

NHH



Norwegian School of Economics

Bergen, Spring 2022

# Right to Restrict?

*A study of legitimacy as a driver of hard paternalistic interventions.*

**Torbjørn Sandmo**

**Supervisor: Alexander W. Cappelen**

Master Thesis, M.Sc. Economics and Business Administration,

Specialisation: Economics (ECN)

NORWEGIAN SCHOOL OF ECONOMICS

This thesis was written as a part of the Master of Science in Economics and Business Administration at NHH. Please note that neither the institution nor the examiners are responsible – through the approval of this thesis – for the theories and methods used, or results and conclusions drawn in this work.

*Abstract*

In this thesis, I study the effect of legitimacy on people's willingness to implement a hard paternalistic intervention towards another person. In an incentivised experiment, participants decide whether to restrict the freedom of a stakeholder to prevent said stakeholder from making a mistake. I vary participants' sense of legitimacy along two dimensions: whether they have been given positive feedback on their performance on a set of cognitive tasks (merit), and whether they have been assigned a leader role in relation to the stakeholder in question (leadership). I find that people become less willing to implement the hard paternalistic intervention after being assigned a leadership role. I do not find significant effects of receiving positive feedback on performance on people's willingness to intervene. My results shed light on how paternalistic preferences may change depending on the degree of hierarchy in interpersonal relationships. Thus, they offer insights into paternalistic motivation in a wide range of hierarchical interpersonal relationships throughout society, such as those between employers and their employees, doctors and their patients, lawyers and their clients or parents and their children.

---

# Contents

<b>Contents.....</b>	<b>3</b>
<b>1. Introduction .....</b>	<b>5</b>
<b>2. Literature Review .....</b>	<b>8</b>
<b>2.1 Paternalism .....</b>	<b>8</b>
2.1.1 Defining Paternalism .....	8
2.1.2 Hard, Soft and Libertarian Paternalism .....	11
2.1.3 Paternalistic Motivation.....	13
<b>2.2 Legitimate Power .....</b>	<b>15</b>
<b>3. Conceptual Framework .....</b>	<b>17</b>
<b>3.1 Spectator Utility Function .....</b>	<b>17</b>
<b>3.2 Types of Spectators .....</b>	<b>18</b>
<b>3.3 Potential Effects of Legitimate Power on Paternalistic Preferences .....</b>	<b>19</b>
<b>4. Methodology.....</b>	<b>22</b>
<b>4.1 Experimental Design.....</b>	<b>22</b>
4.1.1 First stage: Spectators.....	22
4.1.2 Second stage: Stakeholders .....	24
<b>4.2 Sample and Data Collection .....</b>	<b>26</b>
4.2.1 Sample Descriptives .....	26
<b>4.3 Empirical Strategy .....</b>	<b>29</b>
4.3.1 Empirical Strategy: Main Regressions .....	29
4.3.2 Hypotheses .....	30
4.3.3 Empirical Strategy: Heterogeneity Analysis .....	30
<b>5. Results.....</b>	<b>32</b>
<b>5.1 Choice of Regression Model .....</b>	<b>32</b>
<b>5.2 Descriptive Results for R.....</b>	<b>33</b>
<b>5.3 Main Regression Results.....</b>	<b>34</b>
<b>5.4 Heterogeneity Analysis .....</b>	<b>38</b>
<b>6. Discussion .....</b>	<b>43</b>
<b>6.1 Discussion of Results Relating to the Main Hypotheses .....</b>	<b>43</b>

---

<b>6.2</b>	<b>Discussion of Exploratory Results .....</b>	<b>44</b>
6.2.1	Risk Preferences .....	44
6.2.2	Political Orientation.....	45
6.2.3	Control Group Intercept.....	46
<b>7.</b>	<b>Conclusion.....</b>	<b>48</b>
	<b>References .....</b>	<b>49</b>
	<b>Appendix A. – Ethical Considerations .....</b>	<b>52</b>
	<b>Appendix B. – Experimental Instructions.....</b>	<b>54</b>
	<b>Appendix C. – Demographical Distribution Across Treatment Groups.....</b>	<b>59</b>

