Machiavelli Preferences Without Blame: Delegating Selfish vs. Generous Decisions in Dictator Games *

Glynis Gawn † Robert Innes ‡

September 2020

Abstract

Does the impulse to delegate a decision only arise when delegation mimics a selfish decision vs. a generous one? We address this question using a dictator experiment with two possible payment allocations and an option to delegate the payment decision to another player. Two delegation treatments are considered, one in which the delegation option is payoff-equivalent to a direct choice of a "selfish" allocation (better for the dictator, worse for the receiver) and another where it is equivalent to the direct choice of generous / equal payments. Dictators exhibit a significantly greater propensity to delegate in the selfish delegation treatment than in the generous delegation treatment. Results are consistent with "Machiavelli preferences" that only favor delegation when it promotes self-interested / other-harming outcomes.

Keywords: Delegation, Moral Preference, Dictator Game

JEL Classifications: D03, D91

*The authors thank two anonymous referees and the Editor, Professor Urs Fischbacher, for meticulous and helpful comments on a prior version of the paper. The usual disclaimer applies. We also thank our kind and supportive colleagues at UC Merced, particularly including Jesus Sandoval-Hernandez, Jason Lee, Justin Hicks, Steve Seltzer and Mark Harris for graciously enabling their students to participate in our experiments.

†Economics, School of Social Sciences, Humanities & Arts, University of California, Merced. Email: ggawn@ucmerced.edu.

‡Corresponding Author: Economics, School of Social Sciences, Humanities & Arts, University of California, Merced, 5200 N. Lake Rd., Merced, CA 95343. Email: rinnes@ucmerced.edu. Phone: (209) 228-7742.

Machiavelli (1532): "Princes should delegate to others the enactment of unpopular measures and keep in their own hands the means of winning favors"

1 Introduction

Delegation to agents is pervasive in economic interactions. Production decisions can be outsourced to delegates. Intermediaries can be hired as delegates in employment and professional services. Delegated decisions often have moral overtones that are potentially central to the delegation decision. For example, corrupt agents may be hired to do the bidding of companies who seek to distance themselves from the agents' decisions (Cullen, 2008). Considering potential moral implications, why might delegation occur and what are its effects?

A large literature in industrial organization identifies explanations for agency relationships driven by agents' specialized knowledge and/or strategic benefits of delegation (e.g., see Bolton & Dewatripont, 2005). An experimental literature shows that delegation can shift the blame for outcomes harmful to others (Bartling & Fischbacher, 2012; Coffman, 2010; Fershtman & Gneezy, 2001; Oexl & Grossman, 2013). Even absent prospects for blame, principal-agent relationships can attenuate preferences for helping others at one's own expense (Hamman, Loewenstein & Weber, 2010) by creating moral wiggle room (Dana, Weber & Kuang, 2007). Reducing other-regarding preferences – and thereby promoting more selfish outcomes – can motivate a delegated vs. direct decision. However, recent research documents a prevailing preference for retaining decision rights rather than delegating them (Feher, Herz & Wilkening, 2013; Bartling, Fehr & Herz, 2014; Owens, Grossman & Fackler, 2014).

An added consideration is suggested by Machiavelli's (1532) dictum – that principals should delegate decisions that are harmful to others and not delegate decisions that are beneficial to others. The dictum is motivated by a desire to avoid blame (by delegating) or to attract credit (by retaining decision responsibility), as stressed and studied in Bartling & Fischbacher (2012). However, even without any scope for blame or credit, individuals may have a preference for a delegated (vs. direct) decision when another person is harmed by the decision, and vice versa if another person is benefited – what we call "Machiavelli preferences." If such preferences prevail, delegation will occur primarily to advance self-interested outcomes: In order to create a market for delegation services, delegates will promote the other-harming behavior that attracts clients, precisely as found in Hamman, Loewenstein & Weber (2010).

Our objective in this paper is to test for Machiavelli-type preferences in experimental dictator games that contain a delegation option but with no scope for blame or credit from others. In a dictator game, each dictator chooses an allocation of money between him or herself and another person. In our case, there are two potential allocations, one of which is more selfish (giving the dictator more money) and the other more generous (giving the other person an equal amount of money). The dictator can opt to implement the decision of another player (a random delegate) or choose directly between the two allocations. The delegation option is either equivalent to a selfish decision (the DEL-SELF treatment) or equivalent to a generous decision (the DEL-GEN treatment), where the equivalence is determined by a probabilistic correspondence in payments to the players. The question is whether dictators delegate more often when a delegated choice mimics a selfish decision vs. when it mimics a generous one.

2 Related Literature and the Role of Responsibility

Because we focus on delegation when there is no scope for punishment or blame, we sidestep the literature on how delegation alters blame and resulting incentives for generosity, notably including Fershtman & Gneezy (2001), Coffman (2011), Bartling & Fischbacher (2012), and Oexl & Grossman (2013). Absent scope for blame, results from the literature most closely related to the present study might be organized in three sets:

Result 1. By distancing a person from a decision that can be relatively more selfish or relatively more virtuous, *delegation can attenuate a preference for a generous (vs. selfish) outcome*. There are a variety of potential mechanisms for creating *moral wiggle room* that excuses selfish behavior, as first identified by Dana, Weber & Kuang (2007).[1] Most rele-

[1]Examples include group decision-making (Charness & Sutter, 2012), ignorance (Grossman, 2015), hiding

vant here is the key Hamman, Loewenstein & Weber (2010) study on how principal-agent relationships promote selfish choices, whether by agents acting more selfishly on behalf of principals than the principals would do if choosing for themselves, or by the principals choosing to switch away from agents who are generous on their behalf. Their results suggest a norm of selfish (principal-favoring) conduct by agents. We interpret this norm as evidence that delegation creates moral wiggle room.[2]

Result 2. Separate from preferences over allocations, there tends to be a net preference for retaining decision rights; *direct decisions are preferred over equivalent delegated ones.* Bartling, Fehr & Herz (2014), Fehr, Herz & Wilkening (2013), and Owens, Grossman & Fackler (2014) identify the positive value subjects place on autonomy, freedom, and self-decision. Charness et al. (2012) identify benefits of ceding control to workers in a firm, consistent with the hidden costs of control found in Falk & Kosfeld (2006).

Result 3. Subjects are often willing to pay to avoid having to make a morally fraught allocation decision (Dana, Cain & Dawes, 2006). Della Vigna, List & Malmendier (2012) find that potential donors avoid solicitors in order to opt out of a donation decision. Lazear, Malmendier & Weber (2012) find that those opting out of dictator games are almost entirely dictators who, due to shame or guilt, would be generous if forced to make an allocation decision. Delegation is a potential strategy to opt out of a dictator decision. In an earlier paper (Gawn & Innes, 2019a), we find that *dictators choosing to delegate are almost entirely those who would otherwise be generous.*

The literature implies a tension between attenuation benefits of delegation (Result 1) and costs of delegation in lost decision rights (Result 2) that leads some dictators to delegate and others not. With "more" of a social preference to attenuate, dictators who would otherwise be generous enjoy larger net benefits of delegation relative to those who would otherwise be

behind nature (Andreoni & Bernheim, 2009), ambiguity (Haisley & Weber, 2010) risk (Exley, 2016), a pre-existing pro-social reputation (Exley, 2018), use of advisers (Coffman & Gotthard-Real, 2019), and less-than-perfect charity performance metrics (Exley, 2020).

[2]See also related evidence that delegation reduces lie aversion (Gawn & Innes, 2019b), charitable giving (Coffman, 2017), and corruption-avoidance (Drugov, Hamman & Serra, 2014). A recent paper by Collins, Hamman & Lightle (2018) shows that commoditizing the decision right dramatically reduces the social preference for generosity.

selfish (Result 3).

In prior literature, when a dictator who would otherwise be generous decides to delegate, the delegation increases the likelihood of the selfish outcome. In principle, this property does not affect the logic of either Result 1 or Result 2. However, if there are Machiavelli preferences, this property is crucial because a delegated decision is *bad news* for the receiver. What if, instead, delegation by the generous has no effect on the likelihood of the selfish outcome and, for the selfish, delegation increases the likelihood of the generous outcome? This is the generous delegation treatment that we consider in this paper. In this treatment, the decision subject to delegation is *good news* for the receiver. Although the same tension persists between the cost of delegation in sacrificing decision rights (Result 2) and the benefit in attenuating virtuous preferences (Result 1), Machiavelli preferences imply a completely different decision environment; the net benefit of retaining decision rights is greater *because the decision is good news*. These are the preference effects that we investigate in this paper, how the choice to delegate is affected by whether the decision is good or bad news. Because Result 1 attenuation effects (c.f., Hamman, Loewenstein & Weber, 2010) equally motivate delegation in both of our treatments, our experiment permits us to pinpoint an alternate Machiavelli preference mechanism for delegation decisions.

Responsibility attribution is central to Machiavelli preferences. Whether others are attributing responsibility for choices (as in Bartling & Fischbacher, 2012; Andreoni & Bernheim, 2009) or individual decision-makers are ascribing responsibility to themselves, delegation can reduce the inference of responsibility for outcomes. In our case, it is the internal mechanisms at work, those highlighted in recent theories of self-signaling and self-inference about social identity (Grossman & van der Weele, 2017; Gneezy, Kajackaite & Sobel, 2018; Dufwenberg & Dufwenberg, 2018). If delegation distances oneself from a virtuous decision (as in our generous delegation treatment), then self-image of a virtuous social identity is *reduced* by a delegated (vs. direct) decision. In contrast, if delegation distances oneself from a selfish decision (as in our selfish delegation treatment), then self-image is *improved* by the attenuation of responsibility achieved with delegation.

Steffel, Williams & Perrmann-Graham (2016) also stress the importance of responsibility and blame for delegation behavior. Steffel et al. (2016) study hypothetical decisions that could affect one person, either their survey respondent or someone else. They find that respondents prefer not to be a delegate / decider when the choice alternatives are unattractive (vs. attractive) and affect another person (vs. the respondent herself). Studies of delegation in experimental dictator games (like ours and related economics literature) instead consider delegation to a third player when there are between-person tradeoffs.

3 The Experiment

We consider a dictator experiment in which there are two payment possibilities for a dictator and his/her anonymous and passive receiver. Baseline payments (the "generous" allocation) are \$3 to each of the two players. Alternatively, a "selfish" payment change can be made, adding \$2 to the dictator's payment, subtracting \$2 from the receiver's payment, and thereby producing a (\$5,\$1) split. The dictator makes a decision that is used to determine which payment allocation is implemented. We use a generous benchmark (the baseline equal split) because prior work documents a preference for delegation primarily by the otherwise-generous (Gawn & Innes, 2019a); the generous benchmark, by potentially promoting otherwise-generous behavior, offers more scope for variation in delegation behavior across treatments.

The experiment is designed to meet three objectives with the simplest possible design and instructions: (i) provide dictators with a choice between a direct decision on payments and a delegated decision that implements the allocation chosen by another dictator, (ii) control for payoff effects of the direct-vs.-delegate choice so that the choice reflects only relative preferences for delegation, not preferences over payments, and (iii) compare treatments in which a direct "generous" – vs., respectively, "selfish" – allocation choice mimics the payoff profile of the delegation option. The comparison in (iii) enables inferences about how the preference for delegation, vs. a payoff-equivalent direct decision, is affected by whether the equivalent decision is favorable to the other player (good news, when the payoff-equivalent decision is generous) vs. unfavorable (bad news, when the equivalent decision is selfish).

The starting point for the experiment's design is a set of control treatments, without punishment/blame, presented by Bartling & Fischbacher (2012). Their controls satisfy property (i) above, with two possible payment profiles (generous and selfish) and two treatments, one with a delegation option and another with no delegation option. Without delegation, the dictator chooses a payment allocation (generous or selfish) and the chosen allocation is implemented. For a dictator who prefers the "selfish" payments, a potential benefit of delegating — by implementing the decision of another player — is the implied reduction in self-ascribed responsibility for a selfish choice, our interest in the present paper. However, there are two offsetting costs: the lost benefit of a direct decision due to a preference for retaining decision rights (c.f., Bartling et al., 2014) and a reduced probability of the preferred selfish allocation.

Gawn and Innes (2019a) control for the second effect – satisfying property (ii) – by structuring direct selfish payoffs that mimic payoffs under delegation. This is done by implementing a direct selfish decision with the same probability as when the allocation decision is delegated. As a result, a direct selfish decision is payoff-equivalent to a delegated decision. We follow this approach in the present experiment. The difference is that we conduct a third treatment, to satisfy property (iii), in which delegation is payoff-equivalent to direct choice of the *generous* option (DEL-GEN), rather than the selfish option (DEL-SELF).

The design is as follows: Dictators are in the role of player A and receivers (in a different classroom) are in the role of player C. Players A and C are randomly and anonymously matched with two other students, players B and D, who are in a similar situation but in different classrooms than A and C. Players B and D also get a base payment of $3 each, and player B decides whether to make the payment change, so that B earns $2 more and D earns $2 less. The experimenter calculates the percentage of all "player B's" who choose to make the payment change (vs. not) in the matched player B's classroom. We call this percentage Q. The B players participate in a parallel dictator game with no delegation option – the DIC experiment described in the text below.

In the first experimental treatment (DEL-SELF), player A has three options:

Option 1: Choose to MAKE the payment change directly ($2 more for player A and $2 less for player C). In this case, the payment change is implemented for players A and C with probability Q.

Option 2: DO NOT make the payment change. In this case, the initial ($3 each) allocation is implemented with probability one.

Option 3: Implement PLAYER B's decision. In this case, if the matched player B chooses to make the payment change for himself/herself and his/her player D, then the payment change is made for players A and C as well. Likewise, if the matched player B chooses not to make the change, then player A and player C receive the initial allocation of $3 each.

Options 1 and 3 produce an identical payment distribution for players A and C that includes the selfish ($5,$1) allocation with probability Q and the initial allocation of $3 each with probability (1-Q). Here, delegation is completely equivalent, in payoff terms, to a direct selfish decision.

In the second treatment (DEL-GEN), player A again has three options. The only difference is that option 1 implements the selfish allocation with probability one (rather than probability Q), and option 2 implements the initial equal-split allocation with probability (1-Q) (rather than probability one). Here, the direct *generous* option 2 and the delegation option 3 produce an identical payment distribution for players A and C; both produce the payment change ($5/$1) allocation with probability Q and the baseline generous payments ($3 each) with probability (1-Q). Delegation in this case is completely equivalent, in payoff terms, to a direct generous decision.

A third treatment (DIC) is a pure dictator game – the same game in which the B players participate. This treatment follows the DEL-SELF design except there is no delegation option 3; dictators choose between option 1 (changing payments with probability Q) and option 2 (keeping the initial equal split with probability one).

Table 1 summarizes the payment probabilities under the different options for the two main treatments and the DIC controls. We distinquish between: (1) *allocations* that are either "selfish" ($5,$1) or "generous" ($3 each); *options* that implement the allocations with defined probabilities and are by direct choice or delegated, either "direct selfish" (option 1), "direct generous" (option 2), or "delegation" (option 3); and *treatments* that vary the option probabilities so that the delegation option is payoff-equivalent to either a "direct selfish" decision (DEL-SELF) or a "direct generous" decision (DEL-GEN).

All three treatments are implemented in each classroom of the experiment. Matched B players (and associated Q probabilities) are taken from a prior experiment that follows the DIC design (Gawn & Innes, 2019a).[3] Experimental instructions make clear that the Q probability is determined from a dictator game in a *different* classroom — the same other classroom from which the matched B player is selected for the two main (DEL-SELF and DEL-GEN) treatments. There are no cross-treatment differences in how Q is determined; for all treatments in any given classroom (DEL-SELF, DEL-GEN, and DIC), the Q probability is determined from prior DIC behavior in exactly the same other classroom. The Q probability for A players is determined by the proportion of B players *choosing* the "direct selfish" *option*, not the proportion of "selfish" *allocations* actually implemented for the player B's. [4] To avoid biasing perceived norms, Q probabilities are not announced to the A players, but rather illustrated in identical instructions across treatments that read: "For example, if half of the player B's choose to make the payment change, and half do not, then Q will be 50 percent." The 50 percent illustration is meant both to avoid a norm in any one direction and to illustrate that both option decisions are made by other dictators.

The B players are not told anything about matched dictators in other classrooms and do not know that their choices have the potential to affect allocations of players A and C. Each

[3]Mean Q from Gawn & Innes (2019a) is 42.7 percent. Q varies across the five B player classrooms from a low of 29 percent to a high of 56 percent, with a lower division average (three courses) of 40.2 percent and an upper division average (two courses) of 50 percent.

