Altruistic or Expected Leadership?

Laboratory Evidence on What Motivates Pro-Social Influence*

Mauricio Fernández-Duque† and Michael Hiscox‡

September 20, 2022

Abstract

We provide experimental evidence to assess whether pro-social leading is motivated by altruism or by a desire to conform with expectations. A sizable minority of individuals act more pro-socially when they can influence others. Whether or not an individual acts more pro-socially when she can influence others, she wants her actions made public only if it will increase others' pro-sociality, which turns out to be a telltale sign of altruism. Despite some evidence that pro-social leading is partly driven by conformity, altruism wins the day.

Keywords: Leadership; reluctance; social expectations; altruism; lab experiment

JEL Classification: C7, C9, D8

*We would like to thank Samuel Asher, Nava Ashraf, Robert Bates, Hernán Bejarano, Iris Bohnet, Marta Ceballada, Herminio José Chanona Tevera, Serena Eastman, Alexander Elbittar, Sonia DiGiannatale, Diego Domínguez, Ryan Enos, David Fernández, Ben Golub, Emilio Gutiérrez, Carlos Lever, Eric Magar, Fernando Pérez Cervantes, Germán Rojas, Alexandra Uribe, Jeffrey Weldon, excellent assistance by Sergio Basilio, Fabiola Bustamante, Silvia de Hoyos, Fátima Trujillo, staff at ITAM, CIDE and Tec De Monterrey campus Santa Fé, and seminar participants at Harvard University and MIT. We would also like to thank the support of the Harvard Decision Science Lab, and in particular Nina Cohodes.

†Corresponding author. CIDE, Carretera México-Toluca 3655, Col. Lomas de Santa Fé, Del. Cuajimalpa, 01210. Email: mauricio.fernandez@cide.edu. Declaration of interests: none.

‡Harvard University. Email: hiscox@fas.harvard.edu. Declaration of interests: none.

In everyday interactions, individuals find themselves in a position where their actions set an example for others, such as when senior employees are asked to step forward first to participate in collective activities, when older siblings are asked to play nicely for the younger children to see, or when organizations ask individuals to become "influencers" by posting on social media that they voted or donated blood. Individuals in these situations sometimes seem to *lead*, in the sense that they act more pro-socially when they can influence others than when they cannot. We will refer to this type of leadership as *contributing extra*. Even if individuals do not contribute extra, they may lead in the sense of *seeking primacy*, which is when they put themselves in situations when their action can influence others, such as employees volunteering to go first in collective activities (see Figure 1).[1] Both types of pro-social leading are often explained as motivated by altruism (Gächter et al., 2012; Karlan and McConnell, 2014; Jack and Recalde, 2015; Karlan and List, 2018). Another motivation commonly used to explain pro-sociality more generally is a social-image concern (Bénabou and Tirole, 2006; Meyer and Tripodi, 2021).

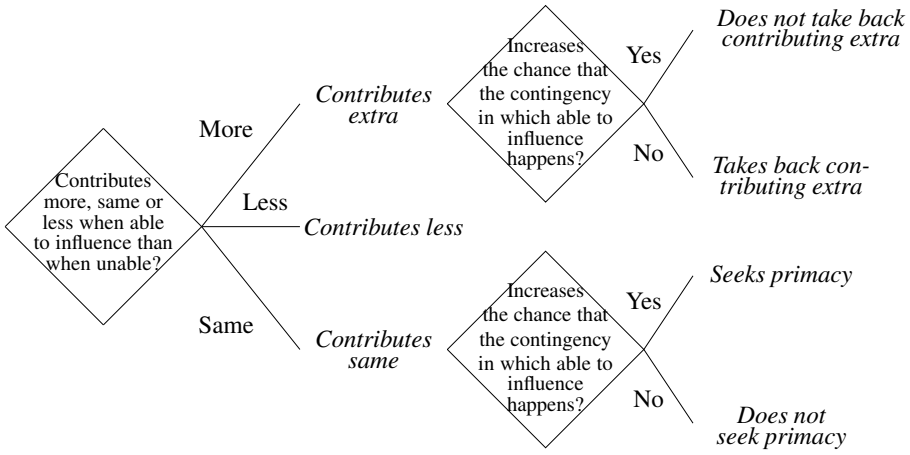

Figure 1: Leadership Nomenclature

In this paper we investigate a third possible motivation for pro-social leading that has received far less attention to date: the desire to conform with expectations, or what individuals believe is

[1] An alternative way to define leadership would be in terms of whether a first-mover's action has an impact on second-movers' actions. In this paper we focus on leadership defined as contributing extra or seeking primacy, since this allows us to study the motivation for leadership independently of the impact it has. Pro-social leading often involves a publicly observed individual taking an action before others take action. However, the observability of the individual is not essential to influence others. Leadership may also be defined as a willingness or propensity towards taking actions that affect others (e.g. Ertac et al., 2020).

the expected behavior in a given situation. To explain our agenda more concretely, suppose a first-mover is taking an action that benefits a welfare-enhancing cause (e.g. by donating time or money to a charity). She may lead by contributing extra or by seeking primacy. Her leadership may be driven by one of three motivations. First, an altruistic concern: a desire to help the cause or a desire to self-signal that she likes helping the cause.[2] Second, a social-image concern: a desire to have a well-regarded reputation in the eyes of others (Buser et al., 2021; Ren et al., 2022). Third, an expectations concern: a desire to go along with what she thinks others expect of her, which is not driven by an image concern. An expectations concern may arise as a heuristic to follow what is expected in a given situation, perhaps out of an ingrained desire to not disappoint others even in situations where one's reputation is not at stake. Research on pro-social behavior has revealed that expectations matter a great deal in general. People tend to behave in pro-social ways when they think it is expected by others, even when the behavior is anonymous (Dana et al., 2006; DellaVigna et al., 2012; Salazar et al., 2022).

We present the results of an experiment that distinguishes between the altruistic and the expectations concerns to lead, while minimizing and holding fixed the social-image concern. This paper is the first, we think, to explore the role that expectations concerns might play in pro-social leading.

To distinguish whether first-movers contribute extra out of expectations concerns or altruistic concerns (which include self-image concerns), we allow first-movers to contribute extra: we ask them to contribute when they can and when they cannot influence a second-mover. After the first-movers make their contribution decisions, we allow them to choose the probability p of implementing the contingency where their contribution can influence others. This allows first-movers to *take back* contributing extra in the sense of minimizing p. If first-movers had instead contributed the same and then they maximize p, we say they seek primacy.

Note that to isolate the motivation for leading, the strategic interaction is kept deliberately

[2]It may seem a bit strange to classify self-signaling as an instance of altruistic concerns, and to separate it from social-image concerns. However, this classification is helpful for showing how the design distinguishes between motivations.

sparse: subjects are anonymous and first-movers do not learn what their second-movers contributed.[3]

Taking back contributing extra is *prima facie* evidence of reluctance—it suggests that the first-mover led pro-socially simply to fulfill an obligation she would rather avoid. However, this is not the whole story. A first-mover may contribute extra, but believe that the contingency where she can influence others does not yield a higher contribution to the public good because second-movers will contribute less in that contingency. This logic may lead altruists to take back contributing extra not because they had reluctantly contributed extra, but in order to increase the impact of their action on the public good. To test for this possibility, we elicited first-movers' beliefs about second-movers' contributions. With these beliefs in hand, we could directly test whether first-movers were motivated by altruistic concerns in their contribution decisions, and in their decisions to take back contributing extra. To more directly test for expectations concerns, we also measured expectations over what first-movers do, and what we call social expectations: expectations over what others expected first-movers do. Eliciting both measures allows us to test whether different definitions of expectations can explain behavior. The decision to alter p and the elicited beliefs will also allow us to distinguish the motives behind contributing extra and seeking primacy.

Our experimental evidence indicates that first-movers contribute extra or seek primacy out of an altruistic concern to increase contributions. Indeed, although the evidence from elicited beliefs provides partial support for expectations concerns, the most solid support is for altruistic concerns. Viewing pro-social leadership in this light helps in the design of policies to increase this type of leadership, and to better understand its welfare implications.

Related literature. Karlan and McConnell (2014) is the closest work to ours, as it is the only experiment we are aware of in which subjects' contributions are displayed either before or after others make their contribution decision. In contrast to their design, we abstract from social image concerns, and ask subjects to make decisions and report beliefs with which to distinguish between

[3]The second-mover will either see the contribution before making his own decision or after, in either case being able to pass judgment on the contribution. This is the sense in which we hold fixed the social image concern. We minimize social image concerns in the sense that the contribution decisions are anonymous.

altruistic and expectations concerns. We defer to Section 3 a comparison of their results with ours.

Other papers also make inferences about first-movers' motivation to contribute extra, by comparing what they contribute as first-movers with what they contribute as second-movers (e.g. Arbak and Villeval, 2013; d'Adda, 2017), or with what they contribute in private (Jack and Recalde, 2015). We believe neither of these are ideal counterfactuals: subjects in a position of second-movers may be influenced by what the first-mover does, while private contributions do not hold the visibility of the first-mover's action constant (as done in our design and in Karlan and McConnell, 2014).

Our work differs from past papers in combining a within-subject design with the elicitation of beliefs and beliefs over beliefs to infer first-movers' preferences. Gächter et al. (2012) also elicited beliefs, finding that first-movers contribute more when they expect others to respond positively to their contribution. However, the impact of a higher slope of response on altruistic first-movers' contributions is ambiguous: a higher slope has an income effect (the level change in the second-mover's contribution) and a substitution effect (how the second-mover's contribution changes as the first-mover's contribution increases) which may go in different directions.

In contrast to Gächter et al. (2012), we develop in Section 1 a test of altruistic concerns in contributing extra which relies on how the first-movers believe the second-mover's level of contribution changes depending on whether the first-mover can or cannot influence the second-mover. We further use elicited beliefs to test for the motivation behind seeking primacy. Although there is some work on seeking primacy (Potters et al., 2005; Rivas and Sutter, 2011; Arbak and Villeval, 2013), our design is the first to consider whether this behavior is affected by expectations concerns.

In our experiment, as in charitable contributions in most natural settings, first-movers receive no material rewards from pro-social leading. Many public goods experiments provide material rewards to a first-mover who can influence others to contribute (e.g. Güth et al., 2007; Komai and Grossman, 2009; Levy et al., 2011; Bracha et al., 2011; d'Adda, 2017; Frackenpohl et al., 2016; Brandts et al., 2016). In these settings, successful leading implies that first-movers receive benefits from others' public goods contributions, and therefore altruism and selfishness are confounded.

1 Experiment Design, Hypotheses and Procedure

1.1 Experimental Design

We assign subjects to be either a *first-mover* who anonymously splits an endowment with a charity, a choice that will be observable to a *second-mover* at some point. First-movers were matched one-to-one to a second-mover, and each subject either was a first-mover once (and not a second-mover), or as a second-mover once (and not a first-mover). First-movers and second-movers both split an endowment of the same size between themselves and a charity. For clarity of exposition, we will refer to the first-movers in the female gender, and to the second-movers in the male gender.

Figure 2 presents the timeline of the experiment.

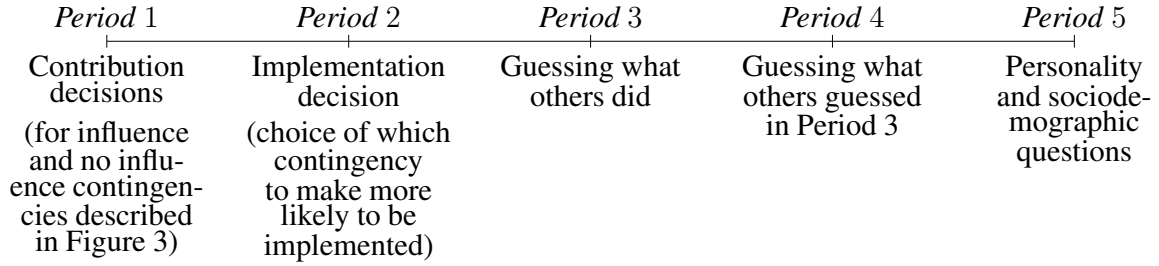

Figure 2: Timeline of the Experiment

Contribution decisions. In period 1, first-movers make a contribution decision which will be seen by a second-mover *before* he makes his own contribution decision. We refer to this sequence of decisions as the *influence contingency*. First-movers also make a contribution decision which will be seen by a second-mover *after* he makes his own contribution decision. We refer to this sequence of decisions as the *no-influence contingency*. The two contingencies are depicted in Figure 3. First-movers learn about both decisions simultaneously, which pilot sessions indicated helped comprehension, and was further done to avoid anchoring the second contribution decision on the first.[4]

[4]Presenting both contingencies simultaneously implies that we cannot compare in isolation the first decision of first-movers who are assigned a influence contingency first and those who are assigned a no-influence contingency first. However, a between-subjects comparison would not have helped the main objective of the paper: to test for altruistic concerns and expectations concerns in leadership behavior. Our test requires observing the same first-movers' decisions in the influence and no-influence contingencies.

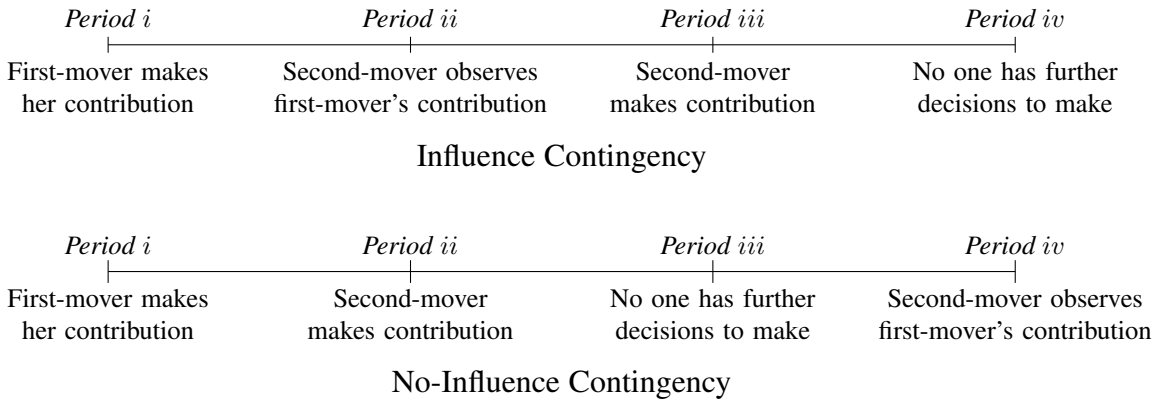

Figure 3: Timelines of Influence and No-Influence Contingencies

Implementation decision. Only one of the contingencies will be implemented. The first-mover's third decision affects the probability of implementation. That is, suppose she had originally decided to contribute x in the influence contingency, and y in the no-influence contingency. Then in her third decision she decides which of the two contingencies is implemented with $2/3$ probability. We refer to this decision as the *implementation decision*.

Belief elicitation. After these choices have been made, we ask the first-movers a series of questions to elicit beliefs about others and measure social expectations, presented as "guessing games" (with monetary prizes for those who came closest to the correct answers). We asked each first-mover to guess how much the average second-mover contributed to the charity in the influence contingency (a guess for each multiple of 10 percent of the endowment that the first-mover could have contributed first) and in the no-influence contingency (one amount to guess). After this they were asked to guess how much the average first-mover contributed in the influence and no-influence contingencies. This is our measure of expectations. To generate measures of social expectations, in the second guessing game we asked each first-mover to guess the average of what others had guessed in the first part of the guessing game. This is our measure of social expectations.[5]

[5] A similar procedure for eliciting second-order beliefs was implemented in Charness and Dufwenberg (2006). Our methodology contrasts with that of Krupka and Weber (2013), who use coordination games to identify social norms. Although the concept of social norms is sometimes used similarly to social expectations, their elicitation method may capture something different than what we have in mind. For instance, subjects may coordinate on a response because it is a salient number or because it is a fair allocation. By asking subjects to guess what others had already guessed, we avoid multiple equilibria that may arise from a motivation to coordinate on the same guess, and the interpretation of the answer is more straightforward.

Personality and socio-demographic questions. After the guessing games, first-movers are asked to answer socio-demographic questions, personality questions and an exit survey.

First-movers are not informed of any details about second-movers other than what we have explained so far.

1.2 Hypotheses

Most of the hypotheses are stated in the pre-analysis plan registered in the AEA RCT registry (AEARCTR-0009191), which were formulated based on the results of the analysis of the US subject pool. We will point out the hypotheses that were not included in the original pre-analysis plan, and the rationale for including them.

Actions. We begin by stating hypotheses which focus on first-movers' actions.

In order to study the motivation behind contributing extra, a necessary condition is for contributing extra to exist. Our first hypothesis is therefore quite natural.

Hypothesis 1 (Contributing extra)**.** *A non-negligible percentage of first-movers contribute extra.*

Despite the results of Karlan and McConnell (2014), who do not find evidence for contributing extra, we believe contributing extra likely happens in some settings.

A complementary hypothesis to Hypothesis 1 is that contributing extra is more common than *contributing less*: contributing more in the no-influence contingency than in the influence contingency. Contributing less can be rationalized under altruistic concerns or by expectations concerns. An altruist may contribute less if she believes the second-mover's contribution decreases with her own contribution. A first-mover who cares about expectations would contribute less if expectations for contributing in the no-influence contingency are higher. However, we expect contributing less is not that common empirically.

Hypothesis 2 (More contributing extra than contributing less)**.** *The proportion of first-movers who contribute extra is higher than the proportion of first-movers who contribute less.*

Hypothesis 1 and 2 follows from the assumption, common in the literature, that contributing extra is altruistic in the sense that first-movers go out of their way to increase some public good.[6] We have already mentioned that beliefs can rationalize contributing less under different underlying preferences, and we will get to beliefs in short order. However, as discussed in the introduction, the literature as a whole mostly emphasizes the interpretation of contributing extra as being altruistic[7] Under this interpretation, the decision of whether or not to take back contributing extra can be taken as *prima facie* evidence to distinguish whether contributing extra is driven by altruistic concerns or expectations concerns: an altruist who contributes extra because it increases a public good would not take back contributing extra.

Hypothesis 3a (*Prima facie* altruistic concerns in contributing extra)**.** *Most first-movers do not take back contributing extra.*

Hypothesis 3b (*Prima facie* expectations concerns in contributing extra)**.** *Most first-movers take back contributing extra.*

What about those who do not contribute extra? Contributing the same amount in both contingencies could in principle be motivated by altruistic concerns, since the income and substitution effects of the influence contingency could cancel each other out (as we will elaborate shortly).[8] If altruistic concerns were driving the contribution decisions, then we should see that those whose contributions have a larger impact on second-movers' contributions should be the ones to seek primacy. We then propose the following, in line with the common interpretation that higher first-movers' contributions have a higher impact on second-movers' contributions:

Hypothesis 4 (More Seeking Primacy Among Higher Contributors)**.** *First-movers who contribute the same amount in both contingencies are more likely to seek primacy if their contribution is higher.*

[6]Note that this is true whether we are using "altruistic concerns" as an intrinsic preference to increase total contributions or a desire to self-signal altruism, and we use the term to encompass both interpretations.

[7]An exception is Varian (1994), who points out that an altruist may contribute less to compel the second-mover to contribute more.

[8]The substitution effect is how second-movers react to a change in the first-mover's contribution within the influence contingency, and the income effect is the level change of the second-mover's contribution between the influence and no-influence contingency.

Note that the pattern of behavior in Hypothesis 4 does not pin down altruistic behavior. On the one hand, the first-movers may not believe seeking primacy increases contributions. On the other hand, there is the alternative explanation, consistent with expectations concerns, that first-movers do not want to publicize behavior that falls short of social expectations. Once again, to distinguish between these stories we must take into account the role of beliefs, to which we now turn.

Beliefs regarding contributing extra: altruistic concerns.[9] Using beliefs to pinpoint altruistic concerns is somewhat tricky. Call the *slope of response* the slope with which the second-mover responds to the first-mover's contribution (we assume the slope is constant for simplicity). It may appear straightforward that an altruistic first-mover contributes more in the influence contingency the higher the slope of response. However, this does not turn out to be true. The slope of response may have counteracting effects on the optimal contribution of an altruistic first-mover, and so the optimal contribution depends on the value of the first-mover's contribution over which the slope pivots.

Despite the counteracting effects of the second-mover's response on the first-mover's contribution, we can make sharp statements if we fix the right values over which to pivot the slope.[10] We define the *counterfactual contribution* as the amount the second-mover would contribute in the influence contingency if the first-mover contributes the same amount she did in the no-influence contingency.[11] If we fix the counterfactual contribution, it is straightforward that an increase in the slope of response increases the alturistic-first-mover's contribution in the influence contingency—there is a pure substitution effect. Similarly, the altruistic-first-mover's contribution in the influence contingency decreases with the counterfactual contribution if the slope is fixed—there is a pure income effect. This leads us to the following.

[9] An altruist motivated by increasing donations will be observationally equivalent to an altruist motivated by self-signaling. This is because the first-mover observes her own actions, so would need to act like an altruist in her contribution and implementation decision to send the right signal to herself.

[10] This hypothesis was not included in the pre-analysis plan. We included it to see whether we could find evidence for altruistic concerns in the decision to contribute extra.

[11] If an altruistic first-mover expects the slope of response to be positive and the counterfactual contribution to be large enough, she will contribute more than in the no-influence contingency. Indeed, she would expect the overall contribution to be larger than in the no-influence contingency, and would receive higher utility than in the no-influence contingency. It is sufficient but not necessary for the counterfactual contribution to equal what the second-mover contributes in the no-influence contingency.

Hypothesis 5 (Beliefs consistent with altruistic concerns in contributing extra)**.** *Holding the slope of response constant, the difference between contributing in the influence and no-influence contingencies increases with the counterfactual contribution.*

Holding the counterfactual contribution constant, the difference between contributing in the influence and no-influence contingencies increases with the slope of response.

Beliefs regarding contributing extra: expectations concerns. Now we turn to how expectations concerns may motivate contributing extra. To capture expectations concerns, suppose a first-mover gets utility that depends on how much of her endowment of 1 she keeps, $u(1-f)$, and in the difference between how much she contributes and the expectations E_c of how much is contributed in that contingency c, $v(f-E_c)$. (In the empirical analysis, we will in fact consider both expectations and social expectations in the place of E_c. Although this confers an extra degree of freedom to the expectations concern hypothesis, we include these tests for two complementary reasons. First, it is not *ex ante* obvious which of the two measures better captures these concerns. Second, because it will turn out that the case for expectations concerns is somewhat weak, and showing both measures makes the weakness of the case more obvious.)

The first-mover's utility is then $u(1-f)+v(f-E_c)$, with u increasing and weakly concave, v increasing and concave, and c equal to either the influence or the no-influence contingency.[12] Note that the shape of u may capture the trade-off between selfish and altruistic concerns in reduced form. A reasonable further wrinkle to this model is that fulfilling expectations concerns are more important in the influence contingency: not fulfilling expectations concerns may be more costly when the action can influence others, even if it is anonymous.[13] With this utility function, it is easy to see the following.

Hypothesis 6 (Beliefs consistent with expectations concerns in contributing extra)**.** *The difference between contributing in the influence and no-influence contingencies increases with (social) expec-*

[12] Charness and Dufwenberg (2006) have a related model of guilt aversion in which a first individual considers his beliefs over a second individual's beliefs. In contrast to our model, the first individual is concerned with the second individual's beliefs over how the first individual's action affects the second individual's payoffs.

[13] Formally, this would be captured by allowing the function v, or the weight given to v, to vary with the contingency.

tations in the influence contingency and decreases with (social) expectations in the no-influence contingency. The first effect is stronger than the second.

Beliefs regarding taking back contributing extra: altruistic concerns. To provide a necessary condition for altruists to take back contributing extra, the key is to focus on the comparison of two beliefs held by the first-mover. First is how much the second-mover contributes in the no-influence contingency. Second is the counterfactual contribution. It turns out that in order for an altruist to take back contributing extra, the second amount must be lower than the first.

Here is the proof of this claim.

Suppose the first-mover's utility increases in how much she keeps of her endowment, and in total contributions to the charity. The first-mover contributes an amount $\hat{f}$ in the no-influence contingency, and believes the second-mover contributes an amount $\hat{s}$. If the first-mover believes the second-mover will contribute an amount $\hat{s}' \geq \hat{s}$ in the influence contingency if she contributes $\hat{f}$, then the first-mover is weakly better off in the influence contingency (strictly if $\hat{s}' > \hat{s}$). Indeed, the first-mover can always contribute $\hat{f}$ and be weakly better off than in the no-influence contingency. Therefore, for an altruistic first-mover to take back contributing extra, she must believe $\hat{s}' < \hat{s}$.

With this claim in hand, we can form the following hypothesis.

Hypothesis 7 (Beliefs consistent with altruistic concerns in taking back contributing extra)**.** *The probability that a first-mover takes back contributing extra is higher if she believes that the second-mover contributes more in the no-influence contingency than the counterfactual contribution ($\hat{s}' < \hat{s}$).*

Beliefs regarding taking back contributing extra: expectations concerns. Since we have hypothesized that (social) expectations affect contribution decisions out of expectations concerns, it follows that an increase in contributions driven by (social) expectations would increase taking back contributing extra.[14] To be consistent with the wrinkle introduced in Hypothesis 6, we will again assume that fulfilling (social) expectations is more costly when the action can influence others.

[14]This hypothesis was not included in the pre-analysis plan. We included it to include a hypothesis that related expectations concerns directly to the decision to take back contributing extra.

Hypothesis 8 (Beliefs consistent with expectations concerns in taking back contributing extra)**.** *The probability that a first-mover takes back contributing extra increases in the (social) expectations in the influence contingency and decreases in the (social) expectations in the no-influence contingency. The first effect is stronger than the second.*

Beliefs regarding seeking primacy: altruistic concerns. If first-movers' implementation decision is driven by altruistic concerns, then those who contribute the same in both contingencies will be more likely to seek primacy if they expect second-movers to contribute more.

Hypothesis 9 (Beliefs consistent with altruistic concerns in seeking primacy)**.** *First-movers who contribute the same amount in both contingencies are more likely to seek primacy if the contribution they expect from second-movers in the no-influence contingency is lower than the contribution they expect in the influence contingency.*

Notice that, in the notation of Hypothesis 7, the condition for first movers to seek primacy in Hypothesis 9 is $\hat{s}' > \hat{s}$. Hypotheses 7 and 9 are therefore closely related.

Beliefs regarding seeking primacy: expectations concerns. We have already suggested an expectations concerns story to explain seeking primacy: that first-movers are willing to make the influence contingency more likely to be implemented as long as it fulfills (social) expectations.[15] We then propose the following:

Hypothesis 10 (Beliefs consistent with expectations concerns in seeking primacy)**.** *First-movers who contribute the same amount in both contingencies are more likely to seek primacy if their contributions in the influence contingency are at least as high as their (social) expectations in that contingency.*

[15]Note that this story is consistent with the "wrinkle" introduced in Hypothesis 6 that fulfilling expectations concerns is more important in the influence contingency. That fulfilling (social) expectations is particularly important (as opposed to surpassing them) seems intuitive, and can be modeled by allowing for v to be zero in the no-influence contingency, for v to be negative in the influence contingency when the contribution is below (social) expectations, and for v to be positive in the influence contingency when the contribution is weakly above (social) expectations.

1.3 Procedure and the Charity

Procedure. The experiment was run in September 2013 through the Harvard Decision Science Lab subject pool (with 93 first-movers), and in May 2022 in the classrooms of three private Mexico City universities: ITAM, CIDE and ITESM Santa Fe campus (with 301 first-movers, although 6 first-movers dropped out at different points of the session). Although we will pool the results, the full list of differences in the implementation of the US and Mexico subject pools are summarized in Appendix B. Participants filled out consent forms before beginning the experiment,[16] their data was collected without identifying information, and subjects knew payment would be done without anyone observing what they received.

To help ensure that our instructions were clear and that levels of comprehension were high, we included a questionnaire screen that quizzed first-movers after they read each key set of instructions. After first-movers answered each of these questionnaires, a screen with the answers would appear, specifying which questions they had answered correctly or incorrectly. They were then shown the instructions again before moving on to the task. First-movers answered the correct question 88% of the time in the US sample, and 94% of the time in the Mexico sample (among those who completed the whole survey).

The Charity. The charity that received contributions from subjects in the experiment was the East Africa Fund of Save the Children. The aim of this fund is to address starvation and malnutrition in East Africa. The remoteness of the charity's activities was chosen in order to make it unlikely that these could impact subjects in any direct way. The activities of this Save the Children fund are conceptualized here as a public good (non-rival and non-excludable) among our subjects who care about helping starving children in East Africa. The instructions stated that Save the Children would not know the origin of the contributions raised via the experiment, to avoid social-image concerns towards the charity.

[16]The US implementation went through an ethics review process in the Harvard Committee on the Use of Human Subjects, and has protocol number IRB13-1169. The Mexico implementation went through an ethics review in the ITAM IRB.

2 Results

The results below pool the evidence from the US and Mexico samples. Many of the results are qualitatively similar across both samples, although the pooled results will mostly be driven by the much larger Mexican sample. We present the full analysis of each sample in Online Appendix D, and summarize the difference in the results across both samples in Section 3.1.

2.1 The 22% of First-Movers Who Contribute Extra Often Take It Back

Summary Statistics. The data for this experiment can be found in Fernández-Duque and Hiscox (2022). There were 394 first-movers, and each first-mover was matched to a second-mover. The average age of first-movers is 23.29 years (standard deviation 8.02). There were 71% of first-movers who were college students, although 74% had completed college in the US sample. A little more than half of first-movers (60%) were male.

Decisions. In Figure 4 we summarize the choices made by the first-movers in the experiment.[17] In the Figure, we plot a histogram of first-movers who contributed more in the influence contingency (that is, the 22% of first-movers who contributed extra), first-movers who contributed the same (70% of first-movers), and first-movers who contributed less (8% of first-movers). We further break down by halves the distribution of those who contributed the same amount in both contingencies: first-movers who contributed at most the median contribution in their respective US or Mexico pool (35% of first-movers, or 50.55 among those who contributed the same amount in both contingencies), and those who contributed more than the median (34% of first-movers, or 49.45% among those who contributed the same amount in both contingencies). Panel A of Table 1 reports the statistics in more detail. On average, first-movers contributed 60.28% of their endowment to charity in the influence contingency compared with 56.03% in the no-influence contingency (significantly different at the 5% level using a paired t-test, p-value=0.0000). Among those who contributed extra, first-movers contributed 25.86 percent more of their endowment in

[17] Tables 6 and 7 in Online Appendix A present the joint frequency of contributions in the influence and no-influence contingencies.

the influence contingency.

	Mean	S.D.	Max	Min	Count
Panel A. Contribution Decisions					
Influence-Contingency Contribution	0.60	0.36	1	0	394
No-Influence-Contingency Contribution	0.56	0.37	1	0	394
Contributed Extra	0.22	0.42	1	0	394
Contributed Same	0.70	0.46	1	0	394
Contributed Less	0.08	0.27	1	0	394
Absolute Difference In Contributions | Contributed Extra	0.25	0.20	1	.01	88
Absolute Difference In Contributions | Contributed Less	0.20	0.13	0.65	0.04	31
Panel B. Implementation Decisions					
Percent Implementing Influence Contingency | Contributed Extra	0.42	0.50	1	0	88
Percent Implementing Influence Contingency | Contributed Less	0.61	0.50	1	0	31
Percent Implementing Influence Contingency | Contributed Same	0.61	0.49	1	0	273
Percent Implementing Influence Contingency | Contributed Same, Below or Equal the Median	0.52	0.50	1	0	137
Percent Implementing Influence Contingency | Contributed Same, Above the Median	0.71	0.46	1	0	136

"Influence-Contingency Contribution" is the percent of her endowment the first-mover contributed in the influence contingency. "No-Influence-Contingency Contribution" is the percent of her endowment the first-mover contributed in the no-influence contingency. "Contributing Extra" is an indicator variable equal to 1 if the first-mover contributed more in the influence contingency than in the no-influence contingency. 'Contributed Same' and "Contributed Less" are defined analogously. "Absolute Difference in Contributions | Contributed Extra" is the absolute difference in contributions (as a percent of her endowment) in the influence and no-influence contingencies among those who contributed more in the influence contingency. "Absolute Difference in Contributions | Contributed Less" is defined analogously. "Percent Implementing Influence Contingency | Contributed Extra" is an indicator variable equal to 1 if the first-mover made the the influence contingency more likely to be implemented, among those who contributed more in the influence contingency than in the no-influence contingency. "Percent Implementing Influence Contingency | Contributed Less", and"Percent Implementing Influence Contingency | Contributed Same"are defined analogously.

Table 1: Summary of Outcome Variables

Note that the number of first-movers who contributed extra, as well as the amount contributed extra, is censored due to the fact that the amount first-movers could contribute was bounded. Among first-movers who contributed the same amount in both contingencies, 58% either always contributed nothing, or always contributed the full endowment. Nevertheless, we have the following:

Result 1 (Contributing extra)**.** *There is a non-negligible percent (22%) of first-movers who contributed extra.*

Result 2 (More contributing extra than contributing less)**.** *The proportion of first-movers who contributed extra is significantly higher than the proportion of first-movers who contributed less (22% versus 8%, p-value of 0.0000 with a paired t-test).*

We can categorize first-movers according to which contingency they implemented. To see this graphically, each bar in Figure 4 is divided into a shaded area and a light area. The shaded area represents the proportion of first-movers within the corresponding bar who implemented the influence

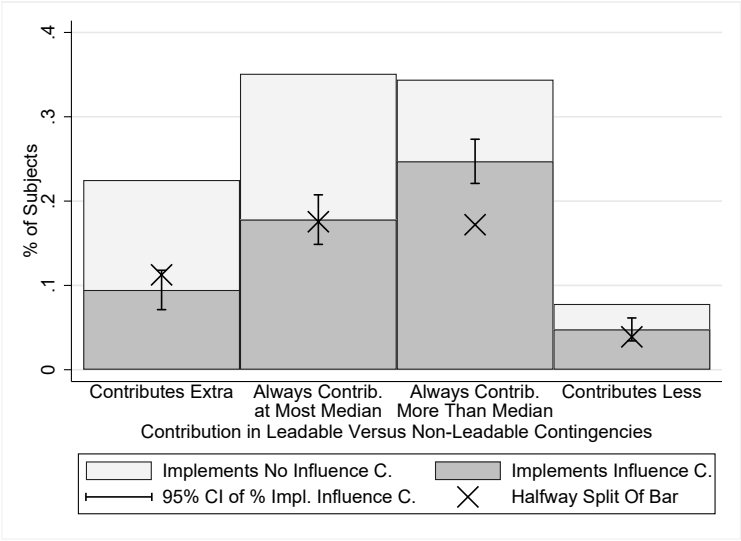

Figure 4: Those Who Contribute Extra Are More Likely to Implement the No-Influence Contingency (Thus Taking Back Contributing Extra)

contingency, and the light area is the proportion who implemented the no-influence contingency. The results can also be seen in Panel B of Table 1. In order to test whether the implementation rate within each bar is different from a halfway split, Figure 4 also marks the halfway split of each bar and the 95% confidence interval of the implementation rate. Of the first-movers who contributed extra, 58% took back contributing extra, although the percentage is not significantly different from 50%.