## Tables

Table 1	<i>Number of observations per treatment group</i>	25
Table 2	<i>Sample statistics</i>	28
Table 3	<i>Main regression</i>	37
Table 4	<i>Heterogeneity table</i>	39
Table 5	<i>Heterogeneity table: Risk and political orientation</i>	42
Table C	<i>Demographic distribution across treatment groups</i>	59

## Figures

Figure 1	<i>Illustration of proposed effect mechanisms from increased legitimate power</i>	21
Figure 2	<i>Proportion of spectators choosing to restrict stakeholder choice over treatment groups</i>	33
Figure 3	<i>Proportion of spectators choosing to restrict stakeholder choice over self-reported risk preference</i>	33
Figure B.1	<i>Structural overview of spectator stage of experiment</i>	54

# 1. Introduction

The degree to which it is legitimate to intervene paternalistically in other people's lives is an important moral question, both for governments in relation to their citizens and for individuals in relation to each other. Paternalism, as described by Gerald Dworkin in the Stanford Encyclopaedia of Philosophy, is "the interference of a state or an individual with another person, against their will, defended or motivated by a claim that the person interfered with will be better off or protected from harm." (Dworkin, 2020) Terminology varies across literature with regards to the two actors involved in a paternalistic intervention. For the purposes of this thesis, I will refer to the actor implementing the paternalistic intervention as the "spectator", and the target of the intervention as the "stakeholder".

The past decade has seen a rise in the popularity of libertarian paternalism in public policy. This is a non-intrusive form of paternalism that furthers the good of a stakeholder without restricting his freedom of choice. (Thaler & Sunstein, 2003; Thaler & Sunstein, 2020; Le Grand & New, 2015) Libertarian paternalism has been popularized through the concept of nudging, where subtle changes are made to a stakeholder's choice environment to help him realise his true preference, by simplifying information, making certain options more salient or removing potential sources of bias. Though not consistently effective, nudges have in some contexts been shown to generate substantial benefits in terms of stakeholder welfare, at a negligible cost to stakeholder freedom. (Mädche, 2019; OECD, 2017)

Not all paternalistic interventions preserve the stakeholder's freedom of choice, however. The focus of this thesis is a class of paternalistic interventions referred to as hard paternalism. These are interventions that restrict the freedom of a stakeholder, without the stakeholder's consent, with the goal of making him better off, while disregarding or diminishing his autonomy. (Pope, 2004) When a hard paternalistic intervention takes place, it is because a spectator has decided that concerns for a stakeholder's welfare should take priority over concerns for that stakeholder's freedom of choice.

Philosophers have long discussed, and continue to discuss, the morality of hard paternalism. (Mill, 1859; Dworkin, 1983; Pope, 2004; Scoccia, 2008; Dworkin, 2020) In his 1859 essay *On Liberty*, John Stuart Mill introduced what is today referred to as the harm principle, stating that: «The only purpose for which power can be rightfully exercised over any member of a civilised community, against his will, is to prevent harm to others. His own good, either

physical or moral, is not a sufficient warrant.» (Mill, 1859, p. 68) This statement constitutes a categorical rejection of hard paternalistic interventions. Others again argue that hard paternalism is permissible in some contexts, treating liberty and autonomy not as inviolable, but rather as a moral good that can be weighed against the potential welfare benefits from an intervention. (Scoccia, 2008; Le Grand & New, 2015)

Philosophers are not alone in wrestling with the morality of hard paternalism however, as practical examples of hard paternalism are found throughout a range of domains in society. Parents must decide to what extent they should restrict the freedom of their children to protect them from harm. University professors must decide whether lectures and problem sets should be made mandatory to save students from hyperbolic discounting. Doctors must decide whether to respect the wishes of a patient, even when they as experts see those wishes as wrong or even dangerous. Hard paternalism is also a highly relevant topic for lawmakers, as paternalistic policies often spark heated political debates, both across the isles in parliaments and in society at large. Examples of hard paternalism in the state-citizen relationship can be prohibiting recreational drugs, raising taxes on unhealthy foods, mandating helmets for motorcyclists or forbidding doctors from administering euthanasia upon request from a terminally ill patient. (Le Grand & New, 2015) As illustrated by the examples above, a facet of many relationships where hard paternalism takes place is that there exists some form of hierarchy between the spectator and the stakeholder, either due to the spectator having some expertise that the stakeholder lacks, or because the spectator has a leadership role in relation to the stakeholder.

A growing body of literature within behavioural economics shows that people value both the freedom and welfare of those around them. (Ambuehl et al. 2020; Bartling et al., 2022; Fehr & Fischbacher, 2003; Weiß et al. 2020) It is when these two concerns are at odds each other that the possibility of a hard paternalistic intervention arises. That would be when a stakeholder's freedom of choice enables him to make a mistake that compromises his welfare. A spectator observing this type of situation must decide whether to exercise power over the stakeholder to prevent the mistake in question from taking place. If the spectator chooses to intervene, the stakeholder's freedom of choice is diminished. If the spectator chooses not to intervene, this allows the stakeholder to act in a way that compromises his welfare.

When a spectator sets out to implement a hard paternalistic intervention, that spectator must have some means of influencing the stakeholder for the intervention to be successful.

Psychologists Bertram Raven and John French define power as the “potential influence of some individual or group O over an individual P.” (Raven & French, 1958b, p.400) In light of this definition, it is evident that power is a prerequisite for hard paternalism. When discussing different bases of power, French and Raven (1959) include what they refer to as legitimate power, “defined as that power which stems from internalized values in P which dictate that O has a legitimate right to influence P and that P has an obligation to accept this influence.” (French & Raven, 1959, p.153) In relation to hard paternalistic interventions, O is here the spectator and P is the stakeholder.

In this thesis, I study how people’s perception about the legitimacy of their own power affects their willingness to implement a hard paternalistic intervention towards a stakeholder. I study this empirically through an online survey experiment, where some spectators are assigned a leader role in relation to a stakeholder, while others receive confirmation that they performed well relative to others on a set of problem-solving tasks. Some spectators also receive a combined treatment, being informed that they were assigned a leader role in relation to the stakeholder because of their strong problem-solving performance. I am interested in studying if and how these different treatments affect spectators’ willingness to implement a hard paternalistic intervention towards a stakeholder.

With this master’s thesis, I hope to contribute to a growing literature on the paternalistic preferences of individuals, by studying the factors that influence people’s decision of whether to restrict another person’s freedom for the sake of that other person’s welfare. Studying how legitimacy affects paternalistic preferences might give us a better understanding of the circumstances under which people perceive hard paternalism as acceptable. Understanding the factors that motivate hard paternalistic interventions is relevant in a wide range of interpersonal relationships, such as those between employers and employees, doctors and patients, lawyers and clients, teachers and students, parents and their children, etc. Insights into when hard paternalism is seen as acceptable can also be of relevance to policy makers in predicting how hard paternalistic policies in the form of prohibitions, mandates or taxes will be received by the general public depending on context.

The thesis is organised in the following seven chapters: 1. Introduction, 2. Literature review, 3. Conceptual framework, 4. Methodology, 5. Results, 6. Discussion, 7. Conclusion.

## 2. Literature Review

In this chapter, I present previous literature on paternalism and legitimate power. In the first section, I present a general definition of paternalism, as well as some ways to classify different types of paternalistic interventions. I also give a brief overview of previous literature on what motivates paternalism, before finally placing the intervention studied as part of this thesis into the classification frameworks presented. In the second section of this chapter, I present a framework for thinking about different bases or sources of power, focusing on the concept of legitimate power as described by psychologists French and Raven (1958a, 1958b, 1959).

### 2.1 Paternalism

One of the most prominent writers on paternalism, philosopher Gerald Dworkin, describes paternalism as “the interference of a state or an individual with another person, against their will, defended or motivated by a claim that the person interfered with will be better off or protected from harm.” (Dworkin, 2020) In this section, I present a more structured definition of paternalism, also from Dworkin. I start by elaborating on the conditions that must be met for an intervention to be regarded as paternalistic. I then go on to discuss distinctions between hard, soft and libertarian paternalism, before presenting previous literature on paternalistic motivation in individuals.

#### 2.1.1 Defining Paternalism

Dworkin defines paternalism as follows:

“X acts paternalistically towards Y by doing (or omitting) Z, if and only if the following three conditions are met:

1. Z (or its omission) interferes with the liberty or autonomy of Y.
2. X does so without the consent of Y.
3. X does so only because X believes Z will improve the welfare of Y (where this includes preventing his welfare from diminishing), or in some other way promote the interests, values or good of Y.” (Dworkin, 2020)



---

Tying this back to the terminology used in this thesis, we say that X is the spectator, Y is the stakeholder, and Z is the paternalistic intervention. In the following, I elaborate on what Dworkin's three conditions entail.

### *First condition*

Dworkin's first condition can be met in two ways: Either through an interference with a stakeholder's liberty, or through an interference with his autonomy. Elaborating on this condition, Dworkin stated that "there must be a usurpation of decision making, either by preventing people from doing what they have decided, or by interfering with the way in which they arrive at their decisions." (Dworkin, 1983, p.107) "Preventing people from doing what they have decided" can be understood as interfering with their liberty, while "interfering with the way in which they arrive at their decisions" can be understood as interfering with their autonomy. On the topic of liberty, economist and philosopher Amartya Sen writes that: "Individual liberty can be seen to require – among other things – that each individual should have a recognized personal sphere in which his preferences and his alone would count in determining the social preference." (1983, p.7) By that it is meant that the liberty of a stakeholder relates to the range of options between which he is left to choose for himself based on his own preferences, without interference from others. Thus, interfering with a choice that was previously left up to the stakeholder constitutes an interference with his liberty.

As for interfering with a stakeholder's autonomy, this can be understood as compromising his voluntariness. In Haworth's (1991) review paper *Dworkin on Autonomy*, two requirements for autonomy are discussed: authenticity and independence.

On authenticity, Haworth writes that "People act authentically, in Dworkin's sense, insofar as they identify with the desires that move them to act." (1991, p.131) There is no generally accepted answer to the question of when a stakeholder ceases to act authentically. The clearer it is that a stakeholder's actions are guided by some force with which he does not identify however, the easier it is to argue that his actions are inauthentic. Addiction is for instance a stronger case for inauthenticity than temptation.

As for the second prerequisite for autonomy, independence, Dworkin makes a distinction between procedural and substantive independence, of which the former is more central to the question of autonomy according to Haworth. Procedural independence relates to the procedure with which a person's motives are formed. If a person has been manipulated into a certain set

of motives or preferences, that calls into question whether their motivations are procedurally independent. (Haworth, 1991) Once again, the line is not clear in terms of under what circumstances one should cease to regard someone's motives as procedurally independent.

To summarize: For an intervention to be paternalistic, it must interfere with either the liberty or autonomy of a stakeholder. The stakeholder's liberty is interfered with if someone prevents him from doing what he wants. The stakeholder's autonomy is interfered with if someone tries to change his preferences or behaviours, given that he has arrived at his motives in an authentic and procedurally independent way. However, it is not entirely clear how the status of these two conditions for autonomy, authenticity and procedural independence, should be determined on a case-to-case basis.

### *Second condition*

Dworkin's second condition states that an intervention is only paternalistic if it is implemented without the stakeholder's consent. Julian Le Grand and Bill New (2015) criticise this second condition for being made redundant by Dworkin's first condition, arguing that consenting to having your liberty or autonomy interfered with is logically incoherent unless one is talking about prior consent. Le Grand and New further argue that imposing prior preferences on a stakeholder who appears to have changed his mind should still be considered paternalism.

I will interpret Dworkin's second condition in this direction: That imposing a stakeholder's previously stated preference still constitutes a case of paternalism, so long as the stakeholder appears to have changed his mind. If the stakeholder explicitly consents to an intervention in the present however, then Dworkin's second condition is no longer met. Also, the absence of consent is enough for an intervention to qualify as paternalistic. Thus, interfering with the liberty or autonomy of a stakeholder under the assumption that he would likely consent is still considered paternalism, so long as consent is not explicitly given.

### *Third condition*

Dworkin's third condition states that the spectator's decision to intervene towards the stakeholder must be motivated by a concern for the stakeholder's own good. If the spectator has significant ulterior motives for intervening that do not relate to her benevolence towards the stakeholder, then the intervention can no longer be said to qualify as paternalistic. It is important to underline that it is the spectator's motivation that determines whether an

intervention constitutes paternalism. An intervention that furthers a stakeholder's own good would not be considered paternalistic unless this is what motivated the intervention. Inversely, an intervention that is harmful to the stakeholder can still be paternalistic, so long as the spectator implementing the intervention was motivated by a genuine belief that the intervention would make the stakeholder better off.

### **2.1.2 Hard, Soft and Libertarian Paternalism**

Hard, soft and libertarian are central classifications of paternalistic interventions in literature. The distinctions between them relate to the concepts of liberty and autonomy. In this section I start by explaining the difference between hard and soft paternalism. I then offer a short presentation of libertarian paternalism, and how it differs from the other two.

In normative philosophical discourse on paternalism, the distinction between “hard” and “soft” paternalism relates to whether an intervention that limits the liberty of a stakeholder also interferes with the autonomy of that stakeholder. (Pope, 2004; Scoccia, 2008; Le Grand & New, 2015; Dworkin, 2020) I quote Pope, who states that: “In short, soft paternalism justifies interference with a subject's liberty to benefit that same subject only if –indeed precisely because– her choices are not substantially voluntary.” (Pope, 2004, p.667-668) In other words, soft paternalism is justified by the perception that the stakeholder's actions are involuntary. Restricting his liberty would therefore not interfere with his autonomy, since his autonomy is already compromised. Proponents of soft paternalism further argue that limiting the liberty of a stakeholder to determine whether he is acting autonomously is justified. If it is determined that the stakeholder's actions are “substantially voluntary” however, then continued interference with his liberty would no longer be classified as soft paternalism.

Instead, such an intervention would classify as hard paternalism. Pope (2004) provides four conditions for an intervention to qualify as hard paternalism: (1) The spectator must limit the stakeholder's liberty (2) motivated by benevolence, while (3) disregarding the stakeholder's contemporaneous preferences, as well as (4) disregarding or deliberately limiting the stakeholder's autonomy. For the purposes of this thesis, we say that the key characteristic of hard paternalism is that it restricts a stakeholder's liberty. The question of whether a liberty restricting intervention is “soft” or “hard” comes down to a question of whether the intervention is motivated by a perception that stakeholder's actions are not “sufficiently

voluntary”. (Pope, 2004; Scoccia, 2008) The question of what qualifies as sufficiently voluntary lies beyond the scope of this thesis, however. I quote Scoccia, who states that: “The point of drawing these two distinctions is to use them to do some normative work—to argue, for example, [...] that only and/or all soft paternalism is permissible.” (Scoccia, 2008, p.359)

As my analysis is positive as opposed to normative, I do not consider it necessary for me to take a strong position on whether choices made by the stakeholders in my experiment qualify as “substantially voluntary” in the eyes of spectators, especially as perceptions about what should qualify as substantially voluntary is likely to vary across my sample. Instead, I will define the intervention available to spectators in my experiment purely in terms of liberty, classifying any paternalistic intervention that restricts a stakeholder’s freedom of choice as hard. In doing so, I acknowledge that my use of the term hard paternalism could be contested. However, as there is no objective measure of what qualifies as “substantially voluntary”, and since perceptions about the autonomy of the stakeholder is likely to be heterogenous across my sample of spectators, I see this as the best solution.

As for paternalistic interventions that do not restrict the liberty of the stakeholder, this is what is referred to as libertarian paternalism. This term was introduced by Richard Thaler and Cass Sunstein (2003), who explain that: “When we use the term libertarian to modify the word paternalism, we simply mean liberty-preserving.” (Thaler & Sunstein, 2020, p.6) Hard paternalism and libertarian paternalism thus have mutually exclusive definitions, since one requires a restriction of liberty to take place while the other requires that liberty is preserved. Libertarian paternalist interventions often involve making changes to the stakeholder’s choice environment to either reduce the potential for bias or to ensure that the stakeholder’s bias will manifest in ways that are beneficial for him. Thaler and Sunstein refer to libertarian paternalistic interventions as *nudges* and emphasise that “To count as a mere nudge, an intervention must be cheap and easy to avoid.” (Thaler & Sunstein, 2020, p.8) Thus, those who actively disagree with the nudge are free to opt out.

To summarize: In this thesis, I concern myself with paternalistic interventions that limit the liberty of the stakeholder. To avoid wrestling with questions regarding who considers whom to be acting substantially autonomously, I simply refer to the liberty-restricting paternalistic intervention that I study in this paper as hard.

---

### 2.1.3 Paternalistic Motivation

In this section, I discuss previous literature on what motivates individuals to implement hard paternalistic interventions towards others. We have established that the main motive for spectators engaging in paternalism is a concern for the target stakeholder's own good. In situations where a stakeholder makes a mistake that leads to a reduction in his welfare, other people, spectators, have an opportunity to intervene to protect the stakeholder from harm or potentially make him better off. This welfare benefit is the argument in favour of paternalistic intervention. However, studies have shown that people place intrinsic value on decision-rights (Bartling et al. 2014; Fehr et al. 2013; Owen et al., 2014) and some of this value is lost to the stakeholder when a paternalistic intervention compromises his right to make his own decisions. This cost to freedom is the argument against paternalistic intervention.

Empirical studies have found that spectators do take into consideration the costs to liberty or autonomy incurred by stakeholders. Spectators show a lower willingness to implement liberty-restricting interventions than liberty-preserving interventions subject to the same potential welfare gain. (Bartling et al., 2022) Spectators are also more willing to intervene towards a stakeholder if that stakeholder has expressed that he does not identify with the desires that drive him towards a certain action. (Ambuehl et al., 2020)

When we think about the spectator's underlying desire to further the stakeholder's own good, this regard for the stakeholder's freedom makes sense. Bartling et al. (2014) find that people assign intrinsic value to decision rights, in that they desire decision rights for its own sake, not just due to the instrumental value that comes with getting to choose between different outcomes. This is consistent with findings in Fehr et al. (2013) and Owen et al. (2014), with both studies showing that people will hold on to decision rights, even when they would achieve a better monetary outcome by delegating decision authority to an agent. Pikulina and Tergiman (2020) also find that people assign intrinsic value to power over others, i.e. that the desire for decision rights is not limited only to choices that concern oneself.

As for motivation in relation to the benefits of paternalism, one should look outside of models of pure self-interest to explain spectator behaviour. Studies have found that one reason why people engage in altruism is that they value fairness. (Fehr & Schmidt, 1999; Almås et al., 2020) Concerns for the wellbeing of others has been found to extend beyond fairness concerns however, to also include empathy (Klimecki et al., 2016) or a desire to feel good about oneself

for having helped someone else, described as a so called “warm-glow” effect. (Anderoni 1990; Ottoni-Wilhelm, 2017)

In terms of paternalistic motivation, the literature also makes distinctions between different classes of paternalism based on how the spectator believes that their intervention will further the stakeholder’s welfare. One such distinction is that between weak and strong paternalism (Dworkin, 2020; Scoccia, 2008), also referred to as means and ends paternalism (Le Grