop slightly as winsorization increases from 1% to 5%, the estimates are significant and very similar in size to the unwinsorized ones reported in Tables 1, 3, 4, and 6. This suggests that the results are not driven by outliers.

Contamination bias: I test whether the main estimates are robust to contamination bias (Goldsmith-Pinkham et al., 2024). The authors argue that in linear regressions with multiple treatments and stratification, each treatment effect is composed of a convex weighted average of the own treatment effect and a non-convex weighted average of the effects of the other treatments - the contamination bias. For example, conditional on strata fixed effects, the ATEs of choice, ability, and money treatments are slightly different when estimated separately than when they are estimated jointly. Goldsmith-Pinkham et al. (2024) propose two bias correction estimators, which are employed in this study. The results are reported in Table 31. The bias-free estimates are very similar in magnitude to those reported in Tables 18, 20, and 23, which indicates that the contamination bias is minimal and unlikely to impact the causal interpretation of the treatments. This application of this method does not allow for clustering standard errors or including control variables, which is why the reported standard errors are different from those reported in Tables 18, 20, and 23. Also, each regression conditions only on strata fixed effects and on no other baseline characteristic or other sets of fixed effects.

Multiple hypotheses testing: In the next robustness check, I focus on multiple hypotheses testing. I computed the Romano and Wolf (2005) step-down adjusted p-values, which control the FWER and allow for dependence among p-values by bootstrap resampling (Table 32 in Appendix B.2). Comparing the model p-values (column 2) with the Romano-Wolf p-values (column 4), there is no change in significance in any of the 12 estimates corresponding to the main 12 hypotheses. Hence, the results are robust to multiple hypothesis testing.

Randomization based inference (RBI): Randomization inference considers what

would have occurred not only under the actual random assignment, but also under all possible random assignments. Simulating many different treatment allocations (1,000 replications), allows me to test whether the observed results would hold across these alternative assignments, providing a more robust check on the significance of the findings without relying on parametric assumptions. The results are shown in Table 33 in Appendix B.2. In all cases, the model and RBI p-values are very similar. This suggests that under 1,000 possible treatment allocations, the results would still be valid.

2.6 Conclusion

In this study, I ran a large-scale natural field experiment with over 4,000 high school students to answer whether and why volunteers increase effort and effort quality when they have the autonomy to choose their tasks. I also explore whether this autonomy effect will persist when choice is removed and benchmark it with the effect of a large monetary reward. I find strong effects of choosing the task on output produced and its quality but no persistence of the effect. Though the effect may not persist, the results suggests that it is cost-effective to empower volunteers with autonomy rather than with monetary rewards.

I provide causal evidence that matching abilities with the task increases quantity and quality, and volunteers choose tasks mainly based on their perceptions of their ability. This result is consistent with the self-determination theory, which posits that people have a psychological need to feel competent and effective, and the impact philanthropy model, in which people make choices they believe will increase the impact of their work. Other mechanisms including preference for choice, preferences for the options associated with the choice, or heightened commitment to do well when the choice is theirs are unlikely to be at work.

In addition to contributing to the economic literature on whether and how more choice enhances effort and prosocial behavior, the implications of these results may also be informative to nonprofits on ways to motivate their voluntary helpers to boost their efforts in the absence of money. An interesting extension of this paper would be to examine the role of autonomy in a setting of paid workers who frequently decide to work overtime and usually receive competitive or non-competitive bonuses. Such workers may prefer to choose their own tasks rather than being assigned to random tasks with a bonus. This may happen because they want to learn or develop a new skill or explore their interests more broadly, thereby bringing meaning to their work—a new frontier in motivating paid workers (Ashraf et al., 2025).

3 Unrealistic Educational Aspirations: Behavioral Insights From Parental Decisions in Albania

This study is part of a larger experimental project in Albania in collaboration with Enkelejda Havari (IESEG, Paris), Rebecca Dizon-Ross (UChicago), and Esmeralda Shehaj Zhabjaku (EBRD). In this paper, I present some of our results that motivate the next steps of our experimental project.

3.1 Introduction

Poor education choices have significant economic consequences. Choosing the wrong track can lead to dropping out (Goux et al., 2017), acquiring skills that do not match labor-market demand, and struggling to find suitable employment (Cervantes and Cooper, 2022), often resulting in NEET (not in education, employment, or training) status or non-standard employment (NSE), with lower wages and greater job insecurity (Cassidy and Gaulke, 2024). These situations can hinder career progression and negatively affect individuals' mental well-being (Llena-Nozal, 2009).

One possible cause of poor educational choices is high parental aspirations, which may lead them to steer their children towards unrealistic educational paths, leading to long-term challenges. In many countries, parents of children in the final grade of lower secondary (middle) school make decisions about their children's choice of secondary school, whether it is academic, vocational, or oriented towards a specific field of study. These decisions are frequently guided by parents' aspirations, e.g., most parents want their children to have a college degree (Leuven and Oosterbeek, 2011). In countries where access to secondary education is not constrained by GPA-based entry criteria (e.g., Spain, Albania, Greece, Italy, Finland, Denmark, and Sweden, among others) - but admission to university is - parents may hold unrealistic aspirations.

What determines track choice after middle school in education systems where enrollment in secondary education is free of ability constraints and unbinding teacher recommendations? Why do most parents prefer the academic track, while fewer opt for vocational schooling? Parents' knowledge and beliefs, as well as educational aspirations, play a crucial part in these decisions (La Ferrara, 2019). The standard economic models of investment in human capital assume that parents have perfect information about their children's academic ability when they make these decisions. However, experimental evidence shows that parents tend to overestimate their children's performance in middle school (Dizon-Ross, 2019). This might also be true for higher levels of education. For instance, many parents aspire to college education for their children and enroll them in the academic track after middle school, unaware that college may not be a feasible option, depending on their academic ability.

This paper provides new evidence that parental misperceptions about academic trajectories and college eligibility are a key driver of secondary track choice. Our analysis combines administrative data on GPA transitions with parents' subjective expectations, allowing us to directly compare perceived and actual academic trajectories. We study parental

aspirations and decision-making in Albania, a setting where parents freely choose between academic and vocational high schools, while university admission depends on performance in high school and the State Matura exams. Institutional details are described in the next section.

Using administrative data on over 9,000 students and a large-scale parent survey, we document four main findings. *First*, students experience a substantial and predictable decline in GPA when transitioning from middle to high school, yet most parents fail to anticipate this drop. Over half of parents expect their child's GPA to remain stable or increase, and fewer than five percent correctly predict their child's expected high school GPA based on historical outcomes. Parents are systematically more optimistic about their own child than about a "typical" child with the same prior performance. *Second*, these misperceptions translate into severe biases in beliefs about college eligibility. Parents substantially overestimate their child's probability of reaching the GPA thresholds required for college applications, with optimism increasing as the thresholds become higher. While parents tend to be pessimistic about the college eligibility of a typical child, they remain highly optimistic about their own child, even when objective probabilities are low. These patterns are consistent with overoptimism and wishful thinking rather than informed updating. *Third*, we show that alternative explanations for low vocational enrollment receive little empirical support. Parents' beliefs about earnings and employment chances across educational tracks are broadly accurate and do not portray vocational education as financially inferior. Nor do parents strongly endorse stigma-based explanations for vocational avoidance. While parents believe that others may view vocational education as low status, explicitly stigma-related concerns are rarely selected as reasons for low enrollment. Incorrect beliefs about college access from vocational education exist, but are limited to a minority of parents. *Finally*, we document a stark mismatch between academic performance and actual track choice. Even students with low middle-school GPA - well below levels consistent with college eligibility - are disproportionately enrolled in academic high schools, while vocational education remains underutilized across the entire GPA distribution.

Taken together, our results suggest that excessive demand for the academic track might be primarily driven by parental misperceptions and unrealistic aspirations regarding academic trajectories and college eligibility, rather than by stigma or misperceptions of labor market returns. These findings underscore the importance of providing parents with clear, data-driven information about academic progression, and realistic post-secondary opportunities, thereby improving educational sorting and reducing inefficient schooling choices.

3.2 Context and Data

3.2.1 Institutional context

The education system in Albania is structured into four main levels (see Figure 14). Pre-university education begins with pre-school (kindergarten) for children aged 3–-6 (non-compulsory). This is followed by primary education, which lasts 5 years (Grades 1–-5), and then lower secondary education for 4 years (Grades 6–-9); together, these 9 years are compulsory. Middle school ends at grade 9.

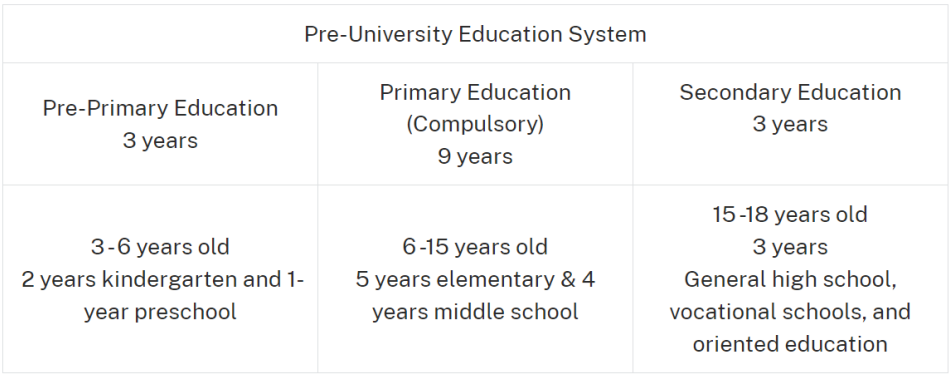


Figure 14: Education levels in Albania

Notes: Vocational programs vary from two to four years. Conditional on successful completion of the State Matura of the Professional (Vocational) State Matura (for those who choose the VET track), students are eligible to apply to universities (colleges).

After completing Grade 9, students enter upper secondary education, which lasts three years (Grades 10–-12) and is divided into general (academic) high schools, vocational schools (VET), and specialized schools (such as those oriented towards arts or sports). Upon graduation, students may continue to tertiary education, which includes universities and professional colleges, typically structured according to the Bologna system (Bachelor, Master, PhD). To do so, they should graduate from the academic high school with a GPA of at least 6.5 (arithmetic average of the overall high school average and the four State Matura exams). The grading system is from 4 (a fail) to 10 (the highest mark). The path to college is displayed in Figure 15. Some programs require a higher GPA, e.g., medicine. Those who complete vocational high school can also apply to college programs, provided they successfully complete the vocational State Matura. However, they can only apply to a very limited range of programs, and they have to compete with those who graduated from the academic track. In general, very few make it to college.

There is one clear trend in track choice after middle school: over 80 percent of students take the academic or oriented track, and very few choose the vocational track (Figure 16). There are several possible reasons why this trend has persisted over the past few years. Aside from the idea of overeducation and unrealistic aspirations, we explore other motives why parents are reluctant to enroll their kids in the VET track. These motives include a potential stigma or low reputation associated with VET schools, incorrect beliefs that a VET degree blocks access to college, or misperceived returns to education and employment opportunities in the vocational track. To explore these motives, we designed a survey for parents, and the next subsection summarizes the data we collected.

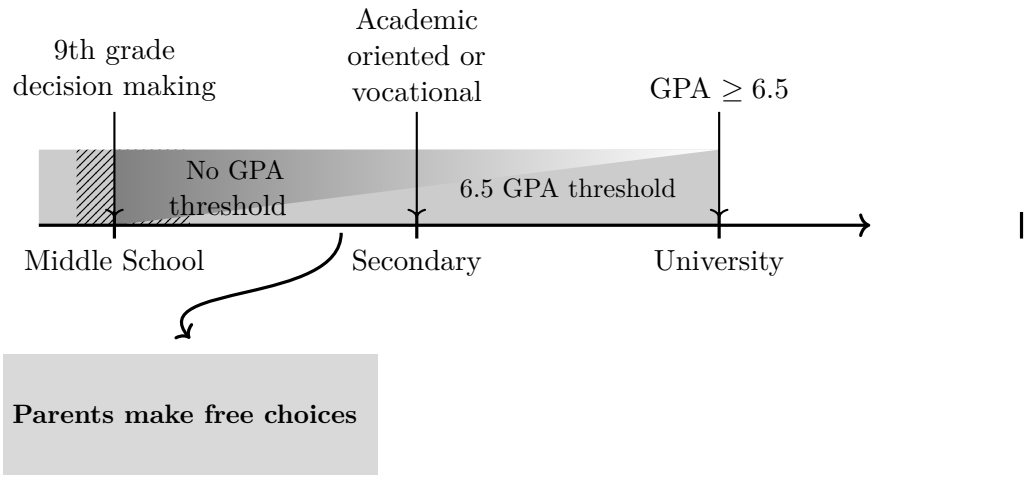


Figure 15: Path to college

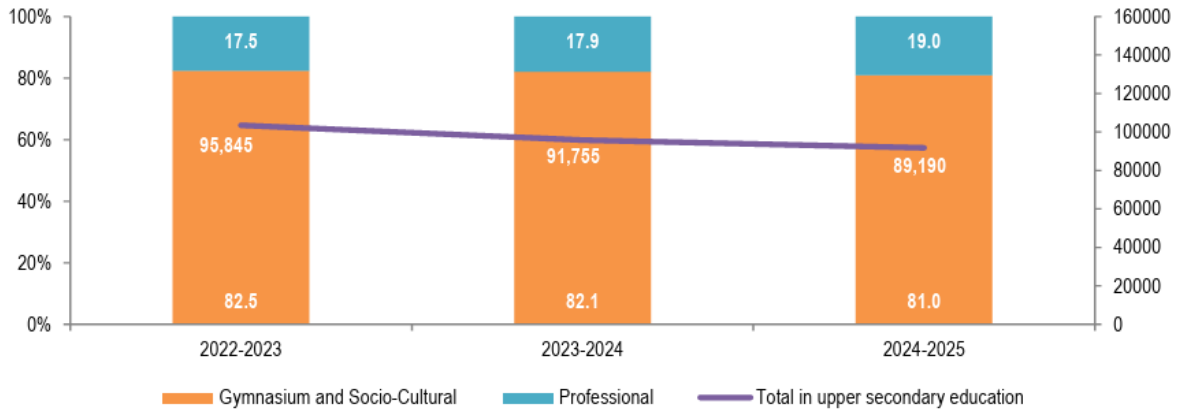


Figure 16: Enrollment in secondary education

Notes: Gymnasium or the Socio-Cultural Gymnasium refers to the academic and oriented tracks which are grouped together because students who complete either have to go through the State Matura if they want to apply to college programs. Professional refers to the vocational track. Students who complete the vocational track and aspire college, have to go through the Professional State Matura.

3.2.2 Data and collaboration with the Ministry of Education

We formed a long-term partnership with the Ministry of Education in Albania in 2024. Throughout this partnership, we received rich administrative data tracking over 9,000 high school graduates from 2021 (their last year of middle school) to 2024 (their last year of high school). The dataset contains the overall GPA at the end of high school (arithmetic average of the GPAs of each of the three high school years), the grades in State Matura exams, the GPA for grade 9, and grades in math, Albanian language, and literature, as well as English language (mandatory State exams at the end of middle school). Then, we calculated the

probability of being eligible for college, conditional on completing the academic track, for each GPA bin at the end of middle school. To check whether college eligibility differs by family background, we matched this data with socio-demographic information on parents, including educational attainment and occupation. In unreported outputs, we show that these characteristics do not impact eligibility chances.

In the summer of 2025, we sent a survey to around 11,000 parents through the official Ministry's channels, but only 4,000 parents fully completed it. The survey was part of a career orientation program recently implemented by the Ministry of Education in all middle schools. The survey was intended to collect information mainly on the following:

- Parents' aspirations and expectations about the future education of their children, including track choice (vocational or academic) and whether they aspire to college education (in Albania or abroad).
- Children's aspirations about their future education (same as above).
- Parents' beliefs about the Grade 9 GPA of their children.
- Parents' beliefs about the probability that a typical child with the same current performance as their child will have a GPA of at least 6.5 at the end of State Matura (eligibility criterion to apply to college).
- Same beliefs but for different GPA thresholds, including 7, 7.5, 8, and 8.5.
- Parents' perceptions of the returns to different types of education, including vocational high school, academic high school, and college.
- Parents' perceptions of future employment after graduating from the academic and VET tracks.
- Parents' perceptions on whether the vocational track is stigmatized.
- Parents' perceptions on why enrollment in the vocational track is low.

Then, we linked the survey data to actual enrollment data. The primary purpose of this survey is to determine whether parents are aware of their children's current performance and whether their aspirations for their children's future education are realistic, i.e., whether they accurately perceive the college eligibility chances of their children. This survey is also useful for shedding light on motives underlying why very few parents choose the vocational track.

3.3 Results

This section presents three sets of results. First, we examine parents' beliefs about their children's academic abilities and educational trajectories, documenting systematic misperceptions regarding GPA progression and college eligibility, as well as stronger optimism toward one's own child compared to a typical student. Second, we examine alternative channels that may explain low enrollment in vocational education, including perceived labor

market returns, stigma, and beliefs about access to college, and show that these factors are unlikely to be primary drivers of track choice. Third, we describe actual secondary-school track choices in relation to prior academic performance, highlighting a substantial mismatch between students' middle-school GPAs and the tracks they ultimately enroll in.

We analyzed a dataset of over 9,000 students who completed the State Matura exams in 2024 (academic year 2023-2024) and tracked their academic records back to 9th grade in middle school. We find that, on average, students experience a drop of 0.87 in their GPA when transitioning from middle school to high school, as measured at the end of the Matura exams (GPA defined according to the Decision of the Council of Ministers). A decline in GPA is observed in 90.74 percent of students. However, Figure 17 reveals substantial parental misperceptions. A majority of parents do not anticipate a decline in their child's GPA upon the transition from middle to high school: 24.47 percent expect no change, and 32.84 percent expect an increase. Only 42.70 percent correctly anticipate a decrease in GPA during this transition.

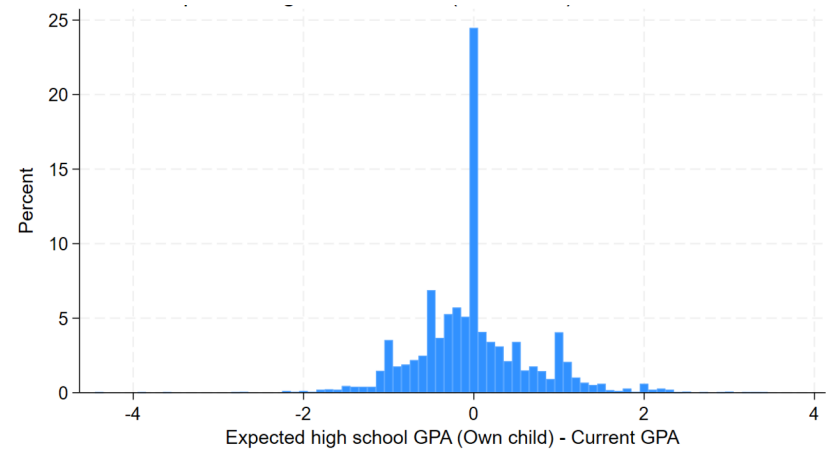
Moreover, parents are more positively biased or optimistic towards their own children, and more realistic towards a typical child with the same current middle school GPA: only 15 percent expect that the 9th-grade GPA of a typical child will remain the same at the end of high school. Although parents received their children's 9th-grade GPA prior to the survey, only half of them could report the actual GPA. Many parents overestimate their children's middle school performance (Figure 18).

Then, using administrative data from the 2023-2024 cohort, we calculated the average high school GPA for each 0.3 bin of 9th-grade GPA and used it as the "actual" high school GPA. We conducted this exercise to demonstrate that parents tend to overestimate their children's expected GPA at the end of the State Matura (Figure 19). Only about 5 percent of parents correctly predict the expected GPA that their child will achieve at the end of the State Matura. Very few underestimate it, and the vast majority overestimate it.