Result 3 (*Prima facie* expectations concerns)**.** *Among first-movers who contribute extra, 58% take back contributing extra, which is not significantly larger than 50% (p=0.0682 in a t-test).*

The implementation decisions of those who contributed the same amount in both contingencies varied with how much they contributed. Among the bottom half, half made the influence contingency more likely to be implemented. Among the top half, significantly more than half made the influence contingency more likely to be implemented. Table 8 in Online Appendix A provides a regression analysis of those who contributed the same amount in both contingencies, with qualitatively similar results.[18]

[18]The significance of the main coefficient disappears when we control for beliefs, demonstrating the importance of taking beliefs into account to explain the first-movers' behavior, as we will do below.

Result 4 (More seeking primacy among higher contributors)**.** *First-movers who contribute the same amount in both contingencies are more likely to seek primacy if their contribution in the influence contingency is higher. Among the bottom half, half seeks primacy (52% implements the influence contingency, significantly indistinguishable from 50%). Among the top half, more than half seeks primacy (71% implements the influence contingency, significantly different from 50% with a p-value of 0.0000). The proportion of first-movers who seek primacy in the top half is significantly different from those in the bottom half (p-value 0.0014 with a two-sample t-test).*

Among those who contributed less, 61% made the influence contingency more likely to be implemented. Although the behavior of those who contribute less is intriguing, we are underpowered to say much of interest about them.

Does leadership work, regardless? We find mixed results to this question. In a companion paper we explore the impact of a passive audience on second-movers' contributions (Fernández-Duque and Hiscox, 2022). The analysis only used data from the US sample. We find that second-movers with a passive audience contribute an extra 21 cents for every additional dollar the first-mover contributes. However, there is no impact of second-movers without an audience. This is also true for our pooled sample. (In the Mexico pool, second-movers never had an audience.) This result may be partly due to censored data. In the Mexico pool alone, 267 of the 317 second-movers contributed the full endowment.

2.2 Beliefs

Before testing the hypotheses related to beliefs, we present some summary statistics. In panel A of Table 2 we report summary statistics for measures relevant to altruistic concerns.

We first consider the slope of response, which captures first-movers' beliefs over how responsive second-movers are to first-movers' contributions in the influence contingency. We calculated the slope of response via a regression for each first-mover.[19] We classified the slopes as positive or

[19]The independent variables in the regression were a constant and the deciles of the endowment the first-mover could have contributed. The dependent variable was the first-mover's prediction of what the second-mover contributes

	Mean	S.D.	Max	Min	Count
Panel A. Beliefs Relevant to Altruistic Concerns					
Slope of Response	0.05	.05	1.1	-1	390
Strictly Positive Response	0.83	0.38	1	0	390
Zero Response	0.06	0.25	1	0	390
Strictly Negative Response	0.11	0.31	1	0	390
Condition for Altruists to Take Back Contributing Extra	0.30	0.46	1	0	390
Seeking Primacy Increases Contributions | Contributed Same	0.35	0.48	1	0	390
Panel B. Beliefs Relevant to Expectations Concerns					
Social Expectations of Influence Contingency	0.57	0.27	1	0	389
Social Expectations of No-Influence Contingency	0.37	0.24	1	0	389
Social Expectations of Influence Contingency | Contributed Extra	0.57	0.21	1	0	85
Social Expectations of Influence Contingency | Did Not Contribute Extra	0.58	0.29	1	0	304
Social Expectations of No-Influence Contingency | Contributed Extra	0.35	0.21	1	0	85
Social Expectations of No-Influence Contingency | Did Not Contribute Extra	0.37	0.25	1.00	0	304
Contribution Satisfies Social Expectations | Contributed Same	0.66	0.47	1	0	273

"Slope of Response" is the slope coefficient of a regression of what the first-mover expects the second-mover contributes (as a percentage of her endowment) for each decile of the first-mover's contribution in the influence contingency. "Strictly Positive Response" is an indicator variable equal to one if Slope of Response is strictly positive. "Zero Response" and "Strictly Negative Response" are defined analogously. "Conditions for Altruists to Take Back Contributing Extra" is an indicator variable equal to one if the first-mover believes that the second-mover contributes more in the no-influence contingency than he would in the influence contingency if the first-mover contributed the same amount as she did in the no-influence contingency. "Seeking Primacy Increases Contributions | Contributed Same" is an indicator variable equal to one if the first-mover believes the second-mover will contribute strictly less in the no-influence contingency than in the influence contingency (given what she contributed), and is limited to first-movers who contributed the same amount in both contingencies. "Social Expectations of Influence Contingency" measures the first-mover's beliefs over others' beliefs regarding the average contribution (as a percentage of the endowment) by first-movers in the influence contingency. "Social Expectations of No-Influence Contingency" is defined analogously. "Social Expectations of Influence Contingency | Contributed Extra" limits the sample to first-movers who contributed more in the influence contingency. "Contribution Satisfies Social Expectations | Contributed Same" is an indicator variable equal to one if the first-mover's contribution is weakly higher than her social expectations in the influence contingency (that is, how much she believes others believe first-movers contribute in the influence contingency), and is limited to first-movers who contributed the same amount in both contingencies. The rest of the variables are defined analogously.

Table 2: Summary Statistics of Beliefs Relevant to Altruistic Concerns and Expectations Concerns

negative if the coefficient was significantly positive or negative at the 5% level.

The slope of response is 0.051 on average, which means that on average first-movers believe second-movers increase their contribution by 51% of a first-mover's increase in contribution (significantly different from zero at the .001% level with a t-test). There are 83% of first-movers who believe the slope of response is strictly positive, 6% who believe second-movers do not respond to a first-mover's contribution, and 11% who believe second-movers respond negatively.

We have identified a condition on beliefs that is necessary for altruistic first-movers to take back extra, which compares the difference in the levels of contribution in both contingencies (Section 1.2). Among all first-movers, there are 30% whose beliefs satisfy this necessary condition.

Among first-movers who contribute the same in both contingencies, 35% believe the second-mover would weakly contribute more in the influence contingency (given her contribution in that

(as a percentage of the endowment) in the influence contingency after the second-mover sees the first-mover contribute the decile specified by the independent variable. The slope of response was then the coefficient on the first-mover's decile of contribution.

contingency). As discussed in Section 1.2, this is a sufficient and weakly necessary condition for an altruist who contributed the same amount in both contingencies to seek primacy.

Next we can examine the social expectations among first-movers. (We mostly focus on social expectations in the body of the text since the results are qualitatively similar to using expectations, and relegate much of the analysis of expectations to Appendix C.) As mentioned earlier, we are able to create measures of social expectations from our two-part guessing game in which first-movers first guessed how much the average first-mover contributed in the influence and no-influence contingencies, and then guessed the average of what others had guessed in the first part.

Panel B of Table 2 summarizes the statistics for these measures of social expectations. In the aggregate, social expectations in the influence contingency are 58% of the endowment, while they are 37% in the no-influence contingency—significantly different at the 0.001% level with a t-test. These values are similar if we disaggregate beliefs among the first-movers who contributed extra, and among those who did not.

Among first-movers who contributed the same amount in both contingencies, 66% believe their contribution is at least as high as their social expectations in the influence contingency. In Section 1.2 we argued that first-movers with expectations concerns would be more likely to seek primacy if their contributions satisfied social expectations.

2.3 Beliefs Support Altruistic Concerns, Provide Some Evidence for Expectations Concerns

We can now examine whether heterogeneity in beliefs can explain contribution decisions.

We first turn to a test of Hypotheses 5 and 6. We estimate the following OLS regression model:

$$Contribution_{i,c} = \beta InfluenceC_{i,c} + \delta \mathbf{X_i} + \gamma InfluenceC_{i,c} \times \mathbf{X_i} + \phi \mathbf{W_i} + \varepsilon_i \qquad (1)$$

where $Contribution_{i,c}$ is the amount contributed by first-mover i in contingency c, $InfluenceC_{i,c}$ is an indicator variable equal to one if c is a influence contingency, $\mathbf{X}_i$ is a vector of the independent

variables of interest, and $\mathbf{W}_i$ is a vector of control variables, including a constant. Notice that even though a unit is the subject-contingency, the error term ε_i is clustered at the subject level. The coefficients of interest are captured by the vector γ, which indicate whether there is a difference in contributions between influence and no-influence contingencies driven by $\mathbf{X}_i$.

	(1) Contribution	(2) Contribution	(3) Contribution
Influence Contingency	1.195	1.176	1.176
	(1.421)	(1.442)	(1.453)
Social Expectations in No-Influence Contingency	0.0467	0.0593	-0.000581
	(0.0503)	(0.0493)	(0.0496)
Social Expectations in Influence Contingency	0.152**	0.0994	0.0584
	(0.0582)	(0.0567)	(0.0542)
Slope of Response	-1.495**	-1.765**	0.170
	(0.411)	(0.375)	(0.746)
Counterfactual Contribution	0.891**	0.909**	0.917**
	(0.0525)	(0.0495)	(0.0438)
Influence Contingency × Social Expectations in No-Influence Contingency	-0.00623	-0.00651	-0.00651
	(0.0353)	(0.0358)	(0.0361)
Influence Contingency × Social Expectations in Influence Contingency	0.167**	0.167**	0.167**
	(0.0380)	(0.0385)	(0.0388)
Influence Contingency × Slope of Response	0.677**	0.679*	0.679*
	(0.260)	(0.263)	(0.265)
Influence Contingency × Counterfactual Contribution	-0.190**	-0.190**	-0.190**
	(0.0423)	(0.0429)	(0.0432)
Dep Variable Mean	58.31	58.34	58.34
Observations	778	776	776
Clusters	389	388	388
R^2	0.694	0.719	0.745
Controls for Socio-Demographics and Personality	No	Yes	Yes
Controls for Beliefs	No	No	Yes

The table reports the γ coefficient vector of model $Contribution_{i,c} = \beta InfluenceC_{i,c} + \delta \mathbf{X_i} + \gamma InfluenceC_{i,c} \times \mathbf{X_i} + \phi \mathbf{W_i} + \varepsilon_i$. An observation is a subject-contingency (i, c), and standard errors reported in parentheses are clustered at the subject level. The dependent variable is the first-mover's contribution as a percentage of the endowment. "Influence Contingency" is an indicator variable for the influence contingency. "Social Expectation in Influence Contingency" is the first-mover's guess of what others guessed was the average contribution in the influence contingency (as a percentage of the endowment). "Social Expectation in No-Influence Contingency" is the first-mover's guess of what others guessed was the average contribution in the no-influence contingency (as a percentage of the endowment). "Slope of Response" is the slope coefficient of a regression of what the first-mover expects the second-mover contributes (as a percentage of her endowment) for each decile of the first-mover's contribution in the influence contingency. "Counterfactual Contribution" is the percentage of the endowment the first-mover believes the second-mover would contribute in the influence contingency had the first-mover contributed the amount she contributed in the no-influence contingency. The Controls for Socio-Demographics and Personality are age, gender, education, student status, economics major, number of past experiments, knowledge of the objective of Save the Children, a battery of questions about how they would behave in different leadership scenarios, risk aversion and a battery of questions about leadership personalities. The Controls for Beliefs are the first-mover's beliefs of what second-movers contribute (as a percentage of the endowment) in the no-influence contingency, and in the influence contingency for every decile of the endowment first-mover can contribute.

* $p < 0.05$, ** $p < 0.01$

Table 3: The Slope of Response, the Counterfactual Contribution and Social Expectations in the Influence Contingency Predict Contributing Extra

The interaction terms in Table 3 capture the vector γ, the relevant vector for testing Hypotheses

5 and 6. The Table tests for both Hypotheses simultaneously, although we get similar results if we run the tests separately.

From Hypothesis 5, we expect contributions in the influence contingency will be increasing in *Slope of Response* holding *Counterfactual Contribution* fixed, and decreasing in *Counterfactual Contribution* holding *Slope of Response* fixed.[20] We run this test considering no controls (column 1), controlling only for all elicited socio-demographic and personality variables (column 2), and also controlling for first-movers' beliefs over second-movers' contributions in the no-influence contingency, and in the influence contingency for every decile contribution of a first-mover (column 3). We find strong support for these predictions: the coefficients are in the expected direction in all specifications.

Result 5 (Beliefs consistent with altruistic concerns in contributing extra)**.** *The difference between contributions in the influence and no-influence contingencies increases with the slope of response (significant between the 5% and the 1% level) and decreases with the counterfactual contribution (significant at the 1% level).*

From Hypothesis 6, we expect contributions in the no-influence contingency will be increasing in *Social Expectations in No-Influence Contingency*. In all specifications, the coefficient of the interaction between *Social Expectations in No-Influence Contingency* and a dummy for the influence contingency is negative (in the expected direction) but insignificant. Hypothesis 6 also predicts that the contribution in the influence contingency will be increasing in *Social Expectations in Influence Contingency*. In all specifications, the corresponding interaction is positive (in the expected direction) and significant.

Result 6 (Beliefs somewhat consistent with expectations concerns in contributing extra)**.** *The difference between contributions in the influence and no-influence contingencies increases with social*

[20]The definition of these terms was provided in Section 1.2, when discussing Hypothesis 5. Briefly, the slope of response captures how second-mover's contribution in the influence contingency changes with an increase in the first-mover's contribution, and the counterfactual contribution captures what the second-mover would have contributed in the influence contingency had the first-mover acted as in the no-influence contingency.

expectations in the influence contingency (significant at the 1% level) and decreases insignificantly with social expectations in the no-influence contingency (p-values between 0.856 and 0.860).

In order to test for Hypothesis 7 and 8, we estimate the following OLS regression model:

$$Y_i = \beta \mathbf{X}_i + \phi \mathbf{W}_i + \varepsilon_i \tag{2}$$

where the $\mathbf{X}_i$ is a vector of independent variables, and $\mathbf{W}$ is a vector of control variables, including a constant. We consider the same mix of controls as in Table 3.

	(1) Takes Back Contr. Extra	(2) Takes Back Contr. Extra	(3) Takes Back Contr. Extra
Condition for Altruists to Take Back Contributing Extra	0.0953*	0.0794	0.0950*
	(0.0421)	(0.0407)	(0.0467)
Expectations in Influence Contingency	0.000324	0.000268	0.00164
	(0.00114)	(0.00111)	(0.00133)
Expectations in No-Influence Contingency	-0.00264*	-0.00278*	-0.00160
	(0.00109)	(0.00110)	(0.00105)
Social Expectations in Influence Contingency	-0.000407	-0.000218	0.000165
	(0.00111)	(0.00106)	(0.00117)
Social Expectations in No-Influence Contingency	0.000731	0.00100	0.00117
	(0.000964)	(0.000998)	(0.000916)
Dep Variable Mean	0.129	0.126	0.126
Observations	389	388	388
R^2	0.049	0.087	0.135
Controls for Socio-Demographics and Personality	No	Yes	Yes
Controls for Beliefs	No	No	Yes

The table reports the coefficient β of the model $Y_i = \beta X_i + \phi \mathbf{W}_i + \varepsilon_i$, with robust standard errors. The dependent variable is an indicator variable equal to one if the first-mover takes back contributing extra. The independent variable, "Condition for Altruists to Take Back Contributing Extra", is an indicator variable equal to one if the first-mover believes that the second-mover contributes more in the no-influence contingency than he would in the influence contingency if the first-mover contributed the same amount as she did in the no-influence contingency. 'Social Expectations in Influence Contingency' is the first-mover's belief of others' belief of what other first-movers contributed on average in the influence contingency (as a percentage of the endowment). 'Social Expectations in No-Influence Contingency' is defined analogously. 'Expectations in Influence Contingency' is the first-mover's belief of what other first-movers contributed on average in the influence contingency (as a percentage of the endowment). 'Expectations in No-Influence Contingency' is defined analogously. The Controls for Socio-Demographics and Personality are age, gender, education, student status, economics major, number of past experiments, knowledge of the objective of Save the Children, a battery of questions about how they would behave in different leadership scenarios, risk aversion and a battery of questions about leadership personalities. The Controls for Beliefs are the first-mover's beliefs of what second-movers contribute (as a percentage of the endowment) in the no-influence contingency, and in the influence contingency for every decile of the first-mover's endowment.

* $p < 0.05$, ** $p < 0.01$

Table 4: Both the Necessary Condition for Altruists to Take Back Contributing Extra and Expectations in No-Influence Contingency Predict Taking Back Contributing Extra

The coefficient vector of interest is β, reported in Table 4, which tests for Hypotheses 7 and 8 simultaneously. In particular, it tests for two versions of Hypothesis 8 jointly: using expectations

and social expectations. To test for Hypothesis 7, the Table includes an independent variable referred to as *Condition for Altruists to Take Back Contributing Extra*. We jointly tested Hypotheses 7 and 8, as well as both versions of Hypothesis 8, for succinctness—we get similar results if we run the tests separately.

There is support for Hypothesis 7. Although the coefficient drops in and out of significance across the different specifications, neither the coefficient nor the p-value change much across the specifications.