[4]The implementation of allocations for the B players depends upon a "parallel Q" for them. However, the player B inferences of their "parallel Q's" are not relevant to A player decisions. All that matters for A players are player B option choices, which depend only on player B allocation preferences. The A player instructions describe this structure with a simple illustration of how Q is calculated (see instruction language in text).

player A is told that their decision does not affect players B and D in any way.

Other design features include the following. To avoid any experimenter demand effects, dictators are not told anything about the purpose of the experiment or its design. The option order is randomly varied in the experiment. Audience effects are avoided by informing A and B dictators that their C and D receivers (respectively) know nothing about payoffs or decisions. For example, the A dictators know that their receivers are not told whether a decision has been delegated or, indeed, what allocation alternatives are available and which one has been selected.[5] The matched receivers are anonymous players in other classrooms who only know that someone else makes a decision that determines payments.[6]

Sample experimental instructions are provided in the Appendix. Instructions are written to be as simple and clear as possible, on a single page with the dictator options (and probabilities) highlighted immediately before / above the decision section where the dictator circles his/her chosen option.

In this experiment, we interpret "delegation" to mean the implementation of a specific other person's decision (our B delegate), as in (for example) Bartling & Fischbacher (2012). Following Gawn & Innes (2019a), the direct choice that we argue is mimicking the delegation choice is one that has exactly the same probability distribution for payments as does the delegation option, but is implemented by a random device, not a specific matched person (player B).

This distinction is similar to that made in a series of papers by Amione & Houser (2011, 2012, 2013) where trust behavior is compared between treatments that implement the decision, respectively, of a specific (randomly selected) matched trustee and the decision outcome selected by a random device that chooses with the same probabilities as does the random matched trustee. Amione & Houser (2012) find that this subtle difference in structure has a profound impact on trust behavior; the specific matched trustee treatment permits betrayal

[5]The instructions to A players (similar to B players) state: "You know what the options are, but your player C DOES NOT. Player C only knows that someone else will make a choice that determines payments to both of you."

[6]The receivers also participate in an unrelated experiment. As a result, final receiver payments do not indicate the precise payment from the present experiment.

that deters trust (Bohnet & Zeckhauser, 2004). For our purposes, the important lesson from Amione & Houser (2011, 2012, 2013) is that the implementation of a decision via a random device vs. a random person can be important, although the reasons are different here. Unlike these studies, there is no potential for betrayal in our framework as there are no responses from other players. Instead, the specific matched player is a delegate and we are interested in how the choice to delegate (vs. choose directly via a random device) is affected by a potential desire to cede self-ascribed responsibility to a specific person. Here, "moral wiggle room" (Dana et al., 2007) is not created solely by complexity or distance, but rather by distancing via another person who can bear responsibility.[7]

4 Hypotheses

The main hypothesis is the Machiavelli conjecture:

Hypothesis 1 (H1) (Machiavelli). The propensity to delegate is higher when delegation is selfish (treatment DEL-SELF) than when it is generous (treatment DEL-GEN).

Three checks speak specifically to the *Machiavelli preference* explanation for H1 – viz, that there is a distinct preference for delegating selfish (vs. generous) decisions. First, perhaps there could be more delegation under the selfish (DEL-SELF) treatment because more people prefer the selfish allocation to begin with. The following hypothesis addresses this conjecture by focusing on *shares* of subjects.

Hypothesis 2 (H2). The propensity to delegate among those preferring selfish in DEL-SELF (either directly or with a delegated decision) is higher than the propensity to delegate for those preferring generous in DEL-GEN (again either directly or with delegation).

H2, if true, implies that H1 could not simply be due to a preponderance of selfish-preferring dictators (the potential delegators in DEL-SELF) vs. generous-preferring dictators (the potential delegators in DEL-GEN).

Second, more subjects might delegate in DEL-SELF due to a stronger preference for

[7] We are indebted to an anonymous referee for articulating this point.

delegation among the selfish. Evidence for or against this *preference distribution effect* can be obtained by measuring *who* is delegating proportionately more in DEL-SELF – those who would otherwise be selfish or, alternately, generous. The following hypothesis, if valid, rules out this explanation for H1:

Hypothesis 3 (H3). In DEL-SELF, the share of the otherwise-selfish who prefer to delegate is not larger than the share of the otherwise-generous who prefer to delegate.

Third, we construct a *strong test* of Machiavelli by examining an implication of treatment-invariant preferences for delegation that *must* hold if there is *not* a Machiavelli preference effect. The test is strong in the sense that the implication may also hold under Machiavelli but can be violated if the Machiavelli preference effect is sufficiently large. A complete derivation of the test / implication is provided in the Appendix.

Hypothesis 4 (H4) (NOT Machiavelli). Delegation preferences are treatment-invariant, implying a (weakly) higher probability of a direct generous (option 2) choice under DEL-SELF than DEL-GEN.

The null hypothesis for H4 is a zero difference in the probability of the direct generous choice (DEL-SELF minus DEL-GEN) vs. the *one-sided* alternative of a negative difference, consistent with Machiavelli. H4 results from the following logic.

Preferences can be defined along two dimensions (see Figure 1): (1) preferences over allocations U (some preferring selfish and some preferring generous) and (2) preferences over delegation d (some preferring direct decisions, as implied by the decision rights literature, and others preferring delegation). If there is no Machiavelli effect, then the two-dimensional preference distribution does not depend upon the treatment, DEL-SELF vs. DEL-GEN. In contrast, Machiavelli logic implies that preferences for delegation are "higher" – i.e., shifted up – under DEL-SELF. [8]

[8]The delegation preferences d include attenuation benefits (c.f., the Hamman et al. (2010) Result 1 of Section 2). A subtle issue arises in interpreting these benefits. (We thank a referee for raising this issue.) If they depend upon Q (see Appendix for discussion), then d will also depend upon Q. The logic we describe here for H4 survives if either of the following is true: (1) Hamman et al. (2010) attenuation results from a generalized reduction in guilt from allowing a selfish allocation to occur, so that d does not vary with Q, or (2) if, alternately, attenuation only occurs when a selfish allocation is actually implemented (so that d rises

Consider choices of the direct generous option 2. Under the DEL-SELF treatment, the generous allocation is only implemented with probability $1 - Q$ under delegation, but with probability one under a direct option 2 decision. Therefore, some subjects with a sufficiently strong preference for the generous allocation will opt for option 2, rather than the delegated option 3, even though they have a positive preference for delegation ($d > 0$); these are the subjects with preferences in area E of Figure 1; for them, option 2 is advantageous because it delivers the preferred generous allocation with higher probability. The same is not true in treatment DEL-GEN, where the generous allocation is implemented with probability $1 - Q$ under either a delegated or direct generous option choice. In DEL-GEN, *any* positive preference for delegation ($d > 0$) leads generous subjects ($U < 0$) to choose the payoff-equivalent delegated option 3; "area E subjects" prefer to delegate. By this logic, a direct generous option choice has a higher probability in the DEL-SELF treatment *so long as* preferences are treatment-invariant (H4). Only if the distribution of allocation / delegation (U/d) preferences depends upon the treatment can DEL-SELF produce a *lower* probability of a direct generous decision. Under Machiavelli logic, the preference for delegation is shifted upward under DEL-SELF, reducing the propensity for a direct generous decision and thereby explaining any violation of H4.

A final hypothesis concerns the extent to which dictators choose to delegate or not in the two treatments. In DEL-SELF, there are competing forces: Machiavelli preferences favor delegation (as the delegated decision is bad for others) and the desire to retain decision rights favors a direct decision (Bartling et al., 2014; Fehr et al, 2013; and Owens et al., 2014). In DEL-GEN, however, both effects favor the direct generous decision over a delegated one.

Hypothesis 5 (H5). (A) In the selfish delegation treatment DEL-SELF, the propensity to delegate is higher, lower or the same as the propensity to choose the payoff-equivalent direct selfish option. (B) In the generous delegation treatment DEL-GEN, the propensity to delegate is lower than the propensity to choose the payoff-equivalent direct generous option.

with Q), then we implicitly assume that the distribution of Q beliefs does not vary with treatment. The latter premise is supported by instructions describing player B decisions and the construction of Q that are identical across treatments. In fact, there are no cross-treatment differences in actual player B decisions.