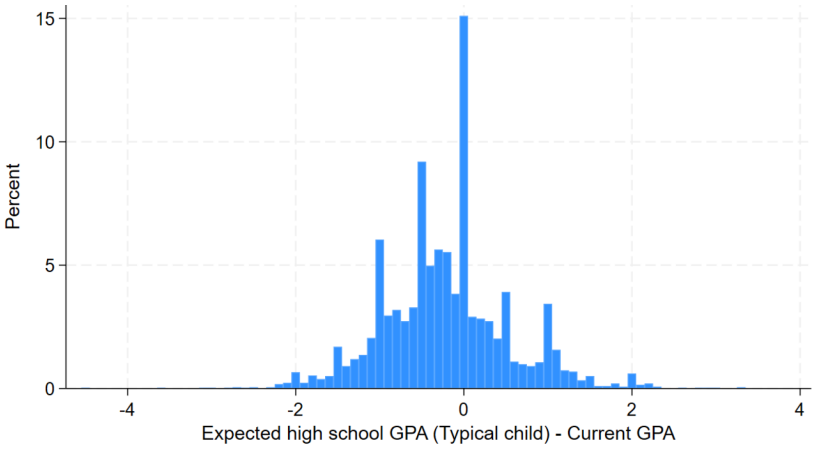
Given parents' overoptimism about their children's potential performance at the end of academic high school, we examine their beliefs about college eligibility, both for their own child and for a typical student with the same GPA, conditional on completing the State Matura. Given the self-reported middle school GPA, parents were asked to estimate the chances that a child with the same middle school GPA would be eligible for college, i.e., graduate from an academic high school with a GPA of at least 6.5. Then, they were asked to repeat the same exercise for different thresholds, corresponding to the eligibility criteria of different programs. We calculated the average probabilities of the high school GPA at the end of State Matura being more than or equal to each threshold for each 0.3 bin of the middle school GPA and used them as the "actual" probabilities. The probabilities are rounded to 10% (Figure 20).

We find substantial heterogeneity in parents' beliefs. Specifically, 24.52 percent of parents overestimate, 57.75 percent underestimate, and only 17.73 percent correctly assess the probability that a typical student with the same 9th-grade GPA as their child will attain a GPA of 6.5 by the time of the State Matura. Clearly, parents are pessimistic that a typical child with the same GPA as their child would be eligible for college.

While we did not elicit parents' beliefs about their own child's probability of reaching



(a) Own child



(b) Typical child

Figure 17: Expected high school GPA vs. current middle school GPA (own child and typical child)

Notes: Figures 23a and 23b display the distribution of the difference between the expected high school GPA of own child and of a typical child and the self-reported current 9th grade GPA.

the 6.5 eligibility threshold, we did ask analogous questions for higher GPA thresholds (Figure 21).

Parents substantially overestimate their child's academic prospects at the end of the State Matura. Specifically, 50.75 percent overestimate the likelihood of achieving a GPA of 7.0, rising to 68.83 percent for a GPA of 7.5 and 73.67 percent for a GPA of 8.0. Unsurprisingly, they are less optimistic about the chances of a typical child making it to college with the same GPA as their own child.

Parents' behavior appears counterintuitive because, as the threshold increases, college eligibility chances should decrease, at least for low to average-performing students. Two interpretations are consistent with these patterns. First, this might reflect parents' hope that their child will put in more effort in high school and meet the eligibility threshold.

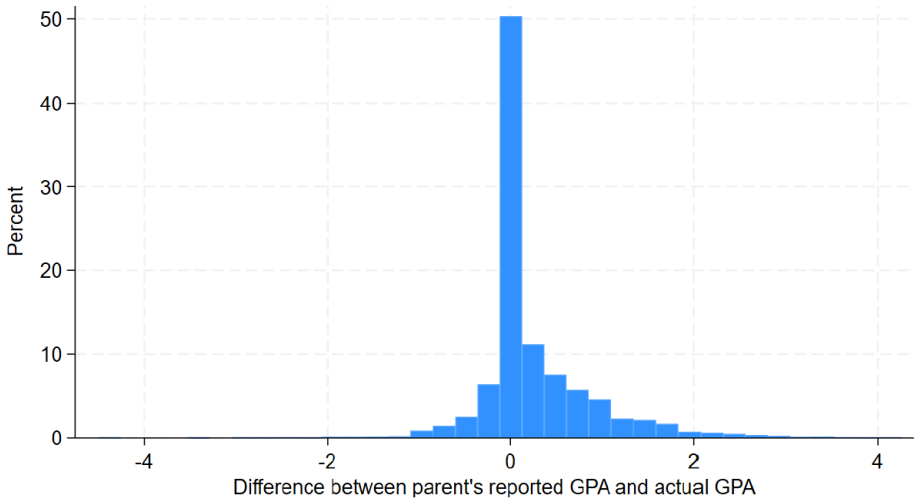


Figure 18: Actual vs. reported middle school GPA

Notes: Figure 18 displays the distribution of the difference between the actual and self-reported 9th grade GPA of own child

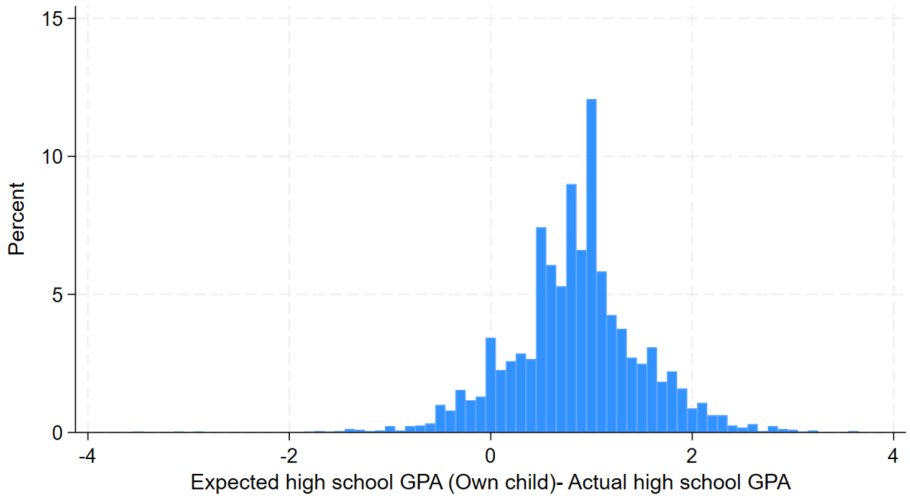


Figure 19: Expected vs. predicted (actual) high school GPA

Notes: Figure 19 displays the distribution of the difference between the expected and predicted (actual) high school GPA of own child

Aspirational hope in this case entails some degree of agency that parents and their children have towards reaching their goals (La Ferrara, 2019). For instance, many parents might pay for private courses or other academic services outside of the education system, hoping that their child's performance will improve. Additionally, they hope to be able to sufficiently encourage their child to put more effort into high school without needing to employ private tutors. Second, their beliefs might not entail any degree of agency, but simple overoptimism that their child is different from the statistical child, similarly to what wishful hoping refers to (Lybbert and Wydick, 2018).

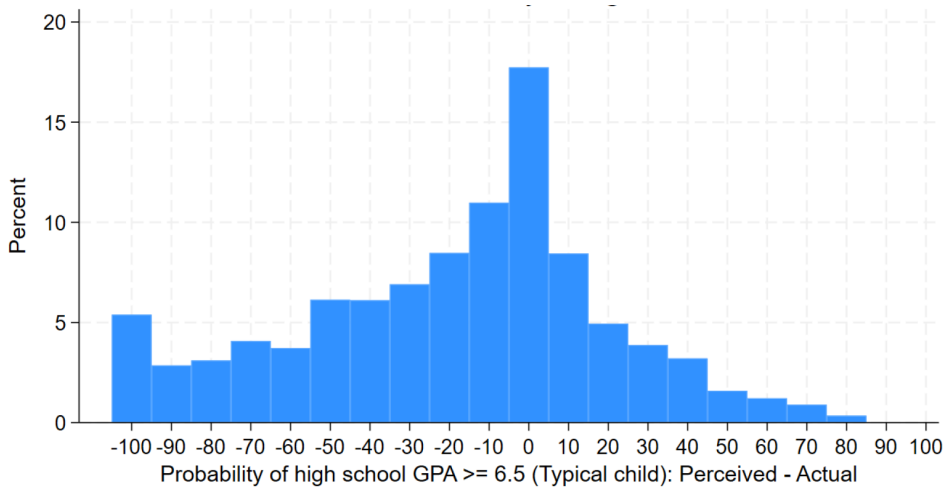


Figure 20: Predicted vs. (actual) college eligibility chances for a typical child

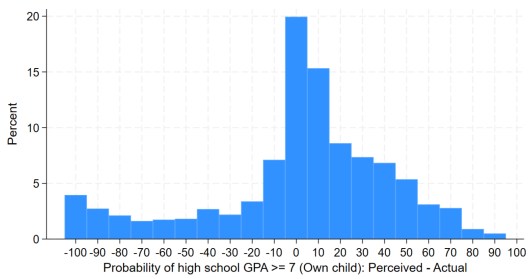
Notes: Figure 20 displays the distribution of the difference between the predicted and actual college eligibility chances for a typical child.

3.4 Other channels: parents' perceptions on low enrollment in the vocational track

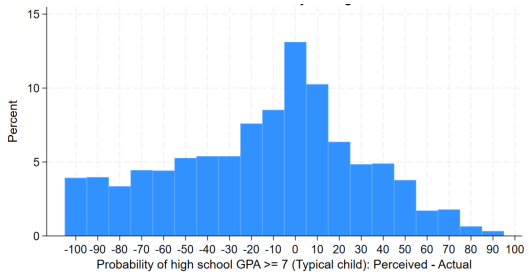
We elicited parents' beliefs about why other parents are reluctant to enroll their children in the vocational track. We consider three potential explanations: misperceived returns to education and employment opportunities across different educational tracks, misperceived stigma associated with vocational education, and incorrect beliefs that the vocational track does not allow access to college. Finally, we asked parents to share their opinions on additional (minor) factors.

We first focus on *perceived returns to education*. Parents were asked to estimate expected earnings associated with three educational pathways: vocational education only, academic high school only, and college completion. For each pathway, parents selected from predefined earnings categories expressed in Albanian lek. Parents' expectations regarding earnings in these hypothetical scenarios are reported in Figure 46 in the Appendix.

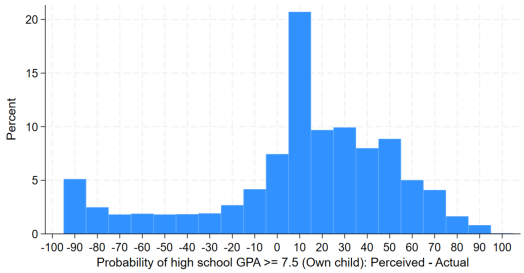
Overall, most parents have realistic expectations about their children's future earnings. Most correctly perceive that, on average, a college degree yields higher earnings than a vocational degree, which in turn is more financially rewarding than completing academic high school alone. Because parents may hold biased beliefs about their own child's future earnings, potentially due to private information about their child's abilities or employment prospects, we designed an alternative belief elicitation strategy based on hypothetical children. Specifically, we presented two scenarios. In the first scenario, a typical child has a zero percent probability of becoming eligible for college, and parents were asked to compare the expected earnings if the child attended a vocational school versus an academic high school. In the second, the typical child has a 100 percent probability of meeting college eligibility requirements, and parents were asked the same question. Parents' responses are



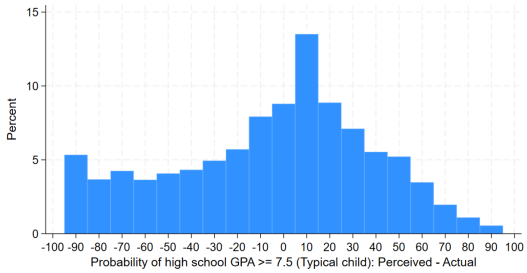
(a) Threshold 7.0; own child



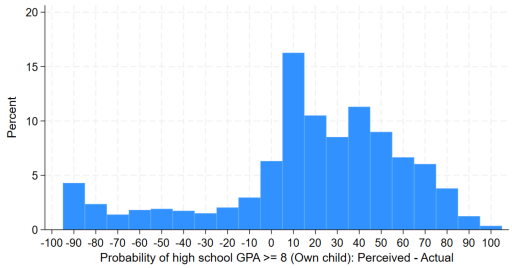
(b) Threshold 7.0; typical child



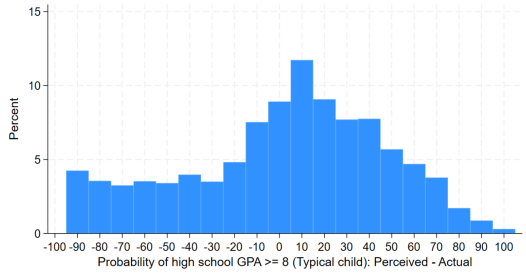
(c) Threshold 7.5; own child



(d) Threshold 7.5; typical child



(e) Threshold 8.0; own child



(f) Threshold 8.0; typical child

Figure 21: Predicted vs. (actual) college eligibility chances (different thresholds) for own and typical child

Notes: Figure 21 displays the distribution of the difference between the predicted and actual college eligibility chances for own child and for a typical child with the same GPA.

summarized in Figure 45 in Appendix C.1. When college eligibility chances are zero, more than 80 percent of parents believe that vocational education yields earnings that are equal to or higher than those from the academic track. Interestingly, even when college eligibility is certain, a substantial share of parents still views vocational education favorably: over 30 percent believe that vocational education yields earnings comparable to those obtained after completing the academic track and college, and approximately 40 percent believe vocational education yields higher earnings. Taken together, these results suggest that parents do not perceive vocational education as inferior to the academic track in terms of labor-market returns. On average, parental beliefs about earnings across educational pathways are well aligned with actual returns and are unlikely to explain low vocational enrollment. Similarly, parents' beliefs about employment chances after the academic and vocational tracks (without continuation to college). Figure 47 in Appendix C.1 suggests that overall, parents

believe that a VET graduate has more chances to find a job than an academic high school graduate. This might suggest that parents are aware that, in general, those who complete the secondary academic track often end up working in basic sales occupations, which are often characterized by low pay and low job security, whereas VET graduates are currently in higher demand.

Second, we elicit parents' beliefs on whether the *vocational track is stigmatized or has a low reputation* compared to the academic track (Figure 48 in Appendix C.1). Only 25 percent agree that the vocational track has a lower reputation than the academic one. The modal response (around 40 percent) is "Neutral", indicating that a large share of parents neither clearly endorse nor reject the idea of stigma. Roughly one-third disagree or strongly disagree, suggesting that many parents do not personally believe vocational education has a lower reputation. Only about one-quarter agree or strongly agree, and the share that strongly agrees is very small. Parents' own beliefs do not indicate a widespread perception of low reputation or stigma. If anything, the dominant feature is ambiguity rather than conviction, rejecting the hypothesis of an internalized stigma. Panel (b) asks parents what percentage of parents in their city or town believe vocational education has a lower reputation. The distribution is centered roughly around 40–-50%, with substantial dispersion. Very few parents believe that almost everyone thinks vocational education has a low reputation, but even fewer think that almost no one does. Parents do not overwhelmingly believe vocational education has a low reputation themselves. However, they believe that many other parents do, which can still discourage enrollment if decisions are shaped by social norms, reputation concerns, or fear of social judgment. To clear out that misperceptions about others might affect track choice, we asked parents to select (based on their opinions and beliefs) one or more factors as potential reasons explaining low enrollment in the vocational track (Table 39 in Appendix C.1).

Taken together, these results suggest that stigma-related concerns are unlikely to be a first-order explanation for low enrollment in the vocational track. Although parents believe that a non-trivial share of other parents view vocational education as having a lower reputation, this belief does not translate into strong endorsement of reputation-based mechanisms when parents are asked to explain vocational avoidance. In particular, explicitly stigma-related channels, such as fear that a child will be perceived as a low achiever or concerns about social peer quality, are selected by relatively few parents. While some parents express concerns about future opportunities, the overall pattern of responses provides little descriptive support for the idea that misperceived social stigma is a central factor shaping track choice. Hence, even if second-order misperceptions exist, they do not appear salient enough to shape perceived enrollment decisions.

Third, we test the hypothesis that parents may hold incorrect beliefs that the *vocational track does not allow access to college*. We asked parents whether they think a child can apply to college after completing vocational high school (Table 40 in Appendix C.1). Most parents (56.23 percent) correctly believe that VET graduates can only apply to selected college programs that match their VET curricula. Only 7.13 percent wrongly believe that vocational education fully prohibits access to college. Surprisingly, many parents (36.64 percent) think that VET graduates can apply to any college. It is possible that these parents mean by "any college" all the programs that are related to the VET degree.

3.4.1 Actual track choice by actual GPA bin

Here, we describe the actual track choice by actual GPA bin (Figure 22). We received data on the actual grade 9 GPA and the track choice. Very few parents enroll their kids in the vocational track. As GPA bins increase, enrollments in the academic track increase, and those in VET decline. Even for the low GPA bins (way below the eligibility threshold), the academic track is more preferred than the vocational one.

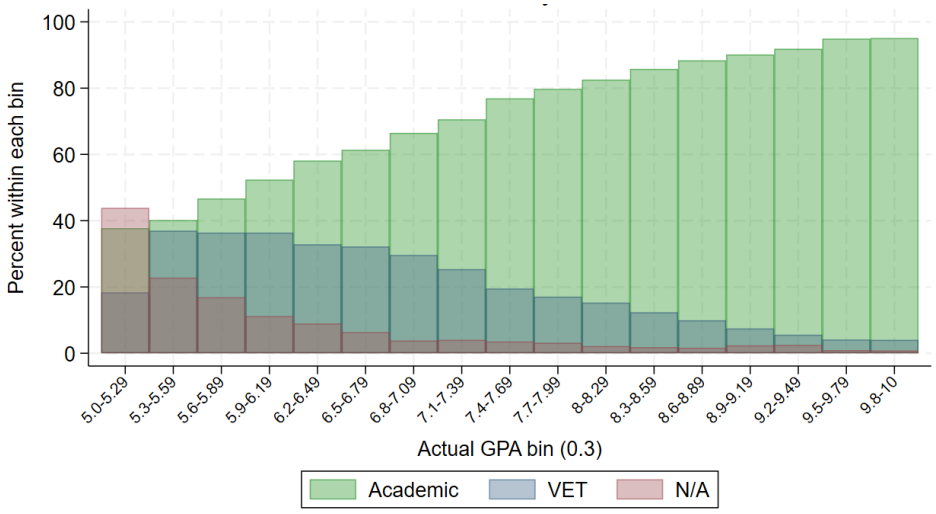
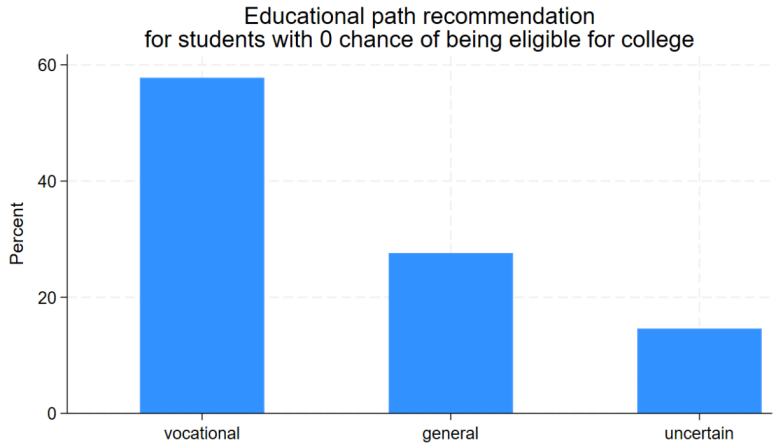


Figure 22: Actual track choice by actual GPA bin

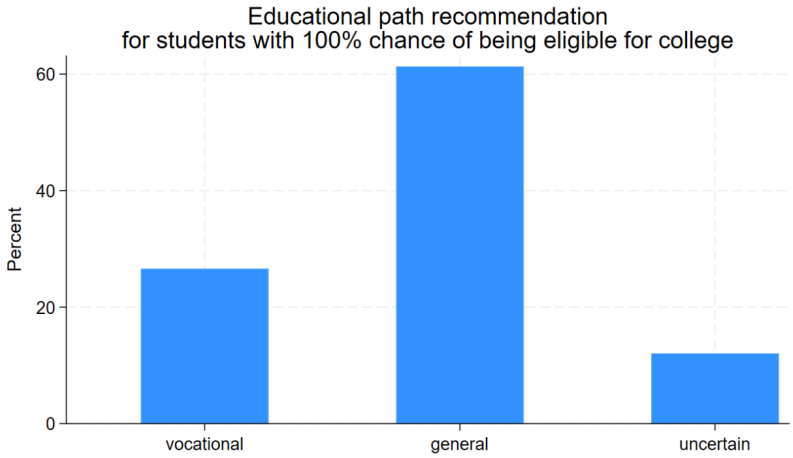
Notes: "Academic" means academic high school, "VET" means vocational school, and "N/A" is outside of the Albanian education system.

If, on average, the GPA from middle school to the end of high school drops by almost 1 point (precisely by 0.87 points) , enrollment in the academic track should be optimal for those with an actual GPA of at least 7.4. However, this is not the case. In the overall sample, around 14 percent aspire only to vocational education, around 22 percent aspire only to the academic high school, and 62 percent aspire to college education. The rest do not aspire to secondary education for their children at all. The national VET enrollment rate has been around 17-18 percent in the last four years (see Figure 16), and it accounts for those who graduated from private middle schools (which are excluded from our sample).

Conditional on knowing whether a child will be eligible for college, most parents provide accurate recommendations (Figure 23). Around 60 percent would recommend the vocational track to those with zero chances of being eligible for college, compared to around 30 percent who would still recommend the academic track. These statistics are reversed when it is known that college eligibility is guaranteed: around 60 percent would recommend the academic track, and around 30 percent would recommend the vocational track. This may be an important step in designing career orientation programs based on accurate predictions. We discuss the potential intervention in the next subsection.



(a) Not eligible



(b) Fully eligible

Figure 23: Parents' recommendations conditional on knowing college eligibility for sure

Notes: Figure 23 displays the track choice that parents would recommend for children who are and are not eligible for college.

3.5 Discussion: Which Education Policies Should Be in Order?

In the absence of rigid tracking systems where students are placed into tracks based on GPA and standardized test scores, some countries have adopted the system of (binding) teacher recommendations (e.g., Germany, France, Austria, Switzerland). One negative aspect of this approach is that teachers may have stereotypes or biased beliefs (Dustmann et al., 2017; Carlana, 2019; Carlana et al., 2022a,b) that prevent them from unlocking the full potential of their students, or they neglect the simple fact that some students demonstrate academic excellence in their later education stages.

Combining the concerns about teacher recommendations and this set of informative results, we are currently collaborating with the Albanian Ministry of Education for an in-

tervention based on individualized statistical information as a tool to help parents adjust their unrealistic aspirations and make better track choices. We plan to run a nationwide natural field experiment in over 250 public middle schools in Albanian municipalities that offer the three secondary education tracks (in collaboration with the Ministry of Education). The main intervention involves informing parents of children in the final grade of middle school about the statistical likelihood that their children will be eligible for college if they complete a general high school education (academic track). Our approach is based on rigorous and stereotype-free recommendations that come from hard evidence rather than teacher perceptions.

While the existing literature has focused on raising aspirations through the use of role models (Nguyen, 2008; Porter and Serra, 2020), career orientation programs, tutoring, and mentoring (Rodriguez-Planas, 2012; Martins, 2014; Goux et al., 2017; Carlana et al., 2022b), our experimental project aims to investigate how adjusting aspirations about future education through *individualized statistical predictions* about the feasibility of parental aspirations impacts the current human capital decisions they make for their children. The literature on individualized information provision and human capital investment in the education context has not gone beyond exploring the role of report cards (Andrabi et al., 2017; Dizon-Ross, 2019), and we are unaware of another study that has aimed to change aspirations through individualized predictions (see La Ferrara, 2019 for a review).

3.6 Conclusion

This paper documents substantial parental misperceptions about children's academic trajectories and college eligibility. These misperceptions coexist with widespread enrollment in the academic track, even among students whose prior academic performance makes college eligibility unlikely. While we do not causally identify the effect of beliefs on track choice, the patterns in the data - together with limited support for alternative explanations - suggest that information frictions play an important role in shaping educational decisions.

Our findings highlight the role of information frictions in education systems where secondary track choice is unconstrained but access to higher education is selective. Policies that provide parents with clear, data-driven projections of academic progression and realistic college eligibility - rather than generic encouragement or stigma-reduction campaigns - may improve educational sorting, reduce inefficient schooling choices, and better align students' pathways with their academic prospects.

Summary

This dissertation studies behavioral and rational mechanisms shaping prosocial behavior, effort provision, and human capital decisions. The first two chapters examine how autonomy affects charitable giving and volunteering effort, distinguishing whether people value agency intrinsically or use it as a self-selection mechanism. The third chapter analyzes parental misperceptions in secondary school track choice in Albania and their implications for educational investment.

Chapter one presents a field experiment in which employees of large Albanian companies were invited to donate to projects administered by Down Syndrome Albania. I vary whether donors can choose the project their donation supports and the amount of information provided. Allowing donors to allocate their gifts substantially increases giving, while the opportunity to obtain additional information on the effectiveness of these projects does not. The evidence indicates that autonomy raises donations primarily because donors can direct funds toward projects they prefer, rather than because they derive intrinsic utility from having choice itself.

Chapter two studies autonomy in a large-scale natural field experiment with 4,390 student volunteers collaborating with a UNICEF-funded NGO. Participants either chose which social issue to write about or were randomly assigned a topic. Task choice increases effort by 33 percent and improves effort quality. I show that the effect operates through ability-task matching: volunteers use autonomy to select tasks they feel more competent performing, rather than responding to intrinsic preferences for choice or heightened commitment.

Chapter three documents systematic parental overestimation of their own child's academic prospects at the time of track choice in Albania. Using administrative data and a large-scale survey, we show that parents are overly optimistic about their child's likelihood of meeting college eligibility thresholds, while holding more pessimistic views about similarly performing peers. These misperceptions coexist with high enrollment in academic tracks among students unlikely to qualify for college, suggesting that biased beliefs and information frictions may contribute to inefficient human capital allocation.

References

Aknin, L. B., Dunn, E. W., and Norton, M. I. (2012). Happiness runs in a circular motion: Evidence for a positive feedback loop between prosocial spending and happiness. Journal of Happiness Studies, 13(2):347–355.

Al-Ubaydli, O. and Lee, M. (2011). Can tailored communications motivate environmental volunteers? A natural field experiment. American Economic Review, 101(3):323–328.

Andrabi, T., Das, J., and Khwaja, A. I. (2017). Report cards: The impact of providing school and child test scores on educational markets. American Economic Review, 107(6):1535–1563.

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. The Economic Journal, 100(401):464–477.

Aretz, B. and Kube, S. (2013). Choosing your object of benevolence: A field experiment on donation options. The Scandinavian Journal of Economics, 115(1):62–73.

Ariely, D., Bracha, A., and Meier, S. (2009). Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. American Economic Review, 99(1):544–555.

Ashraf, N., Bandiera, O., and Jack, B. K. (2014). No margin, no mission? A field experiment on incentives for public service delivery. Journal of Public Economics, 120:1–17.

Ashraf, N., Bandiera, O., Minni, V., and Zingales, L. (2025). Meaning at work. Working paper w33843, National Bureau of Economic Research.

Bénabou, R. and Tirole, J. (2003). Intrinsic and extrinsic motivation. The Review of Economic Studies, 70(3):489–520.

Bidee, J., Vantilborgh, T., Pepermans, R., Huybrechts, G., Willems, J., Jegers, M., and Hofmans, J. (2013). Autonomous motivation stimulates volunteers' work effort: A self-determination theory approach to volunteerism. Voluntas: International Journal of Voluntary and Nonprofit Organizations, 24(1):32–47.

Bloom, N., Liang, J., Roberts, J., and Ying, J. Z. (2015). Does working from home work? Evidence from a Chinese experiment. The Quarterly Journal of Economics, 130(1):165–218.

Borgloh, S., Dannenberg, A., and Aretz, B. (2013). Small is beautiful – Experimental evidence of donors' preferences for charities. Economics Letters, 120(2):242–244.

Brock, J. M., Lange, A., and Ozbay, E. Y. (2013). Dictating the risk: Experimental evidence on giving in risky environments. American Economic Review, 103(1):415–437.

Brown, A. L., Meer, J., and Williams, J. F. (2019). Why do people volunteer? An experimental analysis of preferences for time donations. Management Science, 65(4):1455–1468.

Bursztyn, L. and Yang, D. Y. (2022). Misperceptions about others. Annual Review of Economics, 14(1):425–452.

Carlana, M. (2019). Implicit stereotypes: Evidence from teachers' gender bias. The Quarterly Journal of Economics, 134(3):1163–1224.

Carlana, M., Ferrara, E. L., and Pinotti, P. (2022a). Implicit stereotypes in teachers' track recommendations. In AEA Papers and Proceedings, volume 112, pages 409–414. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Carlana, M., La Ferrara, E., and Pinotti, P. (2022b). Goals and gaps: Educational careers of immigrant children. Econometrica, 90(1):1–29.

Cassidy, H. and Gaulke, A. (2024). The increasing penalty to occupation-education mismatch. Economic Inquiry, 62(2):607–632.

Cervantes, C. V. and Cooper, R. (2022). Labor market implications of education mismatch. European Economic Review, 148:104179.

Chandler, D. and Kapelner, A. (2013). Breaking monotony with meaning: Motivation in crowdsourcing markets. Journal of Economic Behavior & Organization, 90:123–133.

Charness, G., Gneezy, U., and Henderson, A. (2018). Experimental methods: Measuring effort in economics experiments. Journal of Economic Behavior & Organization, 149:74–87.

Charness, G. and Grieco, D. (2019). Creativity and incentives. Journal of the European Economic Association, 17(2):454–496.

Clary, E. G., Snyder, M., Ridge, R. D., Copeland, J., Stukas, A. A., Haugen, J., and Miene, P. (1998). Understanding and assessing the motivations of volunteers: A functional approach. Journal of Personality and Social psychology, 74(6):1516.

Conrads, J., Irlenbusch, B., Reggiani, T., Rilke, R. M., and Sliwka, D. (2016). How to hire helpers? Evidence from a field experiment. Experimental Economics, 19(3):577–594.

Cryder, C. E., Loewenstein, G., and Scheines, R. (2013). The donor is in the details. Organizational Behavior and Human Decision Processes, 120(1):15–23.

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. Journal of Personality and Social Psychology, 18(1):105–115.

DellaVigna, S. and Pope, D. (2018). What motivates effort? Evidence and expert forecasts. The Review of Economic Studies, 85(2):1029–1069.

Dizon-Ross, R. (2019). Parents' beliefs about their children's academic ability: Implications for educational investments. American Economic Review, 109(8):2728–2765.

Drouvelis, M. and Marx, B. M. (2022). Can charitable appeals identify and exploit belief heterogeneity? Journal of Economic Behavior & Organization, 198:631–649.

Duncan, B. (2004). A theory of impact philanthropy. Journal of Public Economics, 88(9-10):2159–2180.

Dustmann, C., Puhani, P. A., and Schönberg, U. (2017). The long-term effects of early track choice. The Economic Journal, 127(603):1348–1380.

Eckel, C. C., Herberich, D. H., and Meer, J. (2017). A field experiment on directed giving at a public university. Journal of Behavioral and Experimental Economics, 66:66–71.

Exley, C. L. (2016). Excusing selfishness in charitable giving: The role of risk. The Review of Economic Studies, 83(2):587–628.

Frey, B. S. and Gallus, J. (2017). Volunteer organizations: Motivating with awards. Economic Psychology, pages 273–286.

Freyer, T., Radbruch, J., Schaube, S., and Strang, L. (2025). The effect of task (mis) matching and self-selection on intrinsic motivation and performance. Journal of Economic Behavior & Organization, 234:107010.

Garcia, T., Massoni, S., and Villeval, M. C. (2020). Ambiguity and excuse-driven behavior in charitable giving. European Economic Review, 124:103412.

Georg, S. and Kube, S. (2012). Goals (th) at work–goals, monetary incentives, and workers' performance. Mpi collective goods preprint 2012/19, *Available at SSRN 2159663*.

Giving USA (2020). Giving usa 2019: The annual report on philanthropy for the year 2019. AAFCR Trust for Philanthropy, Indianapolis, Indiana.

Gneezy, U., Keenan, E. A., and Gneezy, A. (2014). Avoiding overhead aversion in charity. Science, 346(6209):632–635.

Gneezy, U. and Rustichini, A. (2000). Pay enough or don't pay at all. The Quarterly Journal of Economics, 115(3):791–810.

Goldsmith-Pinkham, P., Hull, P., and Kolesár, M. (2024). Contamination bias in linear regressions. American Economic Review, 114(12):4015–4051.

Goux, D., Gurgand, M., and Maurin, E. (2017). Adjusting your dreams? High school plans and dropout behaviour. The Economic Journal, 127(602):1025–1046.

Hackman, J. R. and Oldham, G. R. (1976). Motivation through the design of work: Test of a theory. Organizational Behavior and Human Performance, 16(2):250–279.

Heist, H. D. and Cnaan, R. A. (2018). Price and agency effects on charitable giving behavior. Journal of Behavioral and Experimental Economics, 77:129–138.

Hustinx, L., Haski-Leventhal, D., Handy, F., et al. (2008). One of a kind? Comparing episodic and regular volunteers at the Philadelphia Ronald McDonald House. International Journal of Volunteer Administration, 25(3):50–66.

Imbens, G. W. and Rubin, D. B. (2015). Causal inference in statistics, social, and biomedical sciences. Cambridge University Press.

Jasper, C. R. and Samek, A. S. (2014). Increasing charitable giving in the developed world. Oxford Review of Economic Policy, 30(4):680–696.

Kamdar, A., Levitt, S. D., List, J. A., Mullaney, B., and Syverson, C. (2015). Once and done: Leveraging behavioral economics to increase charitable contributions. Science of Philanthropy Initiative Working Paper No. 25.

Karlan, D. and Wood, D. H. (2017). The effect of effectiveness: Donor response to aid effectiveness in a direct mail fundraising experiment. Journal of Behavioral and Experimental Economics, 66:1–8.

Kessler, J. B., Milkman, K. L., and Zhang, C. Y. (2019). Getting the rich and powerful to give. Management Science, 65(9):4049–4062.

Kosfeld, M. and Neckermann, S. (2011). Getting more work for nothing? Symbolic awards and worker performance. American Economic Journal: Microeconomics, 3(3):86–99.

Krawczyk, M. and Le Lec, F. (2010). 'Give me a chance!' An experiment in social decision under risk. Experimental Economics, 13(4):500–511.

Krügel, J. P. and Meemann, C. (2024). Money or autonomy? evidence from a gift-exchange experiment. SSRN Working Paper, No. 5009025.

La Ferrara, E. (2019). Presidential address: Aspirations, social norms, and development. Journal of the European Economic Association, 17(6):1687–1722.

Lamberton, C. (2013). A spoonful of choice: How allocation increases satisfaction with tax payments. Journal of Public Policy & Marketing, 32(2):223–238.

Lamberton, C. P., De Neve, J.-E., and Norton, M. I. (2014). Eliciting taxpayer preferences increases tax compliance. SSRN Working Paper No. 2365751.

Leuven, E. and Oosterbeek, H. (2011). Overeducation and mismatch in the labor market. Handbook of the Economics of Education, 4:283–326.

Levitt, S. D. and Neckermann, S. (2014). What field experiments have and have not taught us about managing workers. Oxford Review of Economic Policy, 30(4):639–657.

Li, S. X., Eckel, C., Grossman, P. J., and Brown, T. L. (2015). Directed giving enhances voluntary giving to government. Economics Letters, 133:51–54.

Llena-Nozal, A. (2009). The effect of work status and working conditions on mental health in four OECD countries. National Institute Economic Review, 209(1):72–87.

Lybbert, T. J. and Wydick, B. (2018). Poverty, aspirations, and the economics of hope. Economic Development and Cultural Change, 66(4):709–753.

Martins, P. S. (2014). Can targeted, non-cognitive skills programs improve achievement? Evidence from EPIS. IZA Discussion Papers No. 5266.

Mertins, V. and Walter, C. (2021). In absence of money: a field experiment on volunteer work motivation. Experimental Economics, 24(3):952–984.

Metzger, L. and Günther, I. (2019). Making an impact? The relevance of information on aid effectiveness for charitable giving. A laboratory experiment. Journal of Development Economics, 136:18–33.

Neckermann, S., Cueni, R., and Frey, B. S. (2014). Awards at work. Labour Economics, 31:205–217.

Nguyen, T. (2008). Information, role models and perceived returns to education: Experimental evidence from madagascar. Unpublished Manuscript, 6(5).

Porter, C. and Serra, D. (2020). Gender differences in the choice of major: The importance of female role models. American Economic Journal: Applied Economics, 12(3):226–254.

Rodriguez-Planas, N. (2012). Longer-term impacts of mentoring, educational services, and learning incentives: Evidence from a randomized trial in the United States. American Economic Journal: Applied Economics, 4(4):121–139.

Romano, J. P. and Wolf, M. (2005). Stepwise multiple testing as formalized data snooping. Econometrica, 73(4):1237–1282.

Ryan, R. M. and Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American Psychologist, 55(1):68–78.

Whillans, A. V. (2016). A brief introduction to the science of fundraising. Departement of Psychology, The University of British Columbia.

Wu, C. and Eftekhar, M. (2024). Does volunteering crowd out donations? Evidence from online experiments. Manufacturing & Service Operations Management, 26(4):1542–1566.

List of Figures

List of Tables

Annexes

A Allocation Choice in Charitable Giving: A Natural Field Experiment

A.1 Sample characteristics

Here, we display basic statistics regarding the sample, including the response rate, characteristics of the employees, and those of the donors.

Table 4: Response statistics

	Control	Allocation Choice	Forced Allocation	Forced Allocation & Link
Observations	278	261	233	270
Response Rate	22.22%	20.88%	18.64%	21.60%
Choice take-up	NA	74.55%	NA	NA
Clicking rate	NA	NA	NA	33.64%

Note: Columns 2-3 indicate the mean, standard deviation and the number of observations for each of the characteristics

Table 5: Sample characteristics

Characteristics	Mean	SD	N
Gender (not specified)	0.022	0.147	1042
Female	0.582	0.494	1042
Male	0.396	9.489	1042
Age	35.670	9.974	1042
Single	0.357	0.479	1042
Married	0.544	0.489	1042
Cohabiting	0.072	0.259	1042
Divorced	0.027	0.162	1042
Undergraduate	0.025	0.156	1042
Graduate	0.183	0.387	1042
Master	0.697	0.460	1042
Doctoral	0.095	0.293	1042
Parent	0.556	0.497	1042
Specialist	0.632	0.482	1042
Manager	0.289	0.453	1042
Executive	0.079	0.269	1042
Aware of PID	0.403	0.491	1042
Aware of DSA	0.602	0.490	1042
Donated to PID	0.575	0.495	1042
Donated to DSA	0.124	0.330	1042

Table 6: Prediction of the likelihood to donate

	(1)	(2)
Predictors	likelihood	SE
Sex not specified	-0.071**	(0.032)
Female	-0.134	(0.093)
Age	0.007***	(0.002)
Married	0.018	(0.063)
Cohabiting	0.148**	(0.063)
Divorced	-0.006	(0.113)
Graduate	0.201**	(0.080)
Master	0.198***	(0.076)
Doctoral	0.092	(0.092)
Manager	0.100***	(0.036)
Executive	0.106*	(0.061)
Parent	0.032	(0.063)
Aware of PID	-0.068**	(0.031)
Aware of DSA	-0.002	(0.033)
Ever donated to PID	0.104***	(0.032)
Ever donated to DSA	0.021	(0.050)
Mean likelihood	43.67%	-
Observations	1,042	
R-squared	0.077	

Note: Robust standard errors in parentheses; *** p < 0.01, ** p < 0.05, * p < 0.1

A.2 Randomization Check

Table 7: Randomization Check (OLS results of comparing each treatment to the control group

	(1)	(2)	(3)
Covariates	choice	force	link
Gender (not specified)	-0.228*	-0.032	-0.158
	(0.131)	(0.139)	(0.136)
Female	0.030	0.007	-0.025
	(0.047)	(0.047)	(0.046)
Married	-0.036	-0.090	0.015
	(0.100)	(0.103)	(0.088)
Cohabiting	-0.009	-0.036	0.078
	(0.094)	(0.095)	(0.087)
Divorced	-0.047	-0.306**	-0.356***
	(0.141)	(0.145)	(0.132)
Graduate	-0.194	-0.025	-0.065
	(0.140)	(0.172)	(0.147)
Master	-0.220*	-0.040	-0.095
	(0.133)	(0.166)	(0.143)
Doctoral	-0.214	0.045	-0.065
	(0.153)	(0.180)	(0.165)
Manager	0.013	-0.023	-0.047
	(0.050)	(0.053)	(0.051)
Executive	-0.039	0.022	0.028
	(0.089)	(0.088)	(0.083)
Age	0.003	-0.001	-0.001
	(0.003)	(0.003)	(0.003)
Parent	-0.016	0.039	-0.009
	(0.099)	(0.103)	(0.086)
Aware of PID	-0.031	-0.028	0.059
	(0.048)	(0.049)	(0.045)
Ever donated to PID	0.036	-0.017	0.028
	(0.048)	(0.049)	(0.047)
Aware of DSA	0.037	-0.023	0.016
	(0.049)	(0.049)	(0.047)
Ever donated to DSA	-0.027	0.028	-0.005
	(0.074)	(0.075)	(0.072)
Joint orthogonality: p-value	0.823	0.867	0.122
Observations	539	511	548

Note: Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 8: Randomization check: control and choice groups

	Control	Allocation choice	Difference	p-value	N
Gender (not specified)	0.032	0.011	0.022	0.101	537
Female	0.576	0.617	-0.041	0.330	537
Male	0.392	0.372	0.020	0.626	537
Age	35.924	36.479	-0.554	0.539	537
Single	0.334	0.352	-0.018	0.662	537
Married	0.557	0.540	0.017	0.667	537
Cohabiting	0.065	0.065	0.000	0.986	537
Divorced	0.043	0.042	0.001	0.953	537
Undergraduate	0.018	0.034	-0.016	0.230	537
Graduate	0.176	0.184	-0.008	0.818	537
Master	0.716	0.690	0.026	0.507	537
Doctoral	0.090	0.092	-0.002	0.935	537
Parent	0.579	0.567	0.012	0.777	537
Specialist	0.618	0.620	-0.002	0.962	537
Manager	0.306	0.318	-0.012	0.759	537
Executive	0.076	0.061	0.014	0.515	537
Aware of PID	0.406	0.387	0.019	0.645	537
Aware of DSA	0.597	0.636	-0.039	0.354	537
Donated to PID	0.565	0.590	-0.025	0.553	537
Donated to DSA	0.122	0.119	0.003	0.900	537

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 9: Randomization check: control and forced allocation groups

	Control	Forced Allocation	Difference	p-value	N
Gender (not specified)	0.032	0.026	0.006	0.659	509
Female	0.576	0.579	-0.003	0.930	509
Male	0.392	0.395	-0.003	0.949	509
Age	35.924	35.056	0.868	0.300	509
Single	0.334	0.399	-0.065	0.131	509
Married	0.557	0.519	0.038	0.389	509
Cohabiting	0.065	0.069	-0.004	0.860	509
Divorced	0.043	0.013	**0.030****	**0.043**	509
Undergraduate	0.018	0.021	-0.003	0.778	509
Graduate	0.176	0.180	-0.004	0.907	509
Master	0.716	0.687	0.029	0.474	509
Doctoral	0.090	0.111	-0.022	0.417	509
Parent	0.579	0.524	0.055	0.209	509
Specialist	0.618	0.644	-0.025	0.560	509
Manager	0.306	0.270	0.035	0.381	509
Executive	0.076	0.086	-0.010	0.670	509
Aware of PID	0.406	0.361	0.045	0.289	509
Aware of DSA	0.597	0.567	-0.030	0.486	509
Donated to PID	0.565	0.541	0.024	0.5880	509
Donated to DSA	0.122	0.124	-0.002	0.941	509

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 10: Randomization check: control and link groups

	Control	Link	Difference	p-value	N
Gender (not specified)	0.032	0.019	0.013	0.305	546
Female	0.576	0.555	-0.02	0.638	546
Male	0.392	0.426	-0.034	0.421	546
Age	35.924	35.156	0.769	0.352	546
Single	0.334	0.348	-0.014	0.737	546
Married	0.557	0.555	0.002	0.963	546
Cohabiting	0.065	0.089	-0.024	0.289	546
Divorced	0.043	0.007	**0.036*** **	**0.008**	546
Undergraduate	0.018	0.026	-0.008	0.526	546
Graduate	0.176	0.193	-0.016	0.623	546
Master	0.716	0.693	0.023	0.552	546
Doctoral	0.090	0.089	0.001	0.966	546
Parent	0.579	0.548	0.031	0.466	546
Specialist	0.618	0.648	-0.030	0.476	546
Manager	0.306	0.259	0.046	0.228	546
Executive	0.076	0.093	-0.017	0.473	546
Aware of PID	0.406	0.452	-0.045	0.284	546
Aware of DSA	0.597	0.604	-0.007	0.875	546
Donated to PID	0.565	0.600	-0.035	0.404	546
Donated to DSA	0.122	0.129	-0.007	0.797	546

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

A.3 Robustness Checks

Here, we present the robustness checks of the main results. We begin by demonstrating that the increase in donations does not come solely from those who donate large amounts, but also from small donors. Then, we show that our results are robust to outliers. Lastly, we show that the main treatment effects are robust to multiple hypothesis testing.

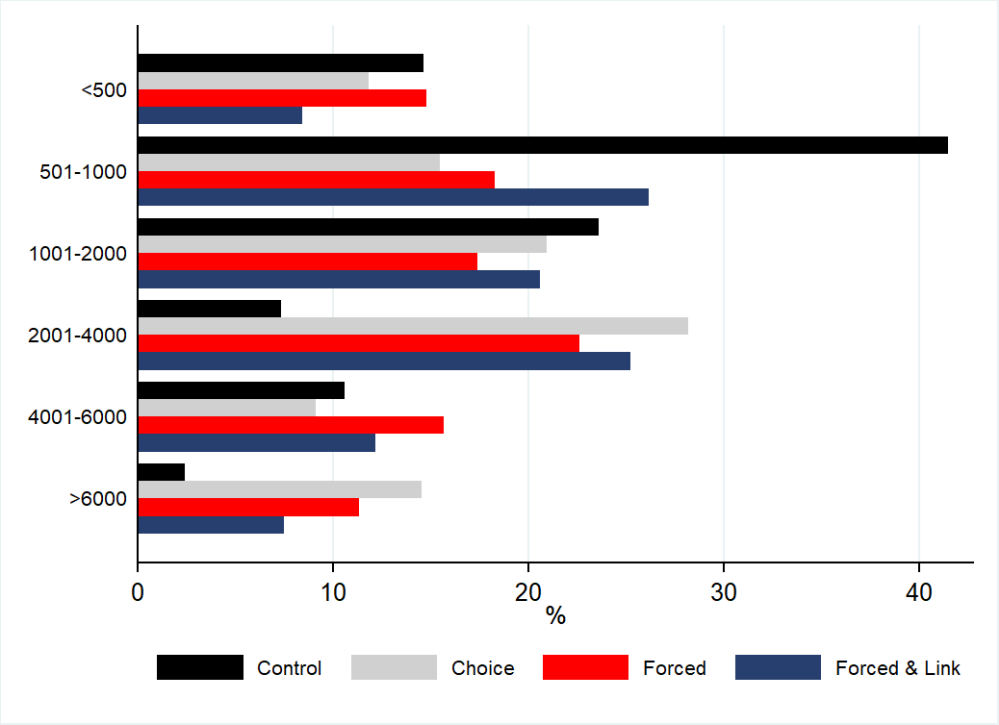


Figure 24: Percentage of gifts by size, conditional on giving

Table 11: Treatment effects on giving using winsorized donations

Treatment	Winsorized 1% Donation	Winsorized 2% Donation	Winsorized 5% Donation
Choice	577.568***	552.725***	314.073**
	(181.470)	(172.930)	(128.303)
Forced Allocation	757.791***	718.656***	499.907***
	(192.268)	(180.310)	(139.324)
Forced Alloc. & Link	293.734*	277.271*	168.286
	(156.173)	(148.303)	(121.060)
Control Mean	810.369***	802.005***	776.621***
	(87.646)	(84.238)	(77.160)
Observations	1,042	1,042	1,042

Note: Robust standard errors in parentheses; *** p < 0.01, ** p < 0.05, * p < 0.1

Table 12: Treatment effects controlling for the familywise error rate (FWER)

Treatment	(1) ATE	(2) p-value
Choice	656.564***	0.002
	(210.386)	
	[217.012]	0.006
Forced Allocation	934.770***	0.000
	(255.054)	
	[249.457]	0.001
Forced Alloc. & Link	424.674**	0.044
	(210.115)	
	[205.849]	0.051
Control mean	810.369	810.369
Observations	1,042	1,042

Note: Robust standard errors in parentheses and clustered standard errors in square brackets; The FWER p-values are

*** p < 0.01, ** p < 0.05, * p < 0.1

A.4 Sub-sample analysis

In this section, we present treatment effects for different sub-samples. This analysis is not causal, but it is used to support the discussion on mechanisms.

Table 13: Treatment effects on giving, conditional on giving

	(1)	(2)	(3)	(4)
Treatment	Donation (AL)	Donation (AL)	Donation (AL)	Donation (AL)
Choice	1,647.896***	1,574.994***	1,728.601***	1,722.406***
	(408.379)	(432.462)	(445.761)	(475.782)
Forced Allocation	1,703.068***	1,701.754***	1,736.320***	1,804.548***
	(453.019)	(427.886)	(460.817)	(431.718)
Forced Alloc. & Link	1,283.730***	1,367.728***	1,359.819***	1,392.493***
	(449.608)	(519.641)	(442.519)	(491.571)
Control mean	1,832.735	1,832.735	1,832.735	1,832.735
Firm FE	No	Yes	No	Yes
Covariates	No	No	Yes	Yes
Observations	455	455	455	455
Choice = Forced: p-value	0.923	0.821	0.990	0.895
Forced = Link: p-value	0.485	0.599	0.536	0.508

Note: Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 14: Differences in giving among sub-groups

Sub-group	Observations	Mean	Difference	p-value
Choice: opt-in (1)	83	3737.647		
Forced (2)	115	3535.803		
(1) - (2)			201.8441	0.754
Choice: opt-out (3)	29	2504.99		
Forced (4)	115	3535.803		
(3) - (4)			-1030.813	0.244
Link: no click (5)	71	3447.744		
Forced (6)	115	3535.803		
(5)-(6)			-88.0592	0.903
Link: click (7)	36	2463.108		
Forced (8)	115	3535.803		
(7) - (8)			-1072.695*	0.053
Control (9)	123	1832.735		
Link: no click (10)	71	3447.744		
(9) - (10)			-1615.009**	0.012
Control (11)	123	1832.735		
Link: click (12)	36	2463.108		
(11) - (12)			-630.3738*	0.068

Note: [a] The notation in parenthesis in each third row of the table represents the difference in giving among the two sub-groups. [b] For comparisons of mean donations between sub-groups 7-8 and 9-10 the two-way mean equality test is performed assuming that the variances are not equal. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 15: Choice take-up and information box click predictors

	(1)	(2)	(3)	(4)
Covariates	choice take-up	SE	link click	SE
Gender (not specified)	0.151	(0.117)	-	-
Female	-0.044	(0.093)	-0.141	(0.097)
Married	0.126	(0.174)	-0.180	(0.187)
Cohabiting	-0.094	(0.177)	-0.173	(0.183)
Divorced	-0.092	(0.375)	-0.615***	(0.199)
Graduate	-0.094	(0.140)	-0.203	(0.178)
Master	-0.228**	(0.102)	-0.306*	(0.169)
Doctoral	-0.419**	(0.198)	-0.294	(0.221)
Manager	0.072	(0.094)	0.091	(0.110)
Executive	0.237*	(0.121)	0.053	(0.138)
Age	0.006	(0.005)	-0.003	(0.005)
Parent	-0.239	(0.150)	0.040	(0.150)
Aware of PID	-0.133	(0.100)	-0.249**	(0.106)
Ever donated to PID	0.021	(0.104)	0.114	(0.107)
Aware of DSA	-0.088	(0.105)	0.060	(0.109)
Ever donated to DSA	0.072	(0.136)	-0.034	(0.157)
Mean take-up/click	0.741	-	0.336	-
Observations	112	-	107	-
R-squared	0.100	-	0.173	-

Note: Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

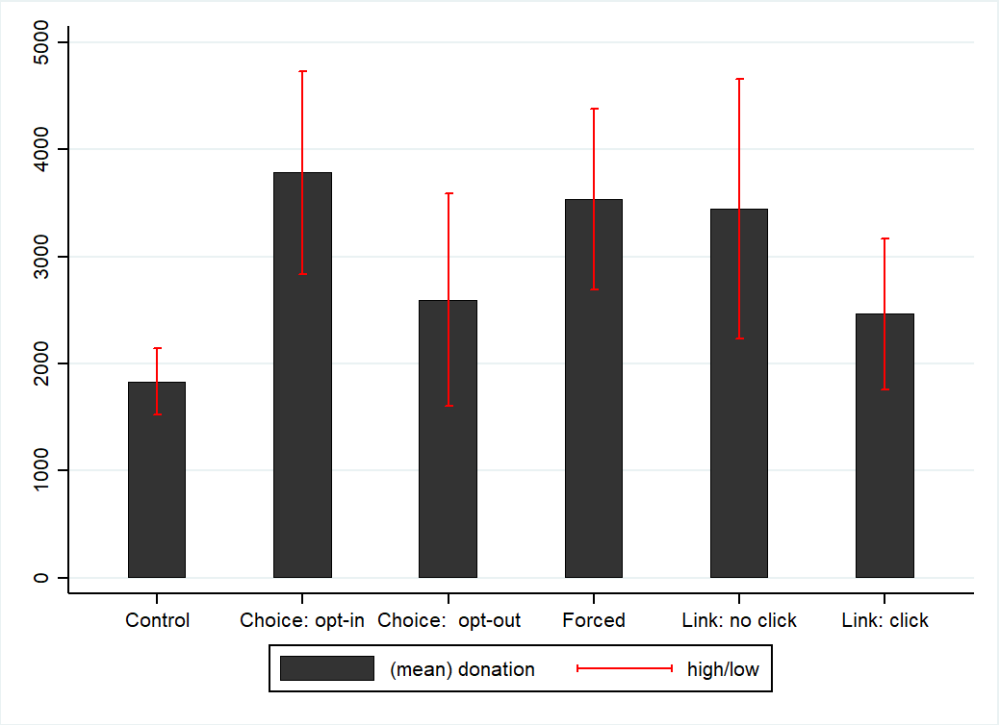


Figure 25: Mean donations by sub-treatment groups

A.5 Follow-up survey

The follow-up survey was conducted in English, but the language used in the main experiment was Albanian; therefore, the call donations and the information in the link were presented to the follow-up participants in their original language. The screenshots of the follow-up survey are displayed below:

In collaboration with Down Syndrome Albania Foundation (DSA), we ran a scientific fundraising campaign. The campaign aimed to raise funds to support three needs of children with Down Syndrome and other intellectual disabilities: therapy, entertainment, and humanitarian aid. We collaborated with the large and medium-sized companies in Albania **that have supported DSA in the past**, including commercial banks, financial institutions, private universities, and distribution and marketing firms, among others. We conducted a short survey with their administrative employees, where in the last question we asked them whether they wanted to donate, and if yes, we asked them to donate in 4 different ways. The invitation reads as follows:

> Fondacioni Down Syndrome Albania (DSA) dhe __(Emri i institucionit)_ po bashkëpunojnë për të mbledhur kontribute për 45 fëmijët me aftësi të kufizuara intelektuale që frekuentojnë qendrën e shërbimeve "PRO PAK" të DSA, për 3 nevoja të ndryshme:
>
> - Argëtim dhe organizim të festave në kuadër të 1 Qershorit dhe Festës së Vitit të Ri (përfshirë dhuratat)
> - Ndihmë humanitare për familjet e varfëra ku jeton një fëmijë me aftësi të kufizuara intelektuale (rroba, ushqim, ilaçe, materiale shkollore)
> - Sigurimin e shërbimeve terapeutike, për ata fëmijë prindërit e të cilëve nuk kanë mundësi të paguajnë (fizioterapi, logopedi, terapi ABA, terapi zhvillimi)

- We simply told to some employees that their donation supports entertainment, humanitarian aid, and therapy.
- We asked some employees **to choose whether they want to allocate** their contribution to entertainment, humanitarian aid, and therapy the way they preferred.
- We asked some employees **to allocate** their contribution to entertainment, humanitarian aid, and therapy the way they preferred.
- Lastly, we asked some employees **to click a link** containing detailed information on how entertainment, humanitarian aid, and therapy help children with Down Syndrome and other intellectual disabilities, before **allocating** their contribution to the three projects the way they preferred.

Through this survey, we would like to receive your feedback about the donors' behavior in this project.

Figure 26: Follow-up survey (1/6)

Have you taught marketing as a course through training or academic classes in the past?

- (●) Yes
- ( ) No

Are you aware of this fundraising campaign?

- ( ) Yes
- (●) No

Have you participated in this campaign?

- (●) Yes
- ( ) No

Now, we will ask you to guess the average donation of different employees, knowing that the ***average donation of those who were simply told that their contribution supports entertainment, humanitarian aid, and therapy is 810 Lek.*** Note that this number includes the zero-donations of those who decided not to donate.

We would like you to answer a few questions about the information in the link that we provided to some of the donors and guess the average donation in the link group.

- Nevoja e argëtimit synon gjithëpërfshirjen e fëmijëve me sindromën Down dhe aftësi të tjera të kufizuara intelektuale, duke mundësuar pranimin e tyre nga bashkëmoshatarët me zhvillim neurotipik dhe anëtarë të komunitetit. Nëpërmjet aktiviteteve gjithëpërfshirëse si mësimi i një vegle muzikore, piktura, teatri etj, apo nëpërmjet aktiviteteve festive dhe lojrave, fëmijët që përfitojmë nga shërbimet ndihen të njëjtë me bashkëmoshatarët e tyre, duke ndikuar kështu pozitivisht në përfshirjen e tyre sociale.
- Ndihma humanitare, konsiston kryesisht në veshje, ushqim apo medikamente për fëmijët me sindromën Down ose me aftësi te tjera të kufizuara intelektuale që jetojnë në familje të varfëra. Kjo mbështetje financiare lehtëson barrën financiare të këtyre familjeve duke rritur kështu mirëqenien në këto familje.
- Shërbimet terapeutike u mundësojnë fëmijëve me sindromën Down dhe aftësi të tjera të kufizuara intelektuale, një përmirësim në cilësinë e jetës, në aftësitë e të kuptuarit dhe të shprehurit, në sjelljen me familjarët dhe bashkëmoshatarët, si dhe në zhvillimin e tyre emocional. Trajtimi terapeutik në Qendrën e Shërbimeve PROPAK ka patur efekt afatgjatë te fëmijët të cilët trajtohen në mënyrë të rregullt dhe që në moshë të re.