Result 7 (Beliefs consistent with altruistic concerns in taking back contributing extra)**.** *The probability that a first-mover takes back contributing extra is higher when the necessary condition for altruists to take back contributing extra holds (p-values between 0.024 and 0.052).*

The case for Hypothesis 8 is harder to make. The only support comes from the significant coefficients in the first two columns. *Expectations in No-Influence Contingency* is negative and significant as expected in the first two columns, but the coefficient drops by almost half in the last column. Further, the other three variables that were expected to matter are insignificant, and often of the wrong sign.

Result 8 (Beliefs weakly consistent with expectations concerns in taking back contributing extra)**.** *The probability that a first-mover takes back contributing extra increases with expectations in the no-influence contingency (p-value between 0.016 and 0.052). However, the expectation in the influence contingency and social expectations are not predictive of the decision to take back contributing extra.*

Table 5 tests Hypotheses 9 and 10, the hypotheses related to first-movers who contribute the same amount in both contingencies. The empirical model is (2), where the dependent variable is *Seeks Primacy*, an indicator variable equal to one if the first-mover seeks primacy (which, again, means to make the influence contingency more likely to be implemented).

To test for Hypothesis 9, the independent variable is *Seeking Primacy Increases Contributions*, an indicator variable equal to one if the first-mover believes the second-mover contributes more in

	(1) Seeks Primacy	(2) Seeks Primacy	(3) Seeks Primacy	(4) Seeks Primacy	(5) Seeks Primacy	(6) Seeks Primacy
Seeking Primacy Increases Contributions	0.227**	0.179*	0.143			
	(0.0682)	(0.0725)	(0.0759)			
Contribution Satisfies Social Expectations				0.0946	0.0665	0.0337
				(0.0650)	(0.0686)	(0.0727)
Social Expectations in Influence Contingency				0.00392**	0.00371**	0.00158
				(0.00108)	(0.00128)	(0.00162)
Social Expectations in No Influence Contingency				-0.000229	-0.000827	-0.00125
				(0.00118)	(0.00137)	(0.00145)
Dep Variable Mean	0.612	0.612	0.612	0.612	0.612	0.612
Observations	273	273	273	273	273	273
R^2	0.041	0.122	0.193	0.074	0.136	0.185
Controls for Socio-Demographics and Personality	No	Yes	Yes	No	Yes	Yes
Controls for Beliefs	No	No	Yes	No	No	Yes

The table reports the coefficient β of the model $Y_i = \beta X_i + \phi \mathbf{W}_i + \varepsilon_i$, restricted to first-movers who contributed the same amount in both contingencies, and with robust standard errors. The dependent variable is an indicator variable equal to one if the first-mover seeks primacy (that is, implements the influence contingency). The independent variable for the first three columns is "Seeking Primacy Increases Contributions", an indicator variable equal to one if the first-mover believes the second-mover will contribute strictly less in the no-influence contingency than in the influence contingency (given what she contributed). The independent variable in the last three columns is "Contribution Satisfies Social Expectations", an indicator variable equal to one if the first-mover's contribution is weakly higher than her social expectations in the influence contingency (that is, how much she believes others believe first-movers contribute in the influence contingency). The regressions in the last three columns include controls for social expectations in the influence and no-influence contingencies. The Controls for Socio-Demographics and Personality are age, gender, education, student status, economics major, number of past experiments, knowledge of the objective of Save the Children, a battery of questions about how they would behave in different leadership scenarios, risk aversion and a battery of questions about leadership personalities. The Controls for Beliefs are the first-mover's beliefs of what second-movers contribute (as a percentage of the endowment) in the no-influence contingency, and in the influence contingency for every decile of the endowment first-mover can contribute.

* $p < 0.05$, ** $p < 0.01$

Table 5: First-Movers Who Contributed Same Amount in Both Contingencies Do Seek Primacy When They Expect It Will Raise Contributions, and it Does Not Matter Whether They Believe Their Contribution Satisfies Social Expectations

the no-influence contingency than in the influence contingency (given her contribution). The three columns use the same mix of controls we have used throughout.

The evidence is somewhat supportive of the altruistic concerns. The first two columns are significant, and the third loses significance. This drop in significance is somewhat worrisome because the value of the coefficient steadily declines as we add coefficients, but neither the change in the coefficient nor in the p-value are very sharp.

Result 9 (Beliefs consistent with altruistic concerns in seeking primacy)**.** *There is some evidence that the probability that a first-mover seeks primacy increases when the first-mover believes the second-mover will increase her contribution in the influence contingency (p-values range between 0.001 and 0.061).*

To test for Hypothesis 10, we use the independent variable *Contribution Satisfies Social Expectations*, an indicator variable equal to one if the first-mover's contribution in the influence contingency is larger than her social expectation in that contingency. The results are shown in the last three columns of Table 5. Again, the three columns are the same mix of controls as before, in addition to controlling for social expectations. The relevant coefficient in each of the columns is insignificant.

Result 10 (Beliefs not consistent with expectations concerns in seeking primacy)**.** *First-movers who contribute the same amount in both contingencies are not more likely to seek primacy if their contributions in the influence contingency are at least as high as their social expectations in that contingency (p-values range between 0.224 and 0.434).*

3 Discussion

If one were to focus solely on the behavioral results, one may conclude that subjects' leadership is driven by expectations concerns. We found that only 22% of our 394 first-movers contributed extra, and among those who contributed extra, 58% took back contributing extra. However, by taking beliefs into account, we are in fact led to conclude that the evidence is strong in favor of altruistic concerns, and weaker for expectations concerns.

Our conclusion is drawn from several tests.

First, we considered the difference in contributions between the contingency in which the first-mover can influence the second-mover and the contingency in which she cannot. As per altruistic concerns, the expected slope and the expected level of second-movers' response to a first-mover's contribution explains the difference in contributions. As per expectations concerns, the difference can be explained by guesses of what first-movers do, and guesses over others' guesses of what first-movers do.

Second, we considered whether elicited beliefs can explain the decision to take back contributing extra. We showed that a necessary belief condition for altruists to take back contributing extra

does indeed predict taking back contributing extra. We found some evidence that expectations concerns can explain the decision to take back contributing extra as well: the expectations of what first-movers do when they cannot influence the second-mover predicts the decision to take back contributing extra, although social expectations are not predictive.

Third, elicited beliefs are broadly supportive of altruistic concerns explaining seeking primacy, but not expectations concerns. We found some evidence that first-movers who contribute the same are more likely to seek primacy the more they contribute. More directly in line with altruistic concerns, first-movers who contribute the same are more likely to seek primacy if they believe the impact of their contribution is higher. To test for expectations concerns, we tested whether first-movers were more likely to seek primacy if their contributions fulfilled expectations or social expectations. We did not find supportive evidence.

It is instructive to contrast our findings with those of Karlan and McConnell (2014). As explained in the introduction, their design is closest to ours. In their experiment they randomly assign subjects to a contingency in which their contributions, publicly displayed, can have an influence on subsequent contributions by others. They compare contributions in this contingency with those made by subjects under the understanding that their contributions will only be made public after others have also made their contribution decisions. They find no evidence for contributing extra. However, they do find that contributions in these two contingencies are higher than private contributions. Based upon this they conclude that the promise of public recognition for contributing increases charitable contributions simply because people want to improve their social image, and not by an altruistic desire to influence others.

In contrast to their work, we did find evidence for contributing extra, and for an altruistic motivation. Our studies differ in some key aspects which may explain the contrasting findings. First, first-movers in their design knew others would see everyone's first-mover contribution. This creates a free-rider problem, since with more first-movers, a single first-mover's contribution becomes a smaller proportion of the contributions second-movers observe. Second, first-mover's names were posted alongside their contribution, so social-image concerns may have crowded out

an already-diluted motivation to lead. Third, we conduct a within-subjects comparison of decisions that can and cannot influence others, while Karlan and McConnell (2014) conduct a between-subjects comparison. A between-subjects design may make it more difficult to find an effect. Our within-subjects design further allowed us to test for each first-mover's motivation for contributing extra.

3.1 Difference Between Mexico and US Samples

Although the experimental results from the Mexican and US samples were similar in a lot of ways, there were some differences. Online Appendix D presents the results for the Mexican and US samples separately.

Mexican subjects generally contributed higher amounts than their American counterparts, which we believe was due to the composition of the samples—the Mexican sample consisted only of college students from private schools, whereas the US sample consisted of a broader range of levels of education, age and socio-economic status. The *prima facie* evidence was more supportive of expectations concerns in the US sample, as 80% of subjects who contributed extra took back contributing extra, compared to 52% in the Mexico sample. The extreme value for the US sample can be explained by the smaller size of the US sample, where only 20 subjects total contributed extra (out of 93 first-movers).

We had analyzed the US data by itself originally, since we ran the first experiment much earlier than the second. In our original analysis, we had concluded that expectations concerns were driving leadership behavior. In the US sample, the predictions regarding altruistic concerns were insignificant, and we had found that social expectations could partly predict contribution decisions and the decision to seek primacy. However, the US sample size was small, which raised concerns about the robustness of the findings.[21] We therefore collected the extra data. The data from Mexico, while qualitatively similar to that of the US, strengthened the patterns in support of altruistic concerns that were present but insignificant in the US data, and weakened the patterns in support

[21] We thank the editorial team for nudging us to raise more data!

of expectations concerns.

It is worthwhile at this point to discuss our decision to present the pooled results. Quite simply, we had no *ex ante* reason to expect different results between the Mexico and US samples. Indeed, as can be verified from our pre-analysis plan (AEARCTR-0009191, developed between the first and second experiments), we had not formulated any hypotheses about a different responses in the samples. Given that the motivation to run the second experiment was to increase the sample size, we found it most natural to pool the results.

Finally, although our view is that the difference in results across samples is explained by the difference in sample size, it is worth noting that there are some explanations for the differences that we cannot fully rule out. First, there may be some differences in the samples themselves that may explain the different results, such as cultural differences, or differences in sociodemographic characteristics. If this is the case, the pooled results should be interpreted as estimating the aggregate patterns across these two disparate groups. Second, the instructions included some changes we believe are small, but that could have systematically affected decisions (full details of both designs can be found in Online Appendix B). For example, in the US sample, we explained the influence and no-influence contingencies, then asked them to choose their contribution in one contingency, and then in the other (the order of the decisions was randomized). In the Mexico sample, we explained the contingencies, and then asked them to choose both their contributions. There is an experimental literature that shows that individuals' choices may be quite different when choices are made sequentially versus simultaneously (e.g. Bohnet et al., 2016). The inconsistency in choices seems driven by a shift in the focus placed on different aspects of the decision, due to the different elicitation methods (Sunstein, 2018). However, we believe this mechanism is not very important in our setting, since in both cases the information about both decisions was presented before any decision was made.

4 Conclusion

Most of us will probably encounter several situations in any given week in which we have an opportunity to lead others by example. We might be asked explicitly to step forward, at the office for example, to support a cause. Or we may just engage with others in contexts in which we are aware that our decision to act in some pro-social way (e.g., stopping to let another driver pull out of a parking spot, waiting and holding a heavy door open for the next person entering a building) can be observed and imitated by others.

In this paper we provide experimental evidence to assess whether pro-social leading is motivated by altruistic concerns or by a desire to conform with expectations. The evidence from our experiment provides some support for expectations concerns, but provides strongest support for altruistic concerns. There is a sizable minority who act more pro-socially when they are in a position to influence others to do the same, and they want to publicize their action only if they believe this extra pro-sociality will indeed lead others to act more pro-socially. The first-movers who do not contribute extra when in a position to influence others also act altruistically, in the sense that they too are more likely to put themselves in a position to influence others when they believe it will lead others to act more pro-socially.

This paper is the first, we believe, to document that individuals act more pro-socially when they can influence others and to isolate the motivation for doing so. Fundraisers should harness the good will of these leaders by providing opportunities to publicize their contributions. Complementary research should try to find observable characteristics of these first-movers, in order to be able to recognize who to provide these opportunities to.

On the other hand, our work does not fully rule out expectations concerns. Further work could complement ours by exogenously assigning the expectations or the social expectations of being in a position to influence others. Setting expectations or social expectations may yet prove to be a useful part of the toolkit of a social planner who want so increase leadership behavior.

References

Arbak, Emrah and Marie-Claire Villeval, "Voluntary leadership: motivation and influence," *Social Choice and Welfare*, 2013, pp. 1–28.

Bénabou, Roland and Jean Tirole, "Incentives and prosocial behavior," *American economic review*, 2006, *96* (5), 1652–1678.

Bohnet, Iris, Alexandra Van Geen, and Max Bazerman, "When performance trumps gender bias: Joint vs. separate evaluation," *Management Science*, 2016, *62* (5), 1225–1234.

Bracha, Anat, Michael Menietti, and Lise Vesterlund, "Seeds to succeed?: Sequential giving to public projects," *Journal of Public Economics*, 2011, *95* (5), 416–427.

Brandts, Jordi, Christina Rott, and Carles Solà, "Not just like starting over-Leadership and revivification of cooperation in groups," *Experimental economics*, 2016, *19* (4), 792–818.

Buser, Thomas, Eva Ranehill, and Roel Van Veldhuizen, "Gender differences in willingness to compete: The role of public observability," *Journal of Economic Psychology*, 2021, *83*, 102366.

Charness, Gary and Martin Dufwenberg, "Promises and partnership," *Econometrica*, 2006, *74* (6), 1579–1601.

Dana, Jason, Daylian M Cain, and Robyn M Dawes, "What you don't know won't hurt me: Costly (but quiet) exit in dictator games," *Organizational Behavior and human decision Processes*, 2006, *100* (2), 193–201.

DellaVigna, Stefano, John A List, and Ulrike Malmendier, "Testing for altruism and social pressure in charitable giving," *The quarterly journal of economics*, 2012, *127* (1), 1–56.

d'Adda, Giovanna, "Relative social status and conformism: Experimental evidence on local public good contributions," *Economics Letters*, 2017, *157*, 31–35.

Ertac, Seda, Mert Gumren, and Mehmet Y Gurdal, "Demand for decision autonomy and the desire to avoid responsibility in risky environments: Experimental evidence," *Journal of Economic Psychology*, 2020, *77*, 102200.

Fernández-Duque, Mauricio and Michael Hiscox, "Audience effects on anonymous pro-social followership," *Economics Letters*, 2022, *212*, 110268.

Fernández-Duque, Mauricio and Michael J Hiscox, "Altruistic or Expected Leadership," *Mendeley Data*, 2022, *V1* (doi:10.17632/8293mvy6dy.1).

Frackenpohl, Gerrit, Adrian Hillenbrand, and Sebastian Kube, "Leadership effectiveness and institutional frames," *Experimental Economics*, 2016, *19* (4), 842–863.

Gächter, Simon, Daniele Nosenzo, Elke Renner, and Martin Sefton, "Who Makes a Good Leader? Cooperativeness, Optimism, and Leading-By-Example," *Economic Inquiry*, 2012, *50* (4), 953–967.

Güth, Werner, M Vittoria Levati, Matthias Sutter, and Eline Van Der Heijden, "Leading by example with and without exclusion power in voluntary contribution experiments," *Journal of Public Economics*, 2007, *91* (5), 1023–1042.

Jack, B Kelsey and María P Recalde, "Leadership and the voluntary provision of public goods: Field evidence from Bolivia," *Journal of Public Economics*, 2015, *122*, 80–93.

Karlan, Dean and John A List, "How can Bill and Melinda Gates increase other people's donations to fund public goods?," *National Bureau of Economic Research*, 2018, *Working Paper No. 17954 (Revised Version).*

_ **and Margaret A McConnell**, "Hey look at me: The effect of giving circles on giving," *Journal of Economic Behavior & Organization*, 2014, *106*, 402–412.

Komai, Mana and Philip J Grossman, "Leadership and group size: An experiment," *Economics Letters*, 2009, *105* (1), 20–22.

Krupka, Erin L and Roberto A Weber, "Identifying social norms using coordination games: Why does dictator game sharing vary?," *Journal of the European Economic Association*, 2013, *11* (3), 495–524.

Levy, David M, Kail Padgitt, Sandra J Peart, Daniel Houser, and Erte Xiao, "Leadership, cheap talk and really cheap talk," *Journal of Economic Behavior & Organization*, 2011, *77* (1), 40–52.

Meyer, Christian Johannes and Egon Tripodi, "Image concerns in pledges to give blood: Evidence from a field experiment," *Journal of Economic Psychology*, 2021, *87*, 102434.

Potters, Jan, Martin Sefton, and Lise Vesterlund, "After you – endogenous sequencing in voluntary contribution games," *Journal of public Economics*, 2005, *89* (8), 1399–1419.

Ren, Yufei, Lin Xiu, and Amy Hietapelto, "Dare to Ask in Front of Others? Women Initiating Salary Negotiations," *Journal of Economic Psychology*, 2022, p. 102550.

Rivas, M Fernanda and Matthias Sutter, "The benefits of voluntary leadership in experimental public goods games," *Economics Letters*, 2011, *112* (2), 176–178.

Salazar, Miguel, Daniel Joel Shaw, Kristína Czekóová, Rostislav Staněk, and Milan Brázdil, "The role of generalised reciprocity and reciprocal tendencies in the emergence of cooperative group norms," *Journal of Economic Psychology*, 2022, *90*, 102520.

Sunstein, Cass R, "On preferring A to B, while also preferring B to A," *Rationality and Society*, 2018, *30* (3), 305–331.

Varian, Hal R, "Sequential contributions to public goods," *Journal of Public Economics*, 1994, *53* (2), 165–186.