5 Logistics

The dictator side of the experiment is conducted in three economics classes at U.C, Merced. 92 students are enrolled in the two main treatments, 44 in treatment DEL-SELF and 48 in treatment DEL-GEN. The parallel dictator game (DIC) is run with 48 additional students. Matched receivers are in two additional courses.

Participation is purely voluntary and has no bearing on course assessment, both of which are indicated to the students before the experiment begins. Participants are instructed not to communicate with each other and are carefully monitored to ensure privacy. Participant decisions are made with complete anonymity. Anonymity is stressed both verbally and in written instructions. Each student is identified for payment by a registration number indicated on both the experimental questionnaire and an attached tag. One week after the experiment is completed, each student presents his or her tag to the experimenter for payment; the dollar cash payment is contained in a closed envelope identified by the registration number so that the experimenter cannot observe who is being paid what. In addition, to avoid any potential concerns for dollar costs to the experimenters or a desire to "please", we convey verbally to the students that we are interested in their decisions for our research and that they should make the decisions that they prefer under the indicated circumstances.

Treatments are randomly assigned by mixing of questionnaires. There is minor variation in the gender composition in the two main treatments. Overall, 53.5 percent of the sample is male, 56.8 percent in the DEL-SELF treatment and 47.9 percent in DEL-GEN; the z-statistic for the difference equals 0.85. 54.2 percent of the DIC sample is male.

6 Results

Figure 2 and Table 2 describe the raw results from the experiment. Both indicate the proportion of dictators who choose each of the three available options (the direct selfish option 1, direct generous option 2, and delegation option 3) in the two treatments, DEL-SELF and DEL-GEN respectively. Table 3 presents regressions that estimate treatment effects (DEL-SELF vs. DEL-GEN) on the propensity to delegate (Hypothesis 1, H1) and to choose the direct generous option 2 (Hypothesis 4, H4) when controlling for gender and course effects. In the control DIC treatment, 41.7 percent of dictators choose the selfish option and 58.3 percent choose the generous option, closely paralleling similar outcomes in Gawn & Innes (2019a).

In the selfish delegation treatment DEL-SELF, 29.5 percent of dictators delegate, compared with 10.4 percent in the generous delegation treatment DEL-GEN. The 19.1 percentage point difference is statistically significant (z=2.30) and consistent with the Machiavelli conjecture (H1).[9] In the regressions (Table 3, panel A), the estimated magnitude and significance of the DEL-SELF (vs. DEL-GEN) treatment effect is very close to that found in the raw proportions data. A male gender is estimated to promote both delegation and selfishness, but not to a statistically significant extent.

One explanation for the higher delegation rate in DEL-SELF could be that, in DEL-GEN, the direct selfish option 1 is more attractive due to a higher probability that the selfish allocation is actually implemented (probability one in DEL-GEN vs. probability Q in DEL-SELF).[10] The results do not provide statistical evidence of this effect, as the proportion of dictators choosing direct selfish is only marginally higher in DEL-GEN (see Table 2 and Figure 2). This direction of effect is also consistent with Machiavelli logic as an upward shift in preferences for delegation under DEL-SELF reduces the propensity to make a direct selfish decision.

Three other pieces of evidence indicate that the support we find for the Machiavelli Hypothesis 1 is due to a preference effect. First, to determine whether the result is due to a level effect – a greater extent of selfishness in DEL-SELF vs. generosity in DEL-GEN – we can compare the *proportion* of those who delegate an equivalent decision in the DEL-

[9] z statistics for cross-treatment differences in Table 2 are normalized Mann-Whitney-Wilcoxon ranked sum U values. Parametric difference-in-mean analogs are similar.

[10] We are grateful to an anonymous referee for making this point. In Figure 1, this effect is illustrated by area B preferences that produce a direct selfish decision under DEL-GEN, but not under DEL-SELF, precisely due to the different probabilities. However, as indicated in the text, this effect is also consistent with Machiavelli logic and, as a result, does not provide an opportunity to test for treatment-invariant preferences against an alternative hypothesis of Machiavelli preferences.

SELF treatment (43.3% of 30 selfish dictators) and in the DEL-GEN treatment (18.5% of 27 generous dictators). Consistent with Hypothesis 2, the difference is large (24.8%) and we reject the null hypothesis that the propensity for delegation by the selfish in DEL-SELF equals the propensity for delegation by the generous in DEL-GEN. The corresponding test statistic is statistically significant (z=1.985, see Appendix).[11] H1 cannot be explained by a higher propensity for selfishness in our experiment.

Second, is evidence for H1 due to a greater preference for delegation by the selfish vs. the generous – a preference correlation, rather than a Machiavelli effect? In an earlier paper (Gawn & Innes, 2019a), we find the opposite, that it is dictators who would otherwise be generous (without a delegation option) who tend to delegate selfish decisions and not those who would otherwise be selfish. To address this issue in the context of the present experiment, we compare DEL-SELF decisions to the parallel dictator game DIC. Consistent with Hypothesis 3, we estimate a *smaller* share of the otherwise-selfish who prefer to delegate – 7.3% (3.0% out of the 41.7% selfish from DIC) – as we do for the otherwise-generous – 45.4% (26.5% out of the 58.3% generous from DIC).[12] We do not find evidence that H1 is explained by a greater propensity for delegation by the selfish.

Third, we can perform a direct test of context-invariant preferences (H4). Contrary to H4, the difference between the proportion of dictators choosing the direct generous option in DEL-SELF vs. DEL-GEN is a negative fourteen percent. The test statistic for rejection of H4 (a one-sided test by construction) is significant at a ten percent level; p values for the test statistic vary from 0.066 to 0.086, depending upon controls (Tables 2 and 3). We reject a necessary implication of *no* Machiavelli preference effect.

[11] In the Appendix, we construct the z statistic for the null hypothesis that the relevant ratio of population propensities – the propensity for delegation as a share of the total propensity for a selfish (respectively, generous) choice in DEL-SELF (DEL-GEN) – is equal in the two treatments, contrary to H2. Note that the test does not evaluate the two self-selected samples described in the text (selfish in DEL-SELF and generous in DEL-GEN), although the latter provide suggestive evidence.

[12] The delegating selfish proportion (3.0%) is estimated by the difference between the proportion selfish in DIC (41.7%) and the proportion directly selfish in DEL-SELF (38.6%). The delegating generous proportion (26.5%) is estimated by the difference between the proportion generous in DIC (58.3%) and the proportion directly generous in DEL-SELF (31.8%). The cross-treatment difference in shares delegating is large (38.1%= 45.4% - 7.3%), although a z statistic for the null hypothesis of a zero difference (approximately standard Normal using standard logic, see Appendix) is insignificant (z=1.144).

Finally, the experiment gives results consistent with Hypothesis 5. In the DEL-SELF treatment, the propensity for delegation is less than for the equivalent direct selfish choice (by a statistically insignificant 9.1 percent). An intrinsic preference for direct decision making deters delegation, dominating any Machiavelli preference for delegating the selfish / bad news decision. However, in the DEL-GEN treatment, direct decision making is strictly preferred, with a propensity to delegate that is 35.4 percentage points less than the propensity to choose the payoff-equivalent direct generous option 2 (z=3.258).[13] Consistent with Hypothesis 5(B), both Machiavelli logic and an intrinsic preference for direct decision-making combine to produce a greater extent of direct (vs. delegated) generous decisions in the DEL-GEN treatment.

7 Conclusion

We present a dictator experiment in which dictators have an option to delegate their decision to another player. In one treatment, a delegated decision is payoff-equivalent to the direct choice of selfish payments that favor the dictator and are correspondingly bad for the receiver. In the other treatment, a delegated decision is payoff-equivalent to the direct choice of generous payments that are good for the receiver. Machiavelli's (1532) famous conjecture is that delegation will be favored when the decision is bad news for others, while direct decisions will be favored when the decision is good news for others. Although this conjecture is generally motivated by the prospect for blame from the receiver (Bartling & Fischbacher, 2012), we are interested in whether primitive preferences for delegation are affected in the same way even when there is no prospect for blame or credit from others.

Our experiment provides evidence in support of the conjecture; the propensity to delegate is almost 20 percentage points higher when delegation mimics a selfish decision vs. when it mimics a generous one. We provide evidence that this result is attributable to a preference effect: A payoff-equivalent delegated (vs. direct) decision is relatively more preferred when the equivalent decision is disadvantageous for the receiver than when it is advantageous.