Figure 27: Follow-up survey (2/6)

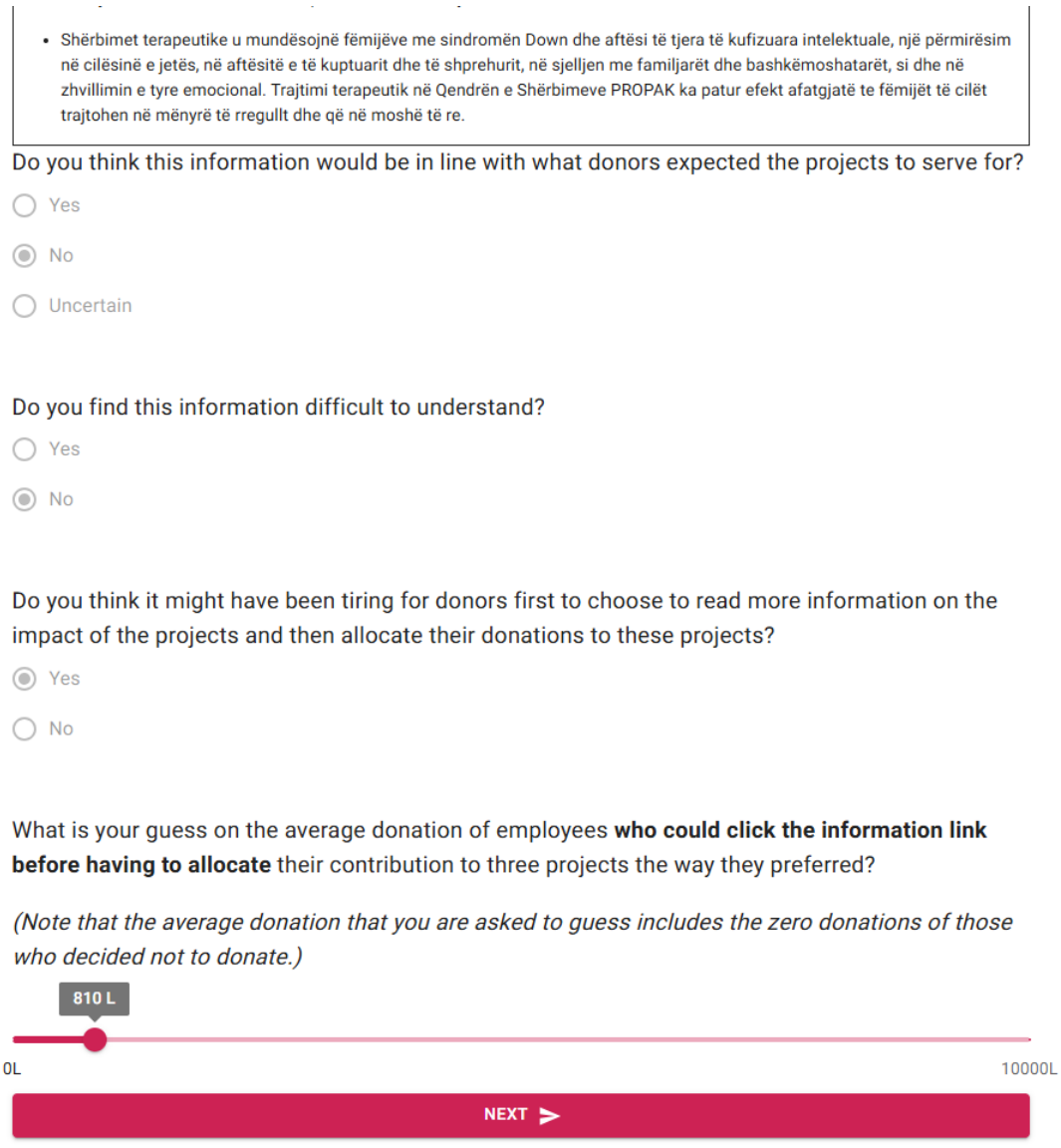


Figure 28: Follow-up survey (3/6)

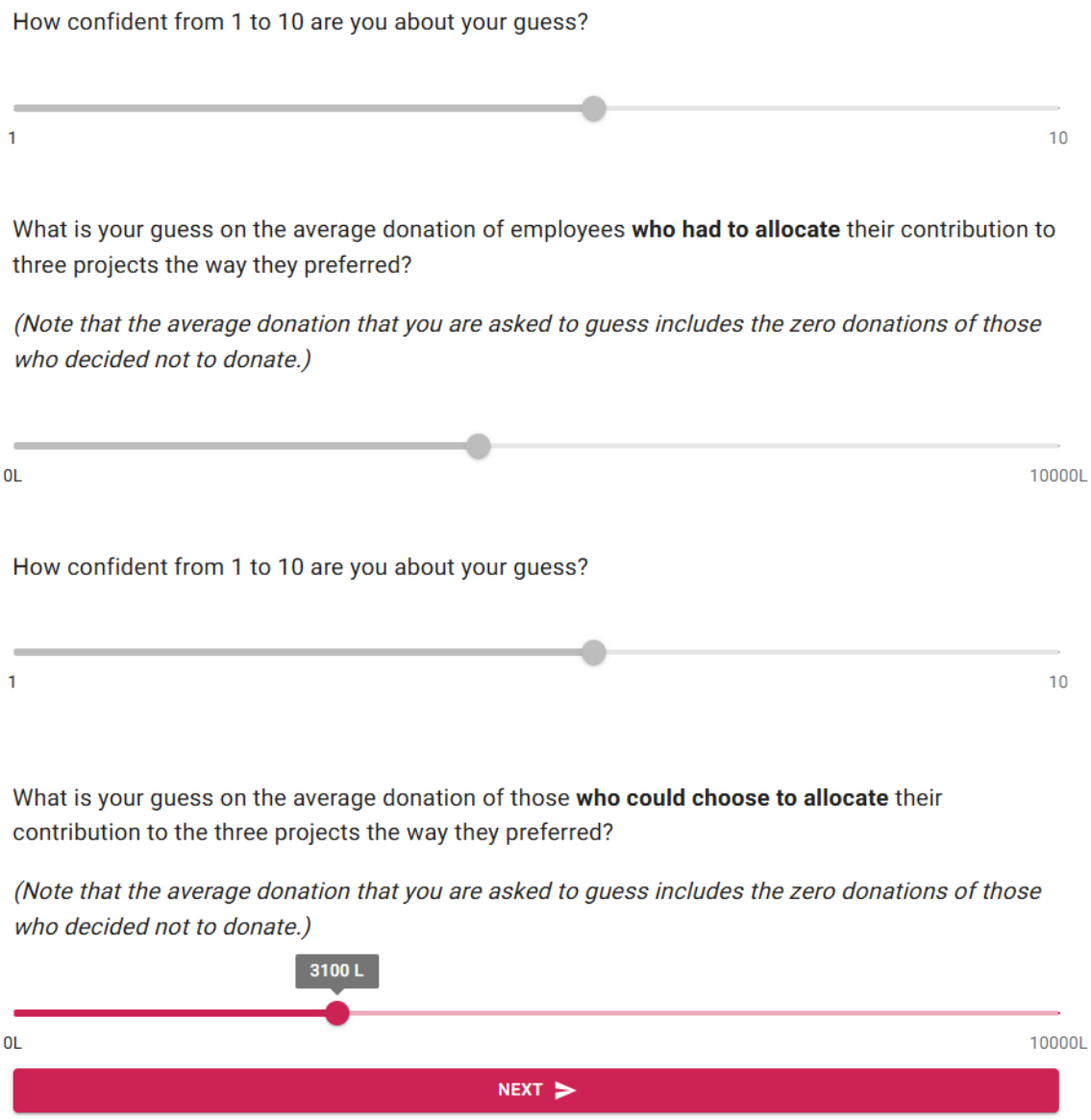


Figure 29: Follow-up survey (4/6)

What is your guess on the average donation of those **who could choose to allocate** their contribution to the three projects the way they preferred?

(Note that the average donation that you are asked to guess includes the zero donations of those who decided not to donate.)

0L 10000L

How confident from 1 to 10 are you about your guess?

1 10

If your guess falls above 810 Lek, why do you think donors **who could choose to allocate** their contribution to the three projects donated on average more than donors **who did not have the option to allocate their contribution**?

- ◯ They could donate more to the projects they liked more
- ◯ Donors like choice: they could allocate their donation if they wanted to, and vice-versa.
- ◯ They trusted the Foundation more because more choice signals the NGO's openness and transparency.
- ◯ Other

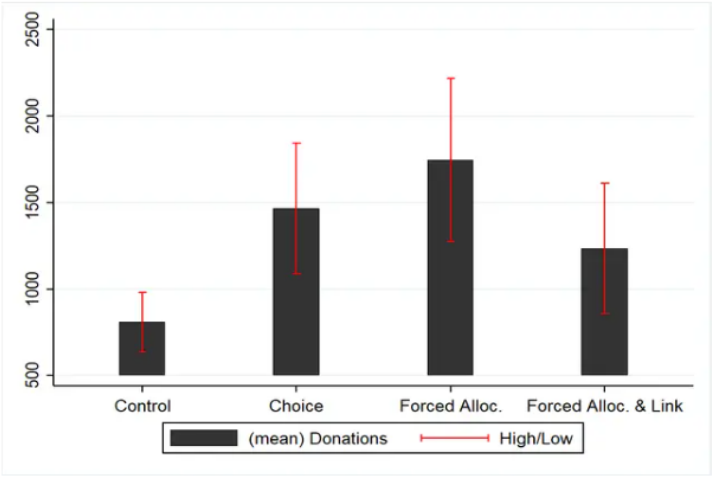
NEXT

Figure 30: Follow-up survey (5/6)

Thank you for your feedback.

Here we present the highlights of this study:

1. Donors who could choose to allocate their contributions to the three projects donated on average 1466 Lek, or 80.9 % more relative to donors who did not have the allocation option.
2. Donors who had to allocate their contributions to the three projects donated on average 1744 Lek, or two times more relative to donors who did not have the allocation option.
3. Donors who had to allocate their contributions to the three projects and that could receive more information about the impact of these projects donated on average 1234 Lek, or 52.3 % more relative to donors who did not have the allocation option.



Legend of the Graph:
Control – donors who did not have an allocation option
Choice – donors who could choose to allocate their donation
Forced alloc. – donors who had to allocate their contribution
Forced alloc. & Link – donors who had to allocate their contribution and that could receive more information on the impact of the three projects.

Figure 31: Follow-up survey (6/6)

Table 16: Follow-up survey: sample characteristics

Characteristics	Mean	SD	N
Marketing experts	0.28	0.450	200
Academic exp.	0.52	0.501	200
Aware	0.53	0.500	200
Participants	0.335	0.473	200

Table 17: Summary of actual average donations and guessed average donations

Group	Observations	Mean	Difference	p-value
Choice (1)	261	1466.93		
Guessed choice (2)	200	1940.75		
(1) - (2)			-473.82	NA
Allocation (3)	233	1745.14		
Guessed allocation (4)	200	1996.85		
(3) - (4)			-251.7	NA
Link (5)	270	1235.04		
Guessed Link (6)	200	1707.9		
(5)-(6)			-472.86	NA
Guessed choice (7)	200	1940.75		
Guessed allocation (8)	200	1996.85		
(7) - (8)			-56.1	0.496
Guessed allocation (9)	200	1996.85		
Guessed Link (10)	200	1707.9		
(9) - (10)			288.95***	0.000

Note: *** p <0.01, ** p < 0.05, * p < 0.1

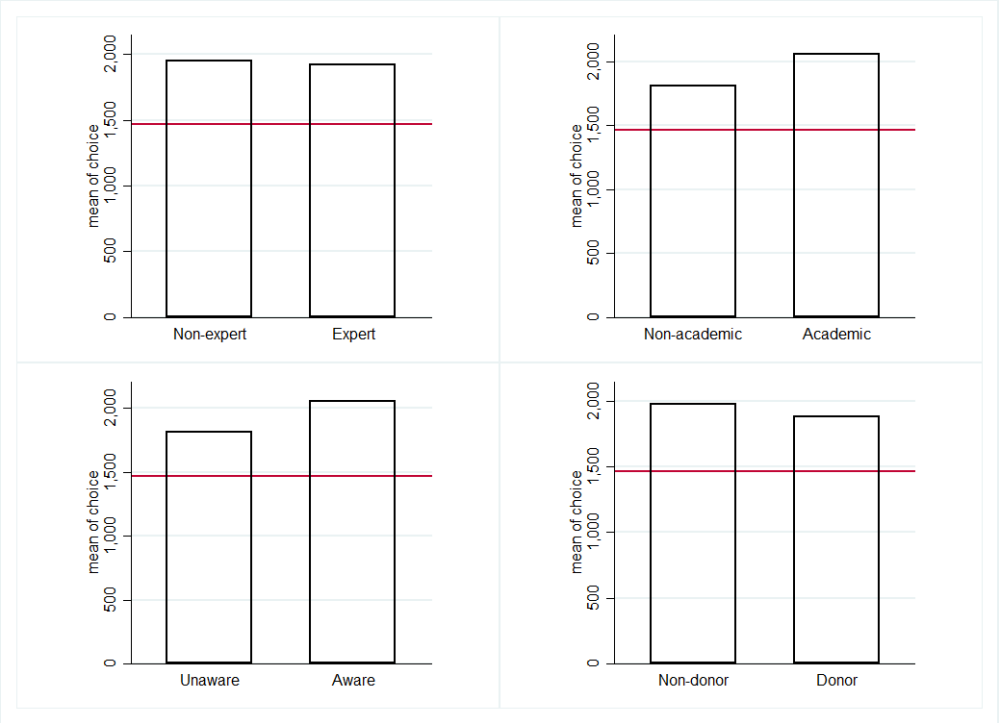


Figure 32: Guessed mean donations in the choice treatment by groups

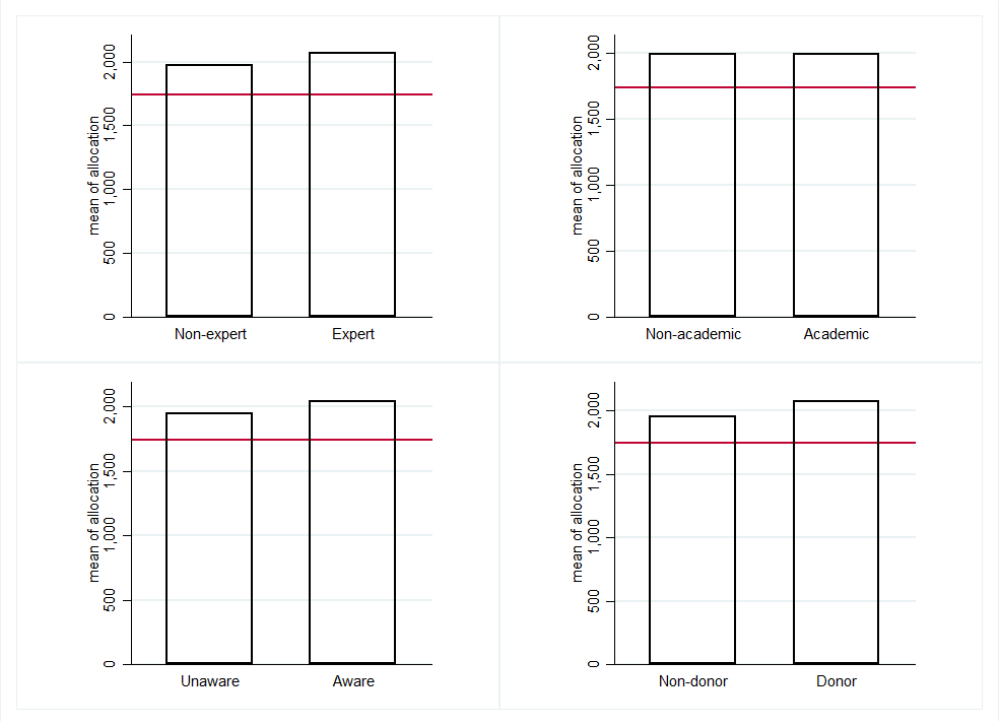


Figure 33: Guessed mean donations in the allocation treatment by groups

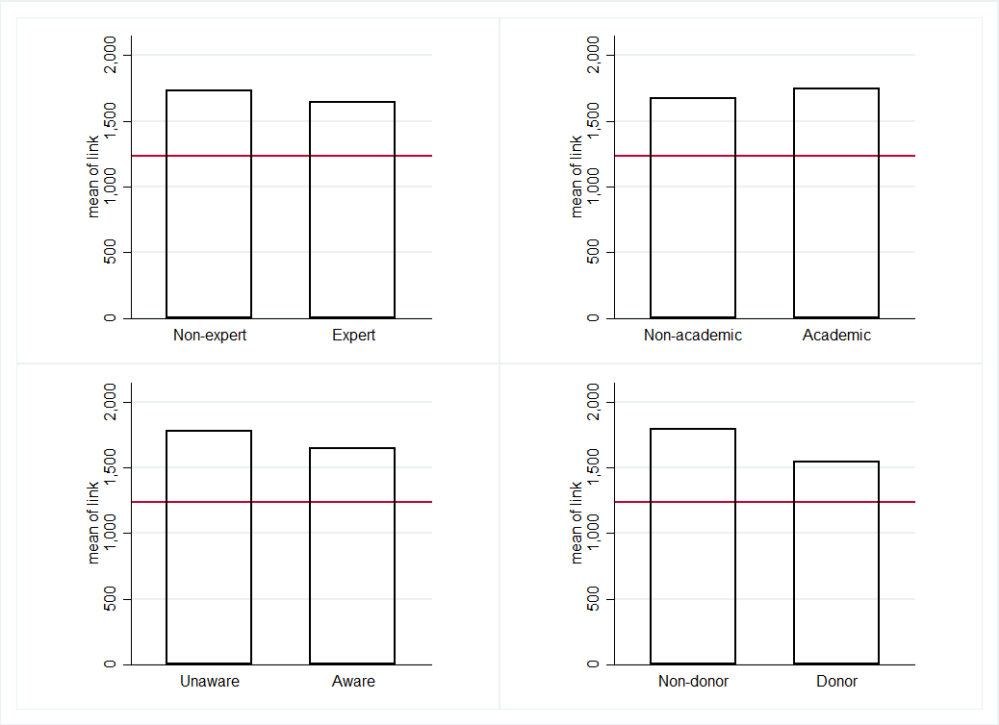


Figure 34: Guessed mean donations in the link treatment by groups

A.6 Email invitation to donation

Below is displayed the email invitation that HR sent to the employees of Deloitte in Albania (and its English translation).

E-mail Subject: **Deloitte në mbështetje të fëmijëve me Sindromën Down!**

Përshëndetje,

Fondacioni Down Syndrome Albania (DSA) dhe Deloitte, po bashkëpunojnë për të rritur ndërgjegjësimin mbi Sindromën Down si dhe për të mbledhur kontribute për 45 fëmijët me aftësi të kufizuara intelektuale të Qendrës së Shërbimeve "PRO PAK" të DSA. Ju lutem, kushtojini **2-3 minuta** kohë plotësimit të anketës së mëposhtme për të mbështetur këtë kauzë. Ne ju sigurojmë se të dhënat tuaja, përfshirë dhurimin do të mbeten tërësisht anonime. Çdo dhurim drejt Fondacionit Down Syndrome Albania është vullnetar!

https://support.dsalbania.org/donate/deloitte/2CHI

Ky link do të qëndrojë i hapur për një javë duke filluar nga dita e sotme. Për çdo pyetje, na kontaktoni tek info@dsalbania.org.

Ju falenderojmë për mbështetjen tuaj!

Figure 35: Email invitation (Albanian)

E-mail Subject: **Deloitte supporting children with Down Syndrome!**

Hello,

The Down Syndrome Albania Foundation (DSA) and Deloitte are collaborating to raise awareness and raise funds to support the 45 children with limited intellectual abilities at the "PRO-PAK" Service Center of DSA. Please take **2-3 minutes** to fill in the survey in the link below to support our cause. We assure you that your data, including your donation will remain anonymous. Every donation to the Down Syndrome Albania Foundation is voluntary!

https://support.dsalbania.org/donate/deloitte/2CHI

This link will remain active for about one week starting from today. Feel free to contact us at info@dsalbania.org for any question.