[13] The z statistic for the within-treatment difference is a normalized value for the one sample Wilcoxon signed rank test.

In confirming a "Machiavelli preference effect," our results highlight the role of responsibility attribution in driving delegation behavior. Bartling & Fischbacher (2012) show that a measure of a person's responsibility for a decision better explains punishment behavior of others at the receiving end than do measures of outcome-based or intention-based social preferences. How others respond to our actions depends upon how responsible they think we are for those actions (see also Engl, 2018). Because delegation reduces the extent to which others assign responsibility to a dictator for the ultimate decision, delegating selfish decisions reduces punishment and is favored for dictators.

In the experiment presented here, delegation does not provide any signal to others about one's moral type; there is simply no scope for punishment or response or any inference by others based on a dictator's allocation or delegation decision. Nevertheless, delegation can reduce one's own sense of responsibility for a morally virtuous or selfish decision and thereby attenuate any self-signal about social identity (Grossman & van der Weele, 2017; Gneezy, Kajackaite & Sobel, 2018; Dufwenberg & Dufwenberg, 2018). This logic, we believe, underpins the evidence we find for the Machiavelli conjecture. Delegation is disadvantageous when it attenuates self-attributed responsibility for a generous decision and is advantageous when it attenuates responsibility for a selfish decision.

We began the paper by noting two competing forces underpinning a principal's choice to delegate an allocation decision. On one hand, delegation reduces any preference for generosity by creating moral wiggle room (Hamman, Loewenstein & Weber, 2010; Dana, Weber & Kuang, 2007). On the other hand, recent research documents a prevailing preference for retaining decision rights – deterring a delegation decision (Bartling, Fehr & Herz, 2014; Fehr, Herz & Wilkening, 2013; Owens & Grossman, 2014). This tradeoff motivates an inquiry into *when* delegation occurs and, in turn, what effects delegation markets will have. The experiment indicates a third force at play in driving delegation decisions: Delegation reduces self-ascribed responsibility for outcomes, motivating the delegation of selfish decisions and not generous ones. Consistent with the benchmark results of Hamman et al. (2010), the implication is that delegation services will be in greater demand by principals when they promote self-interested outcomes that harm those with whom the hired delegates interact.

8 References

Aimone, J. & D. Houser. 2011. "Beneficial Betrayal Aversion." *PLoS ONE*, 6(3): e17725. doi:10.1371/journal.pone.0017725.

Aimone, J. & D. Houser. 2012. "What You Don't Know Won't Hurt You: A Laboratory Analysis of Betrayal Aversion." *Experimental Economics*, 15, 571-588.

Aimone, J. & D. Houser. 2013. "Harnessing the Benefits of Betrayal Aversion." *Journal of Economic Behavior & Organization*, 15, 571-588.

Andreoni, J. & B. D. Bernheim. 2009. "Social Image and the 50-50 Norm." *Econometrica*, 77(5), 1607-36.

Bartling, B., E. Fehr & H. Herz. 2014. "The Intrinsic Value of Decision Rights." *Econometrica*, 82, 2005-39.

Bartling, B. & U. Fischbacher. 2012. "Shifting the Blame: On Delegation and Responsibility." *Review of Economic Studies*, 79 (1), 67-87.

Bohnet, I. & R. Zeckhauser. 2004. "Trust, Risk and Betrayal." *Journal of Economic Behavior & Organization*, 55, 467-484.

Bolton, P., & M. Dewatripont. 2005. *Contract Theory.* MIT Press.

Charness, G., R. Cobo-Reyes, N. Jimnez, J. Lacomba, & F. Lagos. 2012. "The Hidden Advantage of Delegation: Pareto Improvements in a Gift Exchange Game." *American Economic Review*, 102(5), 2358-2379.

Charness, G. & M. Sutter. 2012. "Groups Make Better Self-Interested Decisions." *Journal of Economic Perspectives*, 26(3), 157-176.

Coffman, L. 2011. "Intermediation Reduces Punishment (And Reward)." *American Economic Journal: Microeconomics*, 3(4), 77-106.

Coffman, L. 2017. "Intermediaries in Fundraising Inhibit Quality-Driven Charitable Donations." *Economic Inquiry*, 55(1), 409-24.

Coffman, L. & A. Gotthard-Real. 2019. “Moral Perceptions of Advised Actions.” *Management Science*, 65(8), 3904-27.

Collins, S., J. Hamman, & J. Lightle. 2018. “Market Interaction and Pro-Social Behavior: An Experiment.” *Southern Economic Journal*, 84(3), 692-715.

Cullen, R. 2008. *The Poverty of Corrupt Nations.* Toronto. Blue Butterfly Publishing.

Dana, J., D. Cain, & R. Dawes. 2006. “What You Don’t Know Won’t Hurt Me: Costly (But Quiet) Exit in Dictator Games.” *Organizational Behavior & Human Decision Processes*, 100(2), 193-201.

Dana, J., R. Weber, & J. Kuang. 2007. “Exploiting Moral Wiggle Room: Experiments Demonstrating an Illusory Preference for Fairness.” *Economic Theory*, 33(1), 67-80.

Della Vigna, S., J. List & U. Malmendier. 2012. “Testing for Altruism and Social Pressure in Charitable Giving.” *Quarterly Journal of Economics*, 127(1), 1-56.

Drugov, M., J. Hamman, & D. Serra. 2014. “Intermediaries in Corruption: An Experiment.” *Experimental Economics*, 17, 78-99.

Dufwenberg, M. & M. Dufwenberg. 2018. “Lies in Disguise: A Theoretical Analysis of Cheating.” *Journal of Economic Theory*, 175, 248-64.

Engl, F. 2018. “A Theory of Causal Responsibility Attribution.” Working Paper, University of Cologne.

Exley, C. 2016. “Excusing Selfishness in Charitable Giving: The Role of Risk.” *Review of Economic Studies*, 83(2), 587-628.

Exley, C. 2018. “Incentives for Pro-Social Behavior: The Role of Reputations.” *Management Science*, 64(5), 2460-71.

Exley, C. 2020. “Using Charity Performance Metrics as an Excuse Not to Give.” *Management Science*, 66(2), 553-63.

Falk, A., & M. Kosfeld. 2006. "The Hidden Costs of Control." *American Economic Review*, 96(5), 1611-1630.

Fehr, E., H. Herz & T. Wilkening. 2013. "The Lure of Authority: Motivation and Incentive Effects of Power." *American Economic Review*, 103, 1325-59.

Fershtman, C., & U. Gneezy. 2001. "Strategic Delegation: An Experiment." *RAND Journal of Economics*, 32(2), 352-68.

Gawn, G., & R. Innes. 2019a. "Who Delegates? Evidence from Dictator Games." *Economics Letters*, 181, 186-9.

Gawn, G., & R. Innes. 2019b. "Lying Through Others: Does Delegation Promote Deception?" *Journal of Economic Psychology*, 71, 59-73.

Gneezy, U., A. Kajackaite & J. Sobel. 2018. "Lying Aversion and the Size of the Lie." *American Economic Review*, 108(2), 419-53.

Grossman, Z. 2015. "Strategic Ignorance and the Robustness of Social Preferences." *Management Science*, 60(11), 2659-2665.

Grossman, Z. & J. van der Weele. 2017. "Self Image and Willful Ignorance in Social Decisions." *Journal of the European Economic Association*, 15, 173-217.

Haisley, E. & R. Weber. 2010. "Self-Serving Interpretations of Ambiguity in Other- Regarding Behavior." *Games & Economic Behavior*, 68(2), 614-625.

Hamman, J., G. Loewenstein, & R. Weber. 2010. "Self-Interest Through Delegation: An Additional Rationale for the Principal-Agent Relationship." *American Economic Review*, 100(4), 1826-46.

Lazear, E., U. Malmendier, & R. Weber. 2012. "Sorting in Experiments with Application to Social Preferences." *American Economic Journal: Applied Economics*, 4(1), 136-63.

Machiavelli, N. 1532. *The Prince.* London: Penguin Classics, 2003.

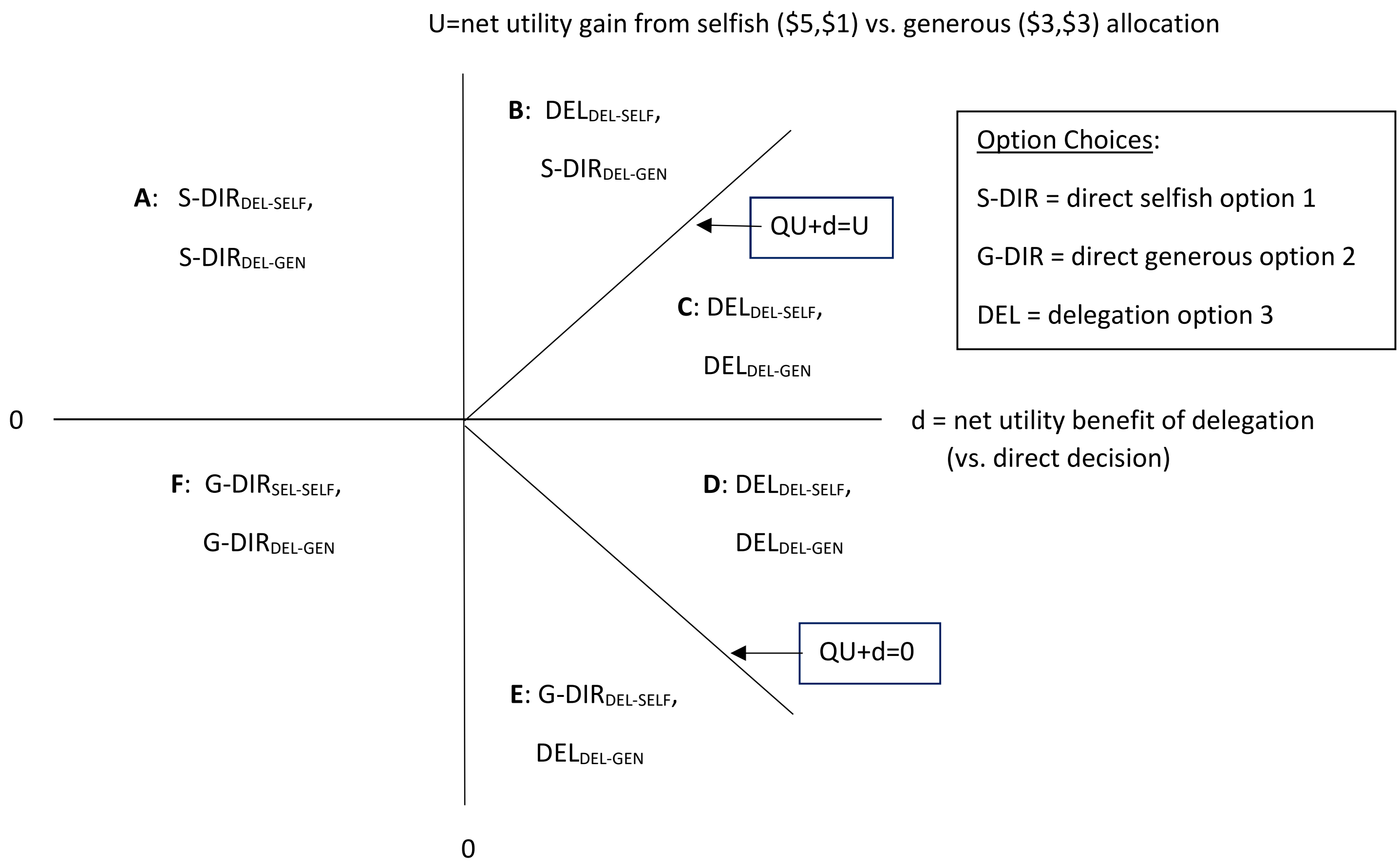
Oexl, R. & Z. Grossman. 2013. "Shifting the Blame to a Powerless Intermediary." *Experimental Economics*, 16, 306-12.

Owens, D., Z. Grossman & R. Fackler. 2014. "The Control Premium: A Preference for Payoff Autonomy." *American Economic Journal: Microeconomics*, 6, 138-61.

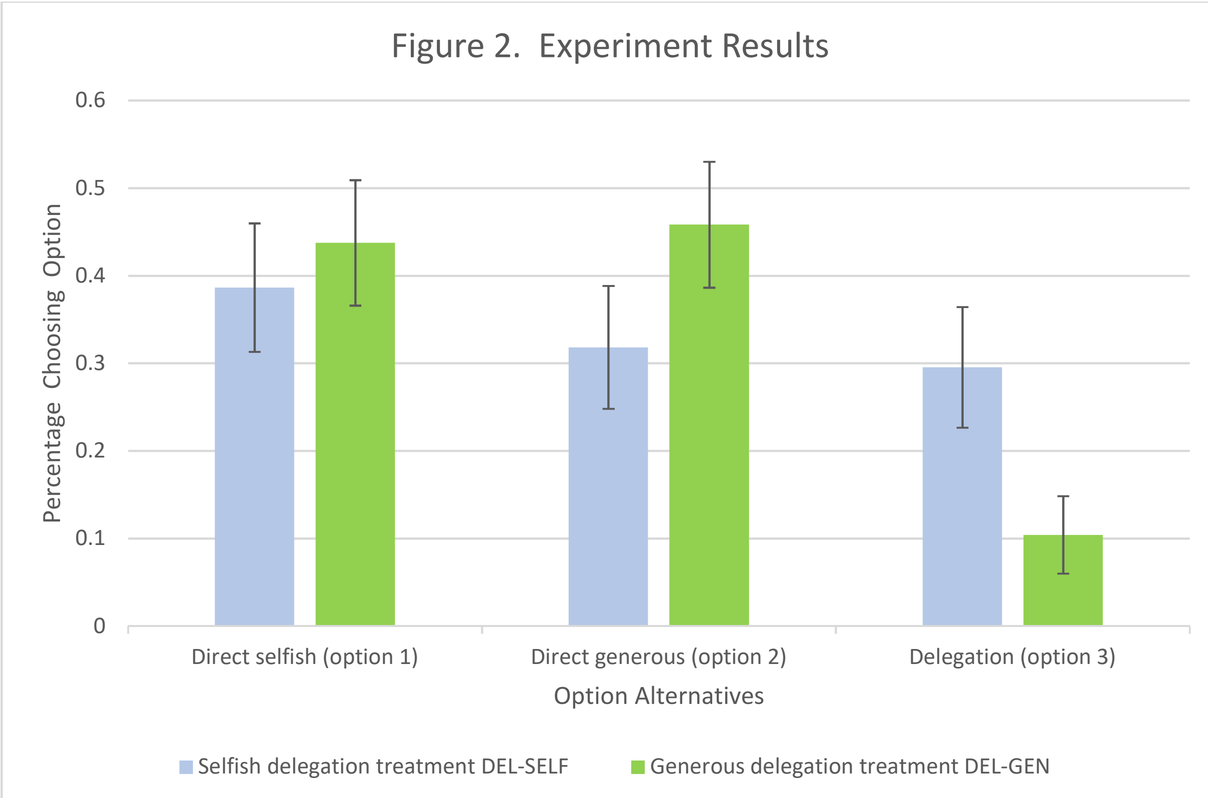
Steffel, M., E. Williams & J. Perrmann-Graham. 2016. "Passing the Buck: Delegating Choices to Others to Avoid Responsibility and Blame." *Organizational Behavior & Human Decision Processes*, 135, 32-44.

Figure 1. Optimal Decisions in Treatments DEL-SELF and DEL-GEN

Under Treatment-Invariant Preferences



Note: Figure 1 depicts optimal option choices for different (U,d) preferences under each of the two main treatments, *assuming* (U,d) preferences are treatment-invariant (contrary to Machiavelli). Available option decisions are direct selfish (S-DIR), direct generous (G-DIR), or delegation (DEL), and the treatment corresponding to each decision (DEL-SELF or DEL-GEN) is indicated in the subscript. Each area (A, B, C, D, E, F) defines (U,d) preferences that support the indicated option choices. Q is the proportion of DIC subjects choosing option 1 in a prior (player B) session / classroom (see Table 1).



Note: Figure 2 depicts the proportions of subjects choosing each of the available options (direct selfish option 1, direct generous option 2, and delegation option 3) in the two main treatments, DEL-SELF (selfish delegation) and DEL-GEN (generous delegation).