Thank you for your support!

Figure 36: Email invitation (English)

B Let Me Choose What I'm Best at: A Natural Field Experiment with Volunteers

B.1 Main Tables

Here, I present the main treatment effects in tabular and regression form, using various specifications.

Table 18: The effect of choosing tasks on effort and effort quality

	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	Nr of messages	Nr of messages	quality (%)	quality (%)	quality (grade)	quality (grade)
Choice	1.484***	1.324***	0.067***	0.053***	0.497***	0.391***
	(0.393)	(0.399)	(0.023)	(0.019)	(0.169)	(0.135)
% change	36.588	32.643	12.249	9.69	8.414	6.619
Control mean	4.056	4.056	0.547	0.547	5.907	5.907
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Task fixed effects	No	Yes	No	Yes	No	Yes
Cohort fixed effects	No	Yes	No	Yes	No	Yes
Controls	No	Yes	No	Yes	No	Yes
Observations	2,454	2,454	2,264	2,264	2,264	2,264
R-squared	0.084	0.168	0.038	0.152	0.031	0.201

Note: Columns (1)-(2) report the estimated choice treatment effect on the relevant number of messages, with and without controls and fixed effects. Columns (3)-(4) report the estimated choice treatment effect on the share of good messages, with and without controls and fixed effects. Columns (5)-(6) report the estimated choice treatment effect on the overall graded quality, with and without controls and fixed effects. The numbers below each treatment effect represent the percentage change caused by the treatment effect. In each regression, standard errors are clustered at the classroom level. *** p<0.01, ** p<0.05, * p<0.1

Table 19: The effect of choosing tasks on the willingness to volunteer in the future

	(1)	(2)	(3)	(4)
Treatment	Future (Y/N)	Future (Y/N)	Email (Y/N)	Email (Y/N)
Choice	0.075**	0.064**	0.041	0.026
	(0.032)	(0.030)	(0.032)	(0.031)
Control mean	0.540	0.540	0.359	0.359
Strata fixed effects	Yes	Yes	Yes	Yes
Cohort fixed effects	No	Yes	No	Yes
Controls	No	Yes	No	Yes
Observations	2,463	2,463	2,463	2,463
R-squared	0.010	0.060	0.022	0.100

Note: Columns (1)-(2) report the choice treatment effects on the willingness to volunteer in the future, with and without fixed effects and control variables. Columns (3)-(4) report the estimated choice treatment effects on willingness to share the email address for future calls. In all regressions, standard errors are clustered at the classroom level. *** p<0.01, ** p<0.05, * p<0.1.

Table 20: The effect of the monetary reward on effort and effort quality

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	Nr. Mess	Nr. Mess	Share	Share	Grade	Grade	Other	Other
Money	1.035**	0.929**	0.015	0.016	0.193	0.168	-0.232	-0.274
	(0.447)	(0.402)	(0.028)	(0.024)	(0.208)	(0.173)	(0.255)	(0.257)
% change	25.518	22.904	2.742	2.295	3.267	2.844	-6.959	-10.168
Control mean	4.056	4.056	0.547	0.547	5.907	5.907	3.334	3.334
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Task fixed effects	No	Yes	No	Yes	No	Yes	No	No
Cohort fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	1,887	1,887	1,617	1,617	1,617	1,617	1,887	1,887
R-squared	0.048	0.142	0.017	0.123	0.012	0.176	0.009	0.063

Note: Columns (1)-(2) report the estimated money treatment effect on the relevant number of messages (Nr. Mess), with and without controls and fixed effects. Columns (3)-(4) report the estimated money treatment effect on the share of good messages (Share), with and without controls and fixed effects. Columns (5)-(6) report the estimated money treatment effect on the overall graded quality (Grade), with and without controls and fixed effects. Columns (7)-(8) report the estimated money treatment effect on the relevant messages in the other task (Other). The numbers below each treatment effect represent the percentage change caused by the treatment effect. In each regression, standard errors are clustered at the classroom level. *** p<0.01, ** p<0.05, * p<0.1.

Table 21: Comparison of effort and effort quality in the choice and money treatments

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	Nr. Mess	Nr. Mess	Share	Share	Grade	Grade	Other	Other
Choice - Money	0.434	0.341	0.067**	0.051**	0.360**	0.268*	0.490**	0.515**
	(0.446)	(0.406)	(0.027)	(0.020)	(0.178)	(0.136)	(0.243)	(0.241)
% change	8.340	6.552	12.249	9.324	5.922	4.409	15.776	17.579
Money (mean)	5.204	5.204	0.547	0.547	6.079	6.079	3.106	3.106
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Task fixed effects	No	Yes	No	Yes	No	Yes	No	No
Cohort fixed effects	No	Yes	No	Yes	No	Yes	No	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	2,035	2,035	1,931	1,931	1,931	1,931	2,044	2,044
R-squared	0.021	0.122	0.025	0.172	0.015	0.219	0.013	0.065

Note: Columns (1)-(2) report the difference in the relevant number of messages (Nr. Mess), with and without controls and fixed effects between the choice and money treatments. Columns (3)-(4) report the difference in the share of good messages (Share), with and without controls and fixed effects between the choice and money treatments. Columns (5)-(6) report the difference in the the overall graded quality (Grade), with and without controls and fixed effects between the choice and money treatments. Columns (7)-(8) report the difference in the number of relevant messages in the other task (Other), with and without controls and fixed effects between the choice and money treatments. In each regression, standard errors are clustered at the classroom level. *** p<0.01, ** p<0.05, * p<0.1.

Table 22: Difference in differences estimates of effort and quality among choice and money treatments relative to the control group

	(1)	(2)	(3)	(4)
Treatment	Nr. of messages	quality (%)	quality (grade)	Nr. of messages (other task)
DiD(Choice-Money)	0.313	0.046**	0.236	0.516**
	(0.407)	(0.021)	(0.145)	(0.238)
Strata fixed effects	Yes	Yes	Yes	Yes
Task fixed effects	Yes	Yes	Yes	Yes
Cohort fixed effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	2,035	1,931	1,931	2,044

Note: Column (1-4) display the difference in differences estimates in the number of relevant messages, quality measures, and effort in the other task between choice and money treatments relative to the control group. In each regression, standard errors are clustered at the classroom level. *** p<0.01, ** p<0.05, * p<0.1.

Table 23: The effect of ability-matching on effort and effort quality

	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	Nr of messages	Nr of messages	quality (%)	quality (%)	quality (grade)	quality (grade)
Ability	0.913**	0.874**	0.064***	0.069***	0.438***	0.456***
	(0.436)	(0.423)	(0.021)	(0.019)	(0.160)	(0.142)
% change	22.509	21.548	11.7	12.614	7.414	7.720
Control mean	4.056	4.056	0.547	0.547	5.907	5.907
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Task fixed effects	No	Yes	No	Yes	No	Yes
Cohort fixed effects	No	Yes	No	Yes	No	Yes
Controls	No	Yes	No	Yes	No	Yes
Observations	2,346	2,346	2,095	2,095	2,095	2,095
R-squared	0.071	0.162	0.037	0.111	0.036	0.167

Note: Columns (1)-(2) report the estimated ability treatment effect on the relevant number of messages, with and without controls and fixed effects. Columns (3)-(4) report the estimated ability treatment effect on the share of good messages, with and without controls and fixed effects. Columns (5)-(6) report the estimated ability treatment effect on the overall graded quality, with and without controls and fixed effects. The numbers below each treatment effect represent the percentage change caused by the treatment effect. In each regression, standard errors are clustered at the classroom level. *** p<0.01, ** p<0.05, * p<0.1.

Table 24: Comparison of matched and exactly matched abilities to tasks between ability and control groups

	(1)	(2)	(3)	(4)	(5)
Comparison groups		Ability-Control		Control	
Sub-sample	Full sample	Matched	Exactly matched	Matched	Exactly matched
Outcome	Nr. of messages	Nr. of messages	Nr. of messages	Nr. of messages	Nr. of messages
Panel A: Ability-Control					
Ability	0.874**	0.242	0.078		
	(0.423)	(0.428)	(0.450)		
Panel B: Control group					
Ability-matched				0.798***	
				(0.189)	
Exact ability-matched					0.891***
					(0.237)
Baseline mean	4.056	4.650	4.803	3.821	3.928
Strata fixed effects	Yes	Yes	Yes	Yes	Yes
Task fixed effects	Yes	Yes	Yes	Yes	Yes
Cohort fixed effects	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Observations	2,346	1,154	735	1,153	1,153
R-squared	0.162	0.170	0.166	0.162	0.160

Note: Columns (1)-(3) in Panel (A) report the estimated ability treatment effect on the relevant number of messages in the full sample, ability-matched and exact ability-matched sub-samples. Columns (4)-(5) in Panel (B) report the effect of ability-matched and exact ability-matched indicators on the number of relevant messages in the control group only. In each regression, standard errors are clustered at the classroom level. *** p<0.01, ** p<0.05, * p<0.1.

Table 25: Task choice regressions

	(1)	(2)	(3)
Task choice	Bullying	Disability	Depression
Panel A: correlation between perceived ability and task choice			
Ability (bullying)	0.056***	-0.029***	-0.027***
	(0.010)	(0.009)	(0.008)
Ability (disability)	-0.018**	0.046***	-0.028***
	(0.008)	(0.006)	(0.008)
Ability (depression)	-0.025**	-0.012*	0.037***
	(0.009)	(0.007)	(0.007)
Panel B: Correlation between perceived awareness and task choice			
Aware (bullying)	0.013	-0.005	-0.008
	(0.008)	(0.008)	(0.007)
Aware (disability)	-0.009	0.003	0.006
	(0.009)	(0.007)	(0.009)
Aware (depression)	-0.005	-0.001	0.006
	(0.008)	(0.005)	(0.008)
Panel C: correlation between knowledge and task choice			
Knowledge (bullying)	0.001	-0.001	-0.000
	(0.001)	(0.001)	(0.001)
Knowledge (disability)	-0.000	0.000	-0.000
	(0.001)	(0.001)	(0.001)
Knowledge (depression)	-0.003**	-0.000	0.003**
	(0.001)	(0.001)	(0.001)
Panel D: correlation between perceived task easiness and task choice			
Easy (bullying)	-0.111**	-0.031	0.142**
	(0.046)	(0.060)	(0.066)
Easy (disability)	-0.104**	0.013	0.091
	(0.042)	(0.067)	(0.064)
Easy (depression)	-0.090**	0.001	0.089
	(0.042)	(0.058)	(0.059)
Mean	0.512	0.195	0.293
Observations	1,301	1,301	1,301
R-squared	0.047	0.056	0.051

Note: Each column measures the correlation of perceived ability (Panel A) perceived awareness (Panel B), knowledge (Panel C) and ease (Panel D) $i \in \{bullying, depression, disability\}$ with the choice of the task. In all panels, standard errors are clustered at the classroom level. *** p<0.01, ** p<0.05, * p<0.1.

Table 26: The effect of choice on effort in the other task and overall enjoyment level

	(1)	(2)	(3)	(4)
Treatment	Nr of messages (other)	Nr of messages (other)	Enjoyment	Enjoyment
Choice	0.266	0.235	0.144	0.136
	(0.214)	(0.217)	(0.220)	(0.213)
Control mean	3.334	3.334	7.400	7.400
Strata fixed effects	Yes	Yes	Yes	Yes
Cohort fixed effects	No	Yes	No	Yes
Controls	No	Yes	No	Yes
Observations	2,463	2,463	2,463	2,463
R-squared	0.006	0.063	0.014	0.039

Notes: Columns (1)-(2) report the choice treatment effects on the number of relevant messages in the other task, with and without fixed effects and control variables. Columns (3)-(4) report the estimated choice treatment effects on the task enjoyment level. In all regressions, standard errors are clustered at the classroom level. *** p<0.01, ** p<0.05, * p<0.1.

Table 27: Comparing effort in the choice and ability groups across different levels of ability-matching

	(1)	(2)	(3)	(4)
Treatment	Nr of messages	Nr of messages	Nr of messages	Nr of messages
	Full sample	Highest ability	Ability-matched	Exact Ability-matched
Choice-Ability	0.382	0.173	0.151	0.173
	(0.433)	(0.409)	(0.427)	(0.434)
Ability mean	5.192	5.192	5.191	5.063
Strata fixed effects	Yes	Yes	Yes	Yes
Task fixed effects	Yes	Yes	Yes	Yes
Cohort fixed effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	2,494	2,167	1,426	922
R-squared	0.126	0.130	0.137	0.119

Note: Column (1) reports the average treatment effect of choosing task relative to being matched with the task corresponding to the highest rated ability. Column (2-4) report the same estimate for different sub-samples corresponding to different degrees of ability - matching. In all regressions, standard errors are clustered at the classroom level. *** p<0.01, ** p<0.05, * p<0.1.

B.2 Robustness Checks

Here, I present a set of robustness checks for the paper's main results.

Table 28: The effect of choice on effort by task

Treatment	(1) Nr of messages (Bullying)	(2) Nr of messages (Depression)	(3) Nr of messages (Disability)
Choice	1.153***	1.451**	1.353***
	(0.433)	(0.596)	(0.337)
Control mean	4.566	4.205	3.449
Strata fixed effects	Yes	Yes	Yes
Cohort fixed effects	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Observations	1,032	766	656
R-squared	0.167	0.152	0.186

Note: Columns (1)-(3) report the choice treatment effect on the number of relevant messages by task, conditioning on fixed effects and control variables. In all regressions, standard errors are clustered at the classroom level. *** p<0.01, ** p<0.05, * p<0.1.

Table 29: Main results excluding "next-day" volunteers

Treatment	(1) Nr. of messages	(2) Nr. of messages	(3) Nr. of messages	(4) Nr. of messages
Choice	1.392***			
	(0.409)			
Money		1.066**		
		(0.423)		
Choice-Money			0.268	
			(0.423)	
Ability				0.921**
				(0.442)
Strata fixed effects	Yes	Yes	Yes	Yes
Task fixed effects	Yes	Yes	Yes	Yes
Cohort fixed effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	2,321	1,765	1,932	2,191
R-squared	0.173	0.149	0.121	0.166

Note: Column (1-4) report the treatment effects on effort exluding those who performed the tasks the next day. In all regressions, standard errors are clustered at the classroom level. *** p<0.01, ** p<0.05, * p<0.1.

Table 30: Main treatment effects on winsorized outcome at the 1% and 5% levels

Treatment	(1) Effort 1%	(2) Effort 5%	(3) Effort 1%	(4) Effort 5%	(5) Effort 1%	(6) Effort 5%	(7) Effort 1%	(8) Effort 5%
Choice	1.482***	1.441***						
	(0.391)	(0.368)						
Money			1.021**	0.843**				
			(0.443)	(0.406)				
Choice - Money					0.447	0.600		
					(0.442)	(0.387)		
Ability							0.908**	0.803**
							(0.433)	(0.386)
Control mean	4.053	3.997	4.053	3.997	5.187	4.933	4.053	3.997
Task fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Strata fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,454	2,454	1,887	1,887	2,035	2,035	2,346	2,346
R-squared	0.084	0.088	0.047	0.042	0.021	0.025	0.071	0.065

Note: Column (1-8) display the difference in mean number of relevant messages for different treatment comparisons and different winsorization levels of the number of relevant messages. In each regression, standard errors are clustered at the classroom level. *** p<0.01, ** p<0.05, * p<0.1.

Table 31: Contamination Bias (corrected estimates)

	Nr. of messages			Quality (%)			Quality(graded)		
Treat	ATE	1-at-a-time	common weights	ATE	1-at-a-time	common weights	ATE	1-at-a-time	common weights
Choice	1.600***	1.625***	1.597***	0.068***	0.065***	0.069***	0.514***	0.512***	0.528***
	(0.149)	(0.145)	(0.152)	(0.013)	(0.013)	(0.013)	(0.078)	(0.077)	(0.080)
Ability	0.947***	0.957***	0.931***	0.063***	0.064***	0.062***	0.430***	0.440***	0.424***
	(0.157)	(0.153)	(0.159)	(0.014)	(0.014)	(0.014)	(0.081)	(0.081)	(0.082)
Money	1.034***	1.041***	1.023***	0.007***	0.015***	0.009***	0.159***	0.194***	0.165***
	(0.205)	(0.200)	(0.204)	(0.016)	(0.016)	(0.016)	(0.100)	(0.099)	(0.099)
Observations	4,381	4,381	4,381	4,026	4,026	4,026	4,026	4,026	4,026

Note: The ATE columns report the unweighted ATEs of all three treatments on each of the three outcomes. Columns "one at a time" report the effects of each treatment estimated separately. Finally, columns "common weights" estimate the efficiently weighted ATEs. Note that standard errors cannot be clustered at the class level. The estimated ATEs are the important part of this table.

Table 32: Multiple hypotheses testing: Romano-Wolf p-values

	(1)	(2)	(3)	(4)
Outcome	Treatment	Model p-value	Resample p-value	Romano-wolf p-value
Nr. of messages	Choice	0.0003	0.0003	0.0007
Nr. of messages	Money	0.0237	0.0057	0.0250
Nr. of messages	Choice-Money	0.3342	0.2199	0.4505
Nr. of messages	Ability	0.0394	0.0090	0.0460
Quality (%)	Choice	0.0051	0.0007	0.0057
Quality (%)	Money	0.6040	0.5042	0.5042
Quality (%)	Choice-Money	0.0139	0.0027	0.0130
Quality (%)	Ability	0.0029	0.0003	0.0033
Quality (grade)	Choice	0.0042	0.0010	0.0043
Quality (grade)	Money	0.3564	0.2666	0.4505
Quality (grade)	Choice-Money	0.0470	0.0157	0.0467
Quality (grade)	Ability	0.0075	0.0017	0.0067
Strata fixed effects		Yes	Yes	Yes
Task fixed effects		Yes	Yes	Yes

Note: Columns (1) represents the treatment dummy. Column (2) reports the p-value of the treatment effect from the estimated model with task and strata fixed effects and clustered standard errors. Column (3) reports the p-value that is obtained with 3000 bootstrap replications. Column (4) reports the adjusted p-values following Romano and Wolf (2005), and accounting for strata fixed effects and clustered standard errors.

Table 33: Randomization-based inference (RBI) p-values

	(1)	(2)	(3)
Treatment	Nr. of messages	Quality (%)	Quality(Graded)
Choice	1.465	0.070	0.505
p-value	(0.000)	(0.003)	(0.003)
RBI p-value	[0.002]	[0.006]	[0.003]
Ability	0.974	0.062	0.452
p-value	(0.032)	(0.004)	(0.006)
RBI p-value	[0.027]	[0.009]	[0.005]
Money	1.068	0.012	0.187
p-value	(0.015)	(0.672)	(0.350)
RBI p-value	[0.032]	[0.648]	[0.296]
Choice-Money	0.396	0.058	0.317
p-value	(0.375)	(0.033)	(0.083)
RBI p-value	[0.430]	[0.030]	[0.099]
Task fixed effects	Yes	Yes	Yes
Strata fixed effects	Yes	Yes	Yes
Observations	4,381	4,026	4,026
R-squared	0.056	0.029	0.023

Note: Columns (1) - (3) present the jointly estimated mean differences in the number of relevant messages, share of good messages and overall graded quality, conditioning on strata and task fixed effects and clustering errors at the classroom level. Below each treatment effect, I report the model p-value in parentheses, and the RBI p-values in square brackets.

B.3 Experimental details

Here I display pictures from the field and translated versions of the experimental sheets.

Figure 37: Details of the experiment

Notes: Students signed a consent to allow the U-Report teachers take pictures of the event.

Below are displayed the translated experimental sheets. Note that the control and ability treatments had identical texts up to the chosen task. Figure 38 displays an example of the experimental sheet where the student was asked to write awareness-raising messages about bullying. If the student was assigned to the ability-matching treatment, then bullying is the cause she felt more able. If she were assigned to the control group, bullying would be assigned randomly.

Code

I

Hello,

The Observatory for the Rights of Children and Youth is working to raise awareness about three social causes: depression, people with disabilities, and bullying. According to psychologists and sociologists, it is essential to raise awareness about these social causes, and it requires the active involvement of youth that has or has not been affected by them to enhance their social and academic lives toward a greater future. You can voluntarily help by writing as many awareness messages against **bullying**. The best motivating messages will be posted in the **U-Report** platform, managed by **UNICEF**, where you can register through scanning the QR-code on the back page, thus becoming a volunteer in a large youth community. U-Report aims to empower youth to share their thoughts on issues that are important to them, to inform them and to shrink the distance between them and (non)government institutions. Moreover, the messages that will not be selected for this campaign, will be filtered and used in similar campaigns in the future. Please, write your messages in **capital letters** in order to increase readability as well as avoid short messages with 2-3 words, e.g., "*Say No to bullying*". Thank you for your contribution in creating a better community for children, youth and for all. **Your voice matters!**

1._____

2._____

3._____

4._____

5._____

6._____

7._____

8._____

9._____

10._____

11._____

12._____

13._____

14._____

15._____

16._____

17._____

18._____

19._____

20._____

Figure 38: Control (Ability) treatment

Code

II

Hello,

The Observatory for the Rights of Children and Youth is working to raise awareness about three social causes: depression, people with disabilities, and bullying. According to psychologists and sociologists, it is essential to raise awareness about these social causes, and it requires the active involvement of youth that has or has not been affected by them to enhance their social and academic lives toward a greater future. You can voluntarily help by writing as many awareness messages against (**choose one between: ☐depression ☐social exclusion of people with disabilities ☐bullying**). The best motivating messages will be posted in the **U-Report** platform, managed by **UNICEF**, where you can register through scanning the QR-code on the back page, thus becoming a volunteer in a large youth community. U-Report aims to empower youth to share their thoughts on issues that are important to them, to inform them and to shrink the distance between them and (non)government institutions. Moreover, the messages that will not be selected for this campaign, will be filtered and used in similar campaigns in the future. Please, write your messages in **capital letters** in order to increase readability as well as avoid short messages with 2-3 words, e.g., "*Say No to bullying*". Thank you for your contribution in creating a better community for children, youth and for all. **Your voice matters!**

1. _____
2. _____
3. _____
4. _____
5. _____
6. _____
7. _____
8. _____
9. _____
10. _____
11. _____
12. _____
13. _____
14. _____
15. _____
16. _____
17. _____
18. _____
19. _____
20. _____

Figure 39: Choice treatment

Code

IV

Hello,

The Observatory for the Rights of Children and Youth is working to raise awareness about three social causes: depression, people with disabilities, and bullying. According to psychologists and sociologists, it is essential to raise awareness about these social causes, and it requires the active involvement of youth that has or has not been affected by them to enhance their social and academic lives toward a greater future. You can voluntarily help by writing as many awareness messages against **bullying**. The best motivating messages will be posted in the **U-Report** platform, managed by **UNICEF**, where you can register through scanning the QR-code on the back page, thus becoming a volunteer in a large youth community. U-Report aims to empower youth to share their thoughts on issues that are important to them, to inform them and to shrink the distance between them and (non)government institutions. Moreover, the messages that will not be selected for this campaign, will be filtered and used in similar campaigns in the future. **The student in your class with the most messages selected by the Observatory will receive a reward of 20 Euros**. If you are the winner, you will receive a reward through the office of your school's principal. Would you like to be considered for the reward if you are the winner?□Yes □No. Please, write your messages in **capital letters** in order to increase readability as well as avoid short messages with 2-3 words, e.g., *"Say No to bullying"*. Thank you for your contribution in creating a better community for children, youth and for all. **Your voice matters!**

1. _____
2. _____
3. _____
4. _____
5. _____
6. _____
7. _____
8. _____
9. _____
10. _____
11. _____
12. _____
13. _____
14. _____
15. _____
16. _____
17. _____
18. _____
19. _____
20. _____

Figure 40: Money treatment

How much did you enjoy writing awareness messages in a scale from 1 to 10? *(0- did not enjoy, 10- fully enjoyed)*

☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10

Would you consider becoming a volunteer in the future? ☐ Po ☐ Jo

If yes, please write down your e-mail address for future notifications:

E-mail:_____

Lastly, take a few minutes to help raise awareness about environmental pollution by writing awareness messages against pollution. This is important since the level of urban waste has recently increased drastically.

1._____

2._____

3._____

4._____

5._____

6._____

7._____

8._____

9._____

10._____

Thank you for your volunteering work!

Scan code-WhatsApp

Figure 41: Other outcomes

B.4 Randomization

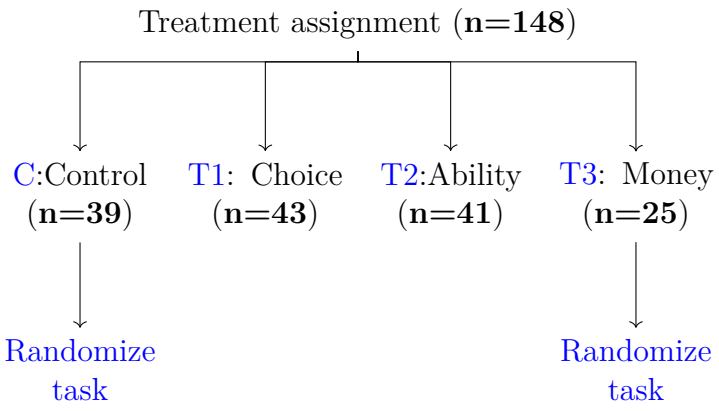


Figure 42: Randomization Diagram

Table 34: Randomization balance tests (choice & control groups)

Characteristics	Control Mean	Choice Mean	Difference	Norm. Diff.	p-value	N
Female	0.561	0.587	-0.026	0.052	0.155	2463
Male	0.439	0.413	0.026	0.052	0.155	2463
Age	16.547	16.744	-0.197	0.204	0.341	2463
Prev. volunt. exp.	0.465	0.488	-0.023	0.046	0.389	2463
Score bullying	53.393	55.234	-1.841	0.106	0.344	2463
Score disability	52.228	53.573	-1.345	0.075	0.681	2463
Score depression	52.278	54.285	-2.007	0.114	0.271	2463
Aware bullying	8.423	8.554	-0.131	0.058	0.374	2463
Aware disability	8.944	8.981	-0.037	0.019	0.684	2463
Aware depression	7.771	7.888	-0.117	0.042	0.522	2463
Ability bullying	8.166	8.159	0.028	0.006	0.655	2463
Ability disability	7.609	7.712	-0.103	0.043	0.233	2463
Ability depression	7.751	7.791	-0.040	0.017	0.602	2463
Easy (bullying)	0.389	0.420	-0.031	0.063	0.246	2463
Easy (disability)	0.382	0.369	0.013	0.028	0.260	2463
Easy (depression)	0.232	0.213	0.019	0.047	0.687	2463
Year 1	0.368	0.300	0.068	0.146	0.445	2463
Year 2	0.323	0.311	0.012	0.026	0.810	2463
Year 3	0.308	0.388	-0.080	0.170	0.319	2463
Language	7.674	7.998	-0.324	0.202	0.087	2463
Literature	7.647	7.961	-0.314	0.197	0.092	2463
Maths	6.888	7.229	-0.341	0.173	0.287	2463
Civic Education	8.491	8.721	-0.230	0.150	0.260	2463
Joint orthog. p-value					0.746	

Note: The labels of the baseline characteristics which were collected through the survey and administrative data from the schools are displayed in the first column. The next three column present the means and difference in means between each treatment arm. The fourth column displays the normalized difference following Imbens and Rubin (2015). The fifth column presents the p-values from the regressions of each baseline characteristic on the treatment dummy and strata fixed effects, clustering standard errors at the classroom level. The last column displays the number of observations. *** p<0.01, ** p<0.05, * p<0.1.

Table 35: Randomization balance tests (ability & control groups)

Characteristics	Control Mean	Ability Mean	Difference	Norm. Diff	p-value	N
Female	0.561	0.574	-0.013	0.026	0.549	2346
Male	0.439	0.426	0.013	0.026	0.549	2346
Age	16.547	16.523	0.024	0.025	0.872	2346
Prev. volunt. exp.	0.465	0.432	0.033	0.067	0.313	2346
Score bullying	53.393	52.086	1.307	0.074	0.239	2346
Score disability	52.228	50.633	1.595	0.088	0.194	2346
Score depression	52.278	51.218	1.060	0.058	0.304	2346
Aware bullying	8.423	8.282	0.141	0.058	0.167	2346
Aware disability	8.944	8.876	-0.068	0.034	0.438	2346
Aware depression	7.771	7.616	0.155	0.054	0.265	2346
Ability bullying	8.166	8.033	0.133	0.057	0.288	2346
Ability disability	7.609	7.657	-0.048	0.019	0.488	2346
Ability depression	7.751	7.726	0.025	0.010	0.911	2346
Easy (bullying)	0.389	0.390	-0.001	0.002	0.914	2346
Easy (disability)	0.382	0.376	0.006	0.012	0.518	2346
Easy (depression)	0.232	0.238	-0.006	0.013	0.535	2346
Year 1	0.369	0.413	-0.044	0.091	0.794	2346
Year 2	0.323	0.281	0.042	0.093	0.754	2346
Year 3	0.308	0.306	0.002	0.004	0.975	2346
Language	7.674	7.864	-0.190	0.120	0.220	2346
Literature	7.647	7.719	-0.072	0.045	0.618	2346
Maths	6.888	6.969	-0.056	0.081	0.895	2346
Civic Education	8.491	8.544	-0.053	0.034	0.881	2346
Joint orthog. p-value					0.589	

Note: The labels of the baseline characteristics which were collected through the survey and administrative data from the schools are displayed in the first column. The next three column present the means and difference in means between each treatment arm. The fourth column displays the normalized difference following Imbens and Rubin (2015). The fifth column presents the p-values from the regressions of each baseline characteristic on the treatment dummy and strata fixed effects, clustering standard errors at the classroom level. The last column displays the number of observations. *** p<0.01, ** p<0.05, * p<0.1.

Table 36: Randomization balance tests (money & control groups)

Characteristics	Control Mean	Money Mean	Difference	Norm. Diff	p-value	N
Female	0.561	0.590	-0.028	0.058	0.228	1887
Male	0.439	0.410	0.028	0.058	0.228	1887
Age	16.547	16.428	0.119	0.114	0.676	1887
Prev. volunt. exp.	0.465	0.436	0.029	0.058	0.705	1887
Score bullying	53.393	54.528	-1.135	0.065	0.490	1887
Score disability	52.228	53.386	-1.058	0.064	0.487	1887
Score depression	52.278	53.692	-1.414	0.080	0.277	1887
Aware bullying	8.423	8.353	0.070	0.029	0.615	1887
Aware disability	8.944	8.811	0.133	0.065	0.341	1887
Aware depression	7.771	7.751	0.020	0.007	0.966	1887
Ability bullying	8.166	8.167	-0.001	0.000	0.731	1887
Ability disability	7.609	7.753	-0.144	0.061	0.196	1887
Ability depression	7.751	7.868	-0.117	0.048	0.233	1887
Easy (bullying)	0.389	0.338	0.051	0.107	0.125	1887
Easy (disability)	0.382	0.429	-0.047	0.095	0.235	1887
Easy (depression)	0.232	0.245	-0.013	0.030	0.452	1887
Year 1	0.369	0.383	-0.014	0.029	0.972	1887
Year 2	0.323	0.329	-0.006	0.013	0.843	1887
Year 3	0.308	0.288	0.020	0.045	0.817	1887
Language	7.674	7.816	-0.142	0.089	0.317	1887
Literature	7.647	7.734	-0.087	0.053	0.470	1887
Maths	6.888	6.888	0.000	0.000	0.960	1887
Civic Education	8.491	8.450	0.041	0.026	0.767	1887
Joint orthog. p-value					0.470	

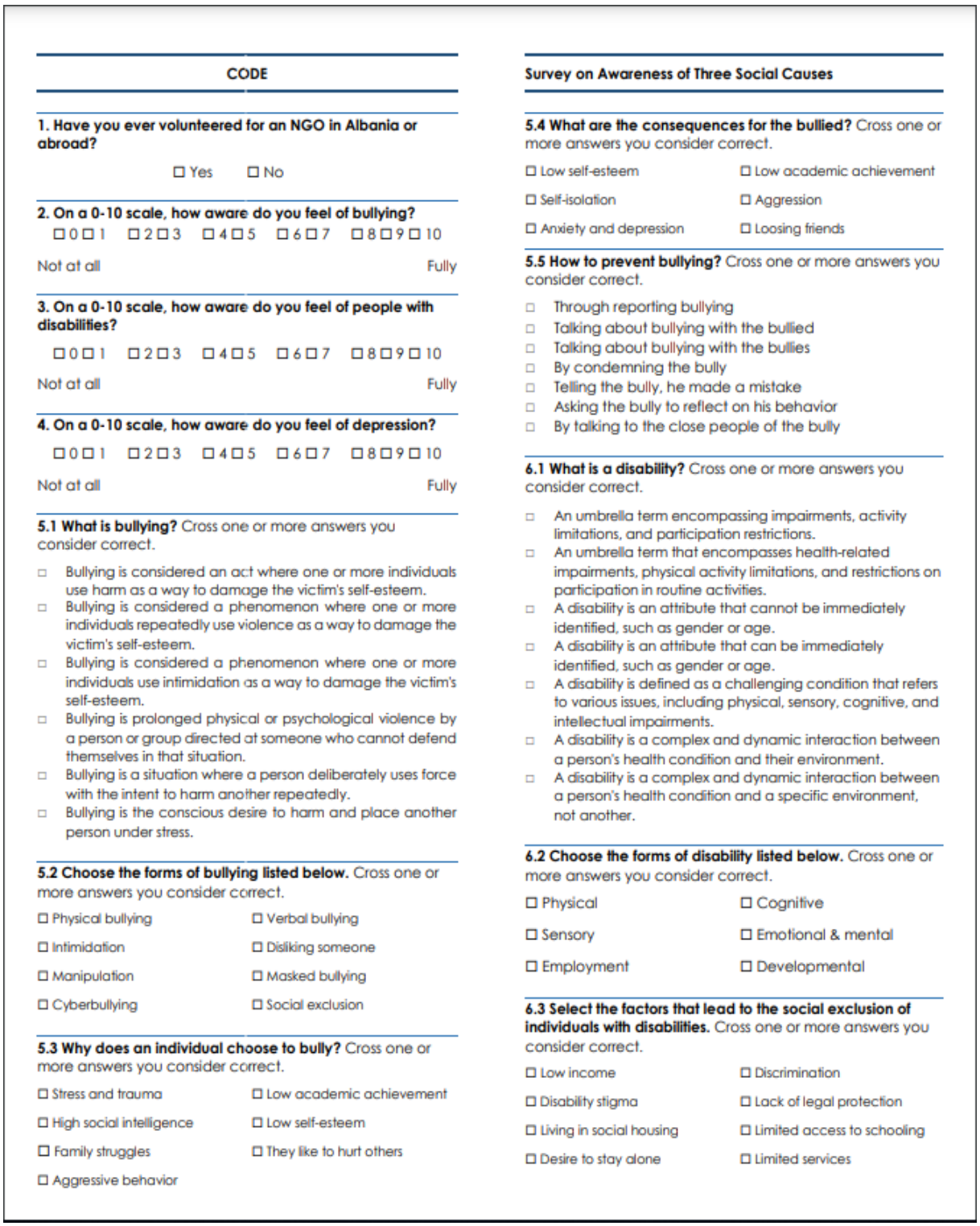
Note: The labels of the baseline characteristics which were collected through the survey and administrative data from the schools are displayed in the first column. The next three column present the means and difference in means between each treatment arm. The fourth column displays the normalized difference following Imbens and Rubin (2015). The fifth column presents the p-values from the regressions of each baseline characteristic on the treatment dummy and strata fixed effects, clustering standard errors at the classroom level. The last column displays the number of observations. *** p<0.01, ** p<0.05, * p<0.1

Table 37: Randomization balance tests (choice & money groups)

Characteristics	Money Mean	Choice Mean	Difference	Norm. Diff	p-value	N
Female	0.590	0.587	0.003	0.006	0.866	2044
Male	0.410	0.413	-0.003	0.006	0.866	2044
Age	16.428	16.744	-0.316	0.297	0.224	2044
Prev. volunt. exp.	0.436	0.488	-0.052	0.104	0.149	2044
Score bullying	54.528	55.234	-0.706	0.043	0.800	2044
Score disability	53.386	53.573	-0.187	0.011	0.981	2044
Score depression	53.692	54.285	-0.593	0.035	0.630	2044
Aware bullying	8.353	8.554	-0.201	0.087	0.150	2044
Aware disability	8.811	8.982	-0.171	0.085	0.199	2044
Aware depression	7.751	7.888	-0.137	0.049	0.592	2044
Ability bullying	8.167	8.159	0.008	0.004	0.806	2044
Ability disability	7.753	7.712	0.041	0.017	0.666	2044
Ability depression	7.868	7.791	0.077	0.033	0.676	2044
Easy (bullying)	0.338	0.421	-0.083***	0.171	0.004	2044
Easy (disability)	0.429	0.369	0.060**	0.124	0.032	2044
Easy (depression)	0.245	0.213	0.032	0.077	0.309	2044
Year 1	0.383	0.300	0.083	0.175	0.442	2044
Year 2	0.330	0.311	0.018	0.039	0.791	2044
Year 3	0.287	0.388	-0.101	0.215	0.299	2044
Language	7.816	7.998	-0.182	0.117	0.202	2044
Literature	7.734	7.961	-0.227	0.143	0.188	2044
Maths	6.888	7.229	-0.341	0.178	0.090	2044
Civic Education	8.450	8.721	-0.271	0.186	0.074	2044
Joint orthog. p-value					0.300	

Note: The labels of the baseline characteristics which were collected through the survey and administrative data from the schools are displayed in the first column. The next three column present the means and difference in means between each treatment arm. The fourth column displays the normalized difference following Imbens and Rubin (2015). The fifth column presents the p-values from the regressions of each baseline characteristic on the treatment dummy and strata fixed effects, clustering standard errors at the classroom level. The last column displays the number of observations. *** p<0.01, ** p<0.05, * p<0.1.

B.5 Supplementary Survey



CODE

1. Have you ever volunteered for an NGO in Albania or abroad?

☐ Yes ☐ No

2. On a 0-10 scale, how aware do you feel of bullying?

☐0 ☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8 ☐9 ☐10

Not at all Fully

3. On a 0-10 scale, how aware do you feel of people with disabilities?

☐0 ☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8 ☐9 ☐10

Not at all Fully

4. On a 0-10 scale, how aware do you feel of depression?

☐0 ☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8 ☐9 ☐10

Not at all Fully

5.1 What is bullying? Cross one or more answers you consider correct.

- ☐ Bullying is considered an act where one or more individuals use harm as a way to damage the victim's self-esteem.
- ☐ Bullying is considered a phenomenon where one or more individuals repeatedly use violence as a way to damage the victim's self-esteem.
- ☐ Bullying is considered a phenomenon where one or more individuals use intimidation as a way to damage the victim's self-esteem.
- ☐ Bullying is prolonged physical or psychological violence by a person or group directed at someone who cannot defend themselves in that situation.
- ☐ Bullying is a situation where a person deliberately uses force with the intent to harm another repeatedly.
- ☐ Bullying is the conscious desire to harm and place another person under stress.

5.2 Choose the forms of bullying listed below. Cross one or more answers you consider correct.

☐ Physical bullying ☐ Verbal bullying
☐ Intimidation ☐ Disliking someone
☐ Manipulation ☐ Masked bullying
☐ Cyberbullying ☐ Social exclusion

5.3 Why does an individual choose to bully? Cross one or more answers you consider correct.

☐ Stress and trauma ☐ Low academic achievement
☐ High social intelligence ☐ Low self-esteem
☐ Family struggles ☐ They like to hurt others
☐ Aggressive behavior

Survey on Awareness of Three Social Causes

5.4 What are the consequences for the bullied? Cross one or more answers you consider correct.

☐ Low self-esteem ☐ Low academic achievement
☐ Self-isolation ☐ Aggression
☐ Anxiety and depression ☐ Loosing friends

5.5 How to prevent bullying? Cross one or more answers you consider correct.

- ☐ Through reporting bullying
- ☐ Talking about bullying with the bullied
- ☐ Talking about bullying with the bullies
- ☐ By condemning the bully
- ☐ Telling the bully, he made a mistake
- ☐ Asking the bully to reflect on his behavior
- ☐ By talking to the close people of the bully

6.1 What is a disability? Cross one or more answers you consider correct.

- ☐ An umbrella term encompassing impairments, activity limitations, and participation restrictions.
- ☐ An umbrella term that encompasses health-related impairments, physical activity limitations, and restrictions on participation in routine activities.
- ☐ A disability is an attribute that cannot be immediately identified, such as gender or age.
- ☐ A disability is an attribute that can be immediately identified, such as gender or age.
- ☐ A disability is defined as a challenging condition that refers to various issues, including physical, sensory, cognitive, and intellectual impairments.
- ☐ A disability is a complex and dynamic interaction between a person's health condition and their environment.
- ☐ A disability is a complex and dynamic interaction between a person's health condition and a specific environment, not another.

6.2 Choose the forms of disability listed below. Cross one or more answers you consider correct.

☐ Physical ☐ Cognitive
☐ Sensory ☐ Emotional & mental
☐ Employment ☐ Developmental

6.3 Select the factors that lead to the social exclusion of individuals with disabilities. Cross one or more answers you consider correct.

☐ Low income ☐ Discrimination
☐ Disability stigma ☐ Lack of legal protection
☐ Living in social housing ☐ Limited access to schooling
☐ Desire to stay alone ☐ Limited services

Figure 43: Supplementary Survey 1/2

6.4 What are some misconceptions about disability? Cross one or more answers you consider correct.

- ☐ A disability defines what kind of individuals bear it.
- ☐ People with disabilities experience constant pain.
- ☐ People with disabilities are not dependent on the help of others.
- ☐ A disability is not a personal tragedy and should not be pitied.
- ☐ People with disabilities are special and should be treated differently.
- ☐ People with disabilities are dependent on the help of others.
- ☐ Individuals with disabilities cannot live a productive life.

6.5 How can you, we improve the social inclusion of people with disabilities? Cross one or more answers you consider correct.

- ☐ By contributing to their self-esteem by showing that you enjoy spending time with them.
- ☐ By avoiding pity as a reason to associate with them.
- ☐ By raising awareness of disabilities.
- ☐ By forming inclusive friendship groups with and without people with disabilities.
- ☐ By giving them space to respect their necessary boundaries.
- ☐ By showing that you are inspired by their strength and courage.
- ☐ By inviting them to social events outside of school/work.

7.1 What is depression? Cross one or more answers you consider correct.

- ☐ Depression is a mental condition of an individual characterized by negative changes, feelings of fear and guilt, as well as obstacles in learning.
- ☐ Depression is one of the most common types of mental health conditions and can develop as early as childhood.
- ☐ Depression is one of the most common types of mental health conditions and often develops alongside anxiety.
- ☐ Depression is a chronic mental health condition that affects an individual's emotional state, the way they think, and how they act.
- ☐ Depression is a chronic or short-term mental health condition that affects an individual's emotional state, the way they think, and how they act.

7.2 What causes depression? Cross one or more answers you consider correct.

- ☐ Excessive stress
- ☐ Inherited genes
- ☐ Family trauma
- ☐ Social isolation
- ☐ Feeling of failure
- ☐ Lack of affection
- ☐ Abuse and violence
- ☐ Competition
- ☐ Anorexia
- ☐ Self-depreciation

7.3 What are the symptoms of depression? Cross one or more answers you consider correct.

- ☐ Changes in appetite
- ☐ Constant anxiety
- ☐ Self-harm
- ☐ Loss of energy
- ☐ Social isolation
- ☐ Sleeping disorder
- ☐ Abuse and violence
- ☐ Aggression
- ☐ Feeling of worthlessness
- ☐ Bad grades at school

7.4 What are some misconceptions about depression? Cross one or more answers you consider correct.

- ☐ People with depression are not normal
- ☐ Depression is not a medical diagnosis
- ☐ Depression is contagious
- ☐ Depression is a medical diagnosis
- ☐ Talking about depression is stigmatized
- ☐ Depression is just excessive stress
- ☐ Depression can be cured with antidepressants

7.5 How to prevent depression? Cross one or more answers you consider correct.

- ☐ By avoiding negative thoughts
- ☐ By seeking medical help
- ☐ By reducing excessive food intake
- ☐ By controlling antidepressants intake
- ☐ By doing sports
- ☐ By talking about our problems to others
- ☐ By focusing more on studying for school
- ☐ Staying in isolation to reflect
- ☐ Eating healthy and eliminating harmful substances

8. If you were to contribute to raising awareness of the three social causes mentioned above, how capable do you feel of raising awareness in society about each one of them on a scale from 0 to 10? (0- not at all & 10-fully)

Depression

☐0 ☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8 ☐9 ☐10

Disabilities

☐0 ☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8 ☐9 ☐10

Bullying

☐0 ☐1 ☐2 ☐3 ☐4 ☐5 ☐6 ☐7 ☐8 ☐9 ☐10

9. How easy/difficult do you think it is to raise awareness of these causes? (1-easy, 2-doable, 3- difficult). Put the numbers in the boxes

☐ Depression ☐ Bullying

☐ Disabilities

Figure 44: Supplementary Survey 2/2

Table 38: Comparison of mean ability, awareness, and knowledge by task

Variable	Mean	Difference	p-value	Obs.
Ability (Bullying)	8.128			4390
Ability (Depression)	7.776			4390
		0.352***	0.000	4390
Ability (Bullying)	8.128			4390
Ability (Disability)	7.677			4390
		0.451***	0.000	4390
Ability (Depression)	7.776			4390
Ability (Disability)	7.677			4390
		0.098***	0.004	4390
Aware (Bullying)	8.412			4390
Aware (Depression)	7.760			4390
		0.652***	0.000	4390
Aware (Bullying)	8.412			4390
Aware (Disability)	8.914			4390
		-0.502***	0.000	4390
Aware (Depression)	7.760			4390
Aware (Disability)	8.914			4390
		-1.154***	0.000	4390
Knowledge (Bullying)	53.777			4390
Knowledge (Depression)	52.825			4390
		0.951***	0.000	4390
Knowledge (Bullying)	53.777			4390
Knowledge (Disability)	52.390			4390
		1.387***	0.000	4390
Knowledge (Depression)	52.825			4390
Knowledge (Disability)	52.390			4390
		0.435**	0.037	4390

Notes: The first column indicates the variables in each paired t test. The second and third columns display the mean and difference in means. The fourth column displays the p-value of the test and the last column indicates the number of observations.

C Unrealistic Educational Aspirations: Behavioral Insights From Parental Decisions in Albania

C.1 Other Channels: Why VET Enrollment is Low?

In this section, we present descriptive evidence on other potential factors that make parents reluctant to enroll their children in the VET track. These factors include perceived returns to different education tracks, the stigma or low reputation associated with the VET track, and inaccurate beliefs that the VET track denies access to college.

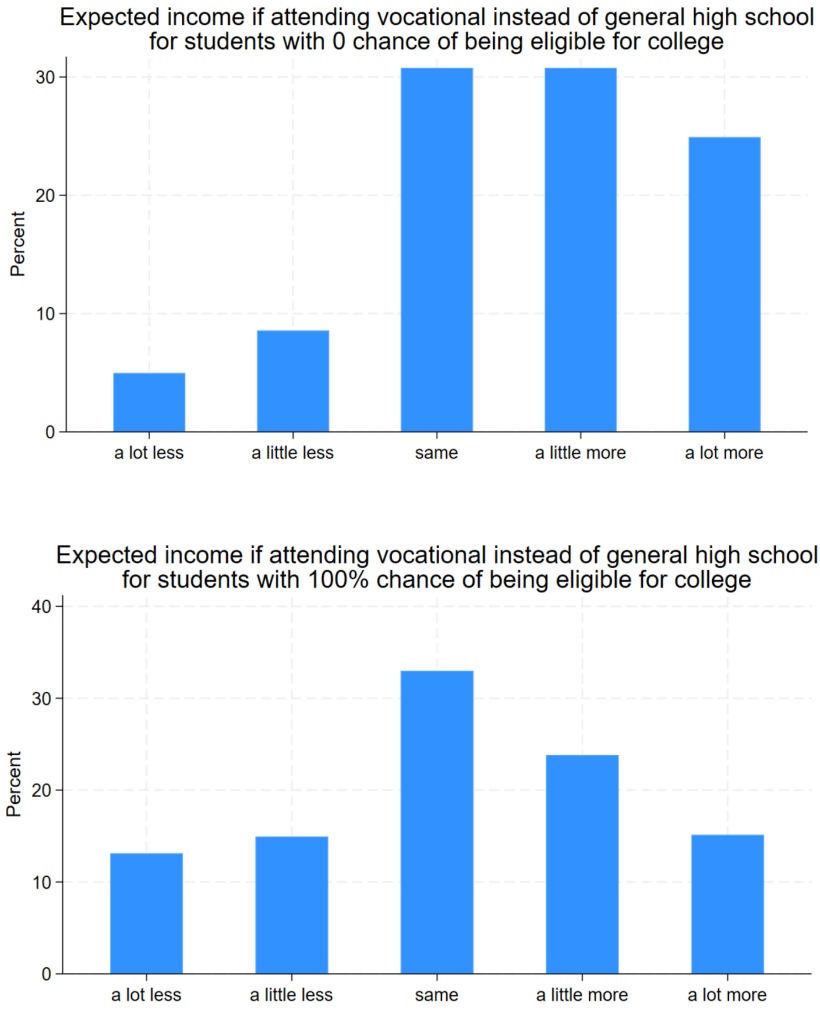


Figure 45: Perceived returns to different schooling levels (different formulation)

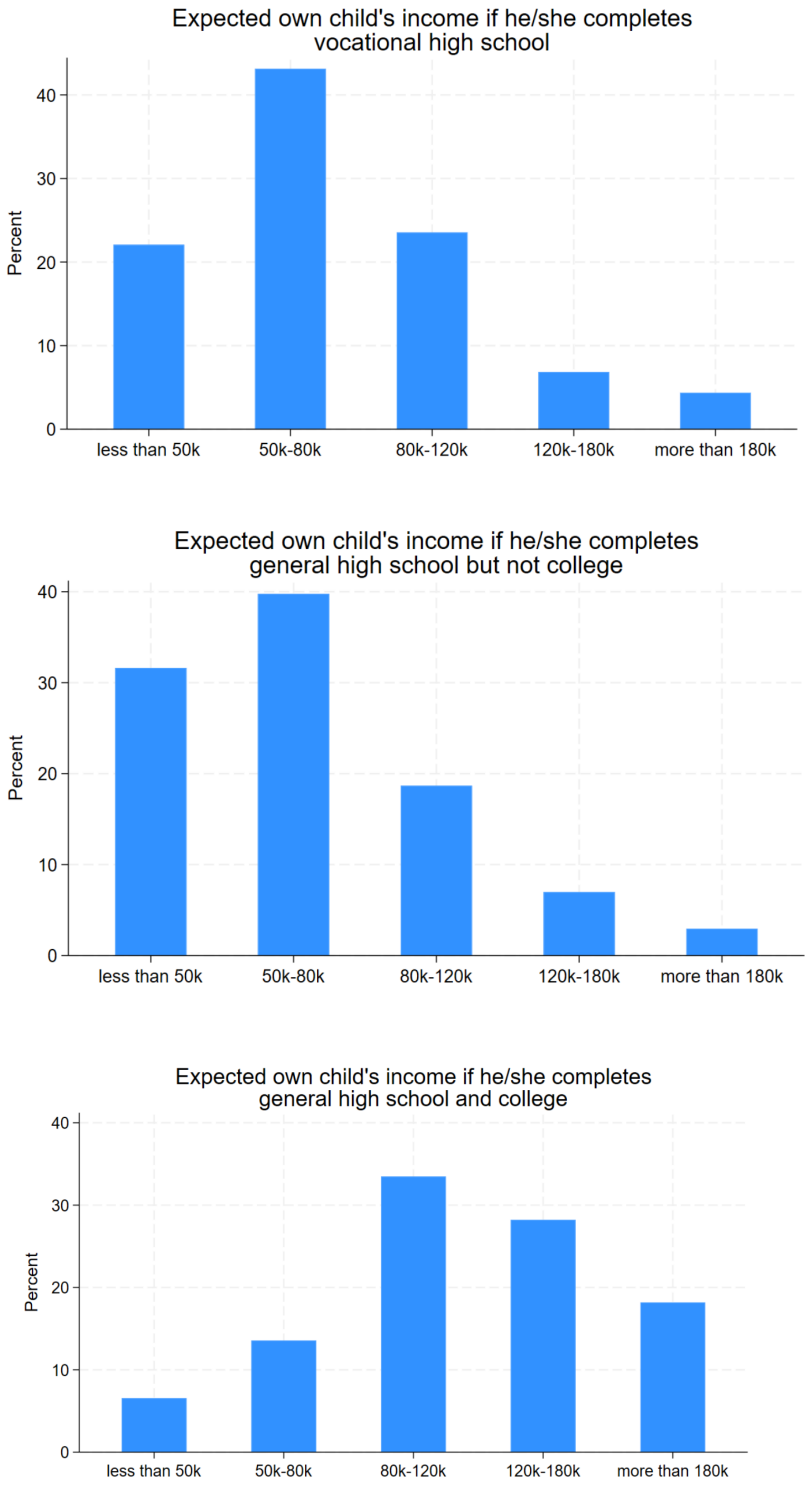


Figure 46: Perceived returns to different schooling levels

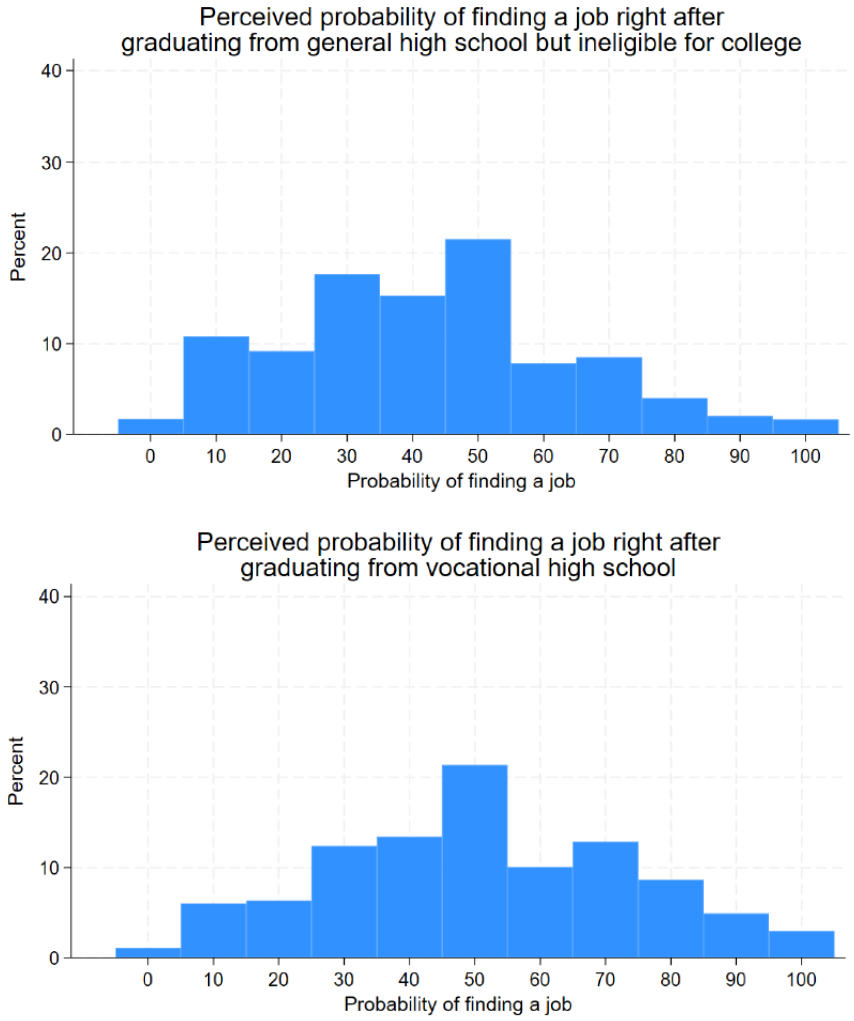


Figure 47: Perceived probability of finding a job right after graduation

Table 39: Parents' opinions on why very few parents enroll their kids in vocational schools

Factor	Yes	No
Low future earnings	14.31%	85.69%
Accessibility or distance from home	10.93%	89.07%
Desire for their child to be surrounded by high achievers	28.42%	71.58%
Fear their child will be perceived as a low achiever	14.61%	85.39%
Belief that general high school increases chances of college enrollment	27.82%	72.18%
Vocational schools are male-dominated	12.37%	87.63%
Vocational education limits future college and career opportunities	40.64%	59.36%

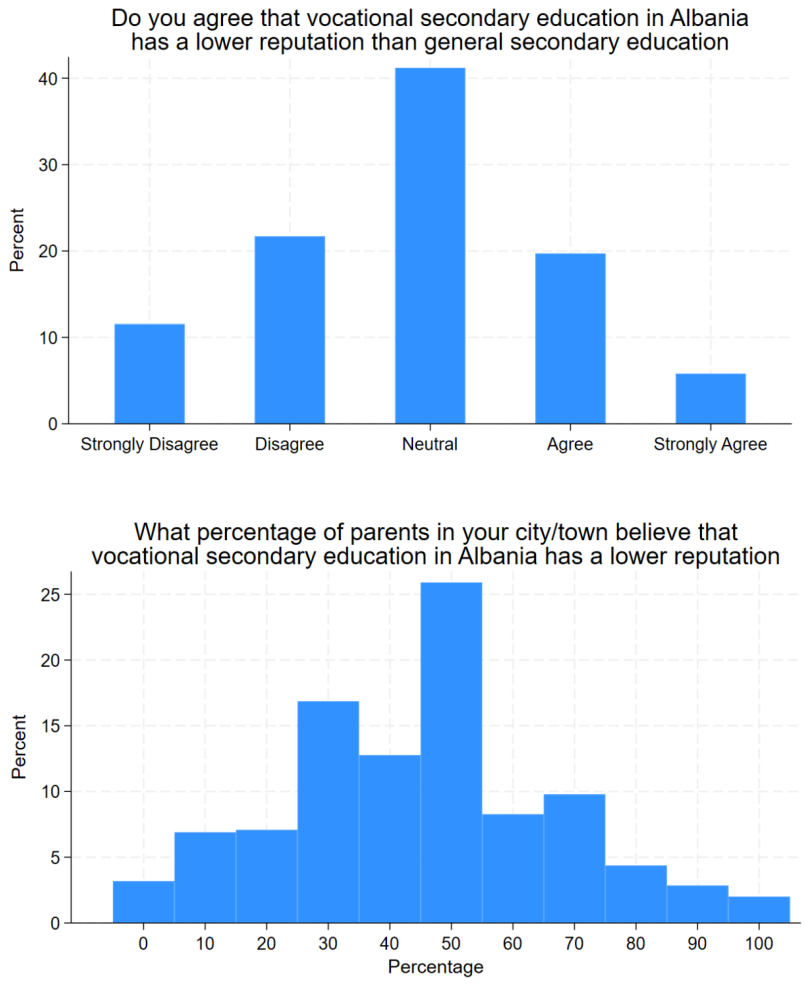


Figure 48: Misperceived stigma about the vocational track

Table 40: Can someone apply to college if they graduate from the vocational track?

Answers	Percentage
Yes, to any college	36.64%
Yes, but only to some colleges	56.23%
No, because only the academic track leads to college	7.13%