Table 1. Payoff Probabilities by Option and Treatment

Treatment→	DEL-SELF		DEL-GEN		DIC	
Payments→	Selfish ($5, $1)	Generous ($3,$3)	Selfish ($5, $1)	Generous ($3,$3)	Selfish ($5, $1)	Generous ($3,$3)
Player A Option Choice ↓	Payment Probabilities ↙ ↘		Payment Probabilities ↙ ↘		Payment Probabilities ↙ ↘	
Direct selfish option 1	Q	1-Q	1	0	Q	1-Q
Direct generous option 2	0	1	Q	1-Q	0	1
Delegation option 3	Q	1-Q	Q	1-Q	--	--

Notes: Table 1 describes the probabilities with which each allocation (selfish ($5,$1) vs. generous ($3,$3)) is implemented under each of the decision options (direct selfish option 1, direct generous option 2, or delegation option 3) available in each of the treatments (DEL-SELF, DEL-GEN, and DIC). Option labels (1,2,3) are varied in the experiment. Q = proportion of DIC subjects choosing the direct selfish option 1 in a prior (player B) session / classroom.

Table 2. Experiment Results

Treatment → Option choice ↓	DEL-SELF (Selfish delegation) (N=44)	DEL-GEN (Generous delegation) (N=48)	Difference (DEL-SELF – DEL-GEN) (z-statistic)
% Direct selfish	0.3864	0.4375	
% Direct generous (H4)	0.3182	0.4583	-0.1401 (-1.368)*
% Delegation (H1)	0.2954	0.1042	0.1912 (2.296)**

Notes: Table 2 indicates the proportion of subjects choosing each of the three available options (direct selfish option 1, direct generous option 2 and delegated option 3) in each of the two (between subject) treatments, DEL-SELF (selfish delegation) and DEL-GEN (generous delegation). Treatments implement payments as described in Table 1. The last column of Table 2 indicates cross-treatment differences in proportions (DEL-SELF minus DEL-GEN). z statistics are normalized Mann-Whitney-Wilcoxon ranked sum U values. **p<0.05 (two-tail), *p=0.086 (one tail).

Table 3. Regressions

Dependent Variable →	(A) Delegation / option 3 (H1) (1=Delegation, 0 otherwise)			(B) Direct generous / option 2 (H4) (1=Direct generous option 2, 0 otherwise)		
Estimation→	OLS	Probit	OLS	OLS	Probit	OLS
	Coefficient (p-value)	Marg. Effect (p-value)	Coefficient (p-value)	Coefficient (p-value)	Marg. Effect (p-value)	Coefficient (p-value)
DEL-SELF Dummy	0.2032 (0.0164)**	0.2041 (0.0190)**	0.1976 (0.0216)**	-0.1542 (0.0657)*	-0.1578 (0.0657)*	-0.1486 (0.0730)*
Male	No	No	0.0808 (0.3327)	No	No	-0.0807 (0.4390)

Notes: N=92 in all models. Table 3 reports coefficient estimates (in OLS linear models) and marginal effects (in probit models) for the DEL-SELF dummy regressor (=1 if treatment is DEL-SELF and =0 if treatment is DEL-GEN). In panel (A), the dependent variable is the delegation (option 3) choice dummy (=1 if delegation option 3 chosen by the dictator, =0 otherwise). In panel (B), the dependent variable is the direct generous (option 2) choice dummy (=1 if direct generous option 2 chosen by the dictator, =0 otherwise). All models include course effects. Robust p-values in parentheses, two-sided in panel A, one-sided on DEL-SELF in panel B. Probit marginal effects at means. **p<0.05 (two-tail), *p<0.10 (one tail).

Appendix 1. Sample Instructions for "Machiavelli Preferences Without Blame"

Sample Instructions for DEL-SELF Experiment

In this experiment, you (player A) are matched with another student in a different classroom (player C). **Each of you obtains a base payment of $3** for this experiment. Based on a decision that you make, this initial set of payments can be changed in the following way:

> <u>**Payment Change**</u>: You can earn $2 more, with player C earning $2 less, for total payments of **$5 to YOU and $1 to PLAYER C**.

You will also be matched with two other students, players B and D, who are in a similar situation as you and player C, but in different economics classrooms. Both players B and D also get a base payment of $3 each, and player B decides whether to make the Payment Change, so that B earns $2 more (total of $5) and D earn $2 less (total of $1).

For your matched Player B's classroom, we will calculate the percentage of *all* "player B's" who choose to make the Payment Change (vs. not). We will call this percentage Q. For example, if half of the player B's choose to make the Payment Change, and half do not, then Q will be 50 percent.

YOU have three options:

> <u>Option 1</u>: YOU choose to MAKE the Payment Change directly ($2 more for you and $2 less for player C).
> *In this case, the Payment Change will be implemented for you and player C with probability Q.*
>
> <u>Option 2</u>: DO NOT make the Payment Change (so that you and C each obtain the initial $3).
> *In this case, your choice – NO PAYMENT CHANGE – will be implemented with 100 percent probability.*
>
> <u>Option 3</u>: Implement PLAYER B's decision.
> *In this case, if your matched player B chooses to make the Payment Change for himself/herself and his/her player D, then the Payment Change will be made for you and player C as well. Likewise, if your player B chooses not to make the change, then the Payment Change will NOT be implemented for you and player C.*

You know what the Options are, but your player C **DOES NOT**. Player C only knows that someone else will make a choice that determines payments to both of you. Your decision does not affect Players B and D in any way.

YOUR DECISION: I choose ***PLEASE CIRCLE ONE***

Option 1	**Option 2**	**Option 3**
(MAKE the Payment Change directly)	(Do NOT make the Payment Change)	(Implement PLAYER B's decision)

Variations in Option Order: A variation on the above instructions switches the order of Options 1 and 3 (with Option 1 becoming the delegation option and Option 3 becoming the Payment Change option).

Sample Instructions for DEL-GEN Experiment: The same as for DEL-SELF, except the menu of options is:

> Option 1: YOU choose to MAKE the Payment Change directly ($2 more for you and $2 less for player C).
> *In this case, the Payment Change will be implemented for you and player C with 100% probability.*
>
> Option 2: YOU choose NOT TO MAKE the Payment Change (so that you and C each obtain $3).
> *In this case, your choice – NO PAYMENT CHANGE – will be implemented with probability 1-Q. The Payment Change will be made for you and player C with probability Q.*
>
> Option 3: Implement PLAYER B's decision.
> *In this case, if your matched player B chooses to make the Payment Change for himself/herself and his/her player D, then the Payment Change will be made for you and player C as well. Likewise, if your player B chooses not to make the change, then the Payment Change will NOT be implemented for you and player C.*

Sample Instructions for Parallel Dictator Game (DIC)

In this experiment, you (player A) will be matched with another student in a different classroom (player C). **Each of you obtains a base payment of $3** for this experiment. Based on a decision that you make, this initial set of payments can be changed in the following way:

> **Payment Change**: You can earn $2 more, with player C earning $2 less, for total payments of **$5 to YOU and $1 to PLAYER C**.

For another session of this experiment (in another economics classroom), we will calculate the percentage of *all* "A players" who choose to make the Payment Change (vs. not). We will call this percentage Q. For example, if half of player A's choose to make the Payment Change, and half do not, then Q will be 50 percent.

YOU have two options:

> Option 1: MAKE the PAYMENT CHANGE ($2 more for you and $2 less for player C).
> *In this case, the Payment Change will be implemented for you and player C with probability Q.*
>
> Option 2: DO NOT make the Payment Change (so that you and C each obtain the initial $3).
> *In this case, your choice – NO PAYMENT CHANGE – will be implemented with 100 percent probability.*

You know what the Options are, but your player C **DOES NOT**. Player C only knows that someone else will make a choice that determines payments to both of you.

YOUR DECISION: I choose ***PLEASE CIRCLE ONE***

Option 1	**Option 2**
(MAKE the Payment Change)	(Do NOT make the Payment Change)

Instructions for Receivers

You make no decision in this experiment. You have been matched with another student in a different classroom. Payments to you and your matched student, for this experiment, are determined based upon a decision made by your matched student.

Appendix 2. Derivation of H4

In this Appendix, we derive the H4 implication of context / treatment invariant preferences described in the text.

Let us assume (contrary to Machiavelli) that delegation preferences are invariant to context (DEL-SELF vs. DEL-GEN), with U = net utility (benefit) of a selfish vs. generous allocation (on support $[\underline{u}, \overline{u}]$, $\underline{u} < 0 < \overline{u}$), d = (context invariant) utility benefit of delegation (on support $[\underline{d}, \overline{d}]$, $\underline{d} < 0 < \overline{d}$), and Q = assessed probability of a selfish choice by player B (on support [0,1]).

Note that d can include any Hamman, Loewenstein & Weber (2010) preference attenuation effect (Result 1, Section 2). There are two ways one might think of attenuation effects here: (1) if delegating reduces guilt from allowing/choosing a selfish allocation, then d incorporates this effect (promoting a higher d that may or may not completely offset a preference for direct decisions) but d will not vary with Q, or (2) if delegating specifically elevates the utility advantage of the selfish (vs. generous) allocation, then d rises with Q because the attenuation benefit is only obtained when the selfish allocation is actually implemented. To illustrate the second interpretation, suppose a represents the increased utility benefit of the selfish (vs. generous) allocation under a delegated (vs. direct) decision, then d includes both the attenuation-free preference for delegation (say, d^*) and the expected attenuation benefit of delegation (Qa), i.e., $d = d^* + Qa$. With this interpretation of attenuation, we assume that the distribution of Q beliefs about player B are treatment-invariant as instructions describing B's decision are identical across treatments. In either case, context-invariant preferences imply that the distribution of d does not depend upon the treatment, DEL-SELF vs. DEL-GEN.

The model described here follows from Von-Neumann-Morgenstern preferences on primitive utility $U^*(A, \delta)$ where $A \in \{S, G\}$ indicates the selfish (S) vs. generous (G) payoff allocation and $\delta \in \{0, 1\}$ indicates a delegated ($\delta = 1$) or direct ($\delta = 0$) decision. Allowing for Hamman et al. (2010) attenuation effects (a = increased benefit of selfish allocation under delegation), context invariant preferences imply $U^*(G, 1) - U^*(G, 0) = d^*$, $U^*(S, 1) - U^*(S, 0) = d^* + a$, and $d = d^* + Qa$. Normalizing $U^*(G, 0) = 0$, we can define $U = U^*(S, 0)$ and the correspondence between options and utilities is as follows.

Assuming context-invariant preferences (H4), subjects obtain respective net utilities $QU/0/QU + d$ under options 1/2/3 in the DEL-SELF treatment; the set of subjects directly choosing the "NO change" option 2 is therefore $\phi_{NC-DEL-SELF} = \{U, d, Q | QU + d < 0$ and $U < 0\}$. In DEL-GEN, subjects obtain $U/QU/QU + d$ under options 1/2/3, and the set choosing "NO change" is $\phi_{NC-DEL-GEN} = \{U, d | d < 0$ and $U < 0\}$.

Figure 1 depicts choices in the two (DEL-SELF and DEL-GEN) treatments for alternate (U, d) preferences. Figure 1 is drawn conditional on the Q belief and probabilities of choices in the experiment (as described in H4) integrate over the belief distribution. In the DEL-SELF treatment, (U, d) preferences within area $(E + F) = \phi_{NC-DEL-SELF}$ produce a direct "NO change" (NC option 2) decision. For the DEL-GEN treatment, the corresponding area is only $F = \phi_{NC-DEL-GEN}$. Because the probability of area E is non-negative, the probability of the "NO change"/option 2 decision is no larger in DEL-GEN than in DEL-SELF so long as preferences are context-invariant (so that areas E and F are the same in the two treatments). H4 follows directly.

Intuitively, area E captures the benefit of a generous (NC) decision in DEL-SELF for those who prefer both the generous allocation ($U < 0$) and delegation ($d > 0$). Under treatment DEL-GEN, where the NC decision is implemented with probability $1 - Q$ under either a delegated or direct generous choice, only those who prefer generous ($U < 0$) and direct decisions ($d < 0$) – those in area F – will opt for the NC option 2.

Only if the distribution of (U, d) preferences is *treatment dependent* can DEL-SELF produce a lower probability of a "NO change" choice. Under Machiavelli logic, the distribution of d preferences is shifted downward / leftward under DEL-GEN (generous delegation) vs. DEL-SELF (selfish delegation); area F becomes larger under DEL-GEN. Machiavelli logic thus permits a higher probability of "NO change" (contrary to H4), but notably does not *imply* a higher probability (with $E \geq 0$). In this sense, a rejection of H4 provides evidence for a strong Machiavelli preference effect.

Appendix 3: Test Statistic for Proportions Delegating

Let p_k = mean of Bernoulli distribution k, q_k = sample mean from k, and $k = sd$ index the sample s ($s \in \{S, G, D\}$, S = DEL-SELF, G = DEL-GEN, and D = DIC) and choice of option d ($d \in \{1, 2, 3\}$), with n_s = number of observations for sample s. Each Bernoulli observation takes a value of one if the decision is d, and zero if not.

The null hypotheses for H2 and H3, respectively, are as follows:

H2: The population proportion of selfish-preferring subjects who delegate in DEL-SELF, $[p_{S3}/(p_{S3}) + p_{S1})]$ equals the proportion of generous-preferring subjects who delegate in DEL-GEN, $[p_{G3}/(p_{G3}) + p_{G2})]$. Rewriting this null: $[p_{S3}p_{G2}] - [p_{S1}p_{G3}] = 0$.

H3: The population proportion of otherwise-generous who delegate, $(p_{D2} - p_{S2})/p_{D2}$, equals the proportion of otherwise-selfish who delegate, $(p_{D1} - p_{S1})/p_{D1}$. Rewriting this null: $[p_{D2}p_{S1}] - [p_{D1}p_{S2}] = 0$.

The following null and alternative hypotheses encompass both H2 and H3. Let $s \in \{k, l\}$, $d \in \{v, w.z\}$, $k \neq l$, $v \neq w$, and $v \neq z$.

$$H_0 : G^* = (p_{kv}p_{lw}) - (p_{kz}p_{lv}) = 0; \qquad H_1 : (p_{kv}p_{lw}) - (p_{kz}p_{lv}) > 0$$

where, for H2, $k = S$, $l = G$, $v = 3$, $w = 2$, and $z = 1$; and for H3, $k = D$, $l = S$, $v = 2$, and $w = z = 1$.

Let $G = (q_{kv}q_{lw}) - (q_{kz}q_{lv})$ = estimated G^*. By independence of observations in the two samples, $E(G) = G^*$. In addition, the variance for G can be written as follows:

$$V(G) = V(q_{kv}q_{lw}) + V(q_{kz}q_{lv}) - 2C,$$

where (1) from the Appendix in Gawn & Innes (2019b),

$$V(q_c q_d) = p_c p_d\{1 + (n_c - 1)p_c + (n_d - 1)p_d - (n_c + n_d - 1)p_c p_d\}/(n_c n_d)$$

for $(c, d) \in \{(kv, lw), (kz, lv)\}$; (2)

$$C = E\{(q_{kv}q_{lw} - p_{kv}p_{lw})(q_{kz}q_{lv} - p_{kz}p_{lv})\}$$

$$= E\{q_{kv}q_{kz}\}E\{q_{lw}q_{lv}\} - (p_{kv}p_{kz}p_{lw}p_{lv})$$

$$= -[n_k + n_l - 1][p_{kv}p_{kz}p_{lw}p_{lv}]/[n_k n_l].$$

and (3) the last equalities use independence and

$$E[q_{sv}q_{sy}] = E\{[\sum_i x_{vi} \sum_j x_{yj}]\}/n_s^2 = E\{[\sum_i x_{vi}x_{yi}[+[\sum_i \sum_{j\neq i} x_{vi}x_{yj}]\}/n_s^2 = [n_s-1][p_{sv}p_{sy}]/n_s$$

for $y \in \{w, z\}$, where x_{di} and x_{dj} are (Bernouli) observations from sample s for choice of option d, $x_{di}x_{ei} = 0$ for $e \neq d$, and x_{di} and x_{ej} are independent for $j \neq i$.

Let $W(G)$ denote the estimated value for $V(G)$, obtained by substituting estimated values (q_{sd}) for true values (p_{sd}) in the variance relation above. By the above properties, the Central Limit Theorem and the Law of Large Numbers, the following statistic is approximately distributed as a standard normal under the null:

$$t_0 = G/\{[W(G)]^{1/2}\} \overset{a}{\sim} N(0, 1) \text{ under } H_0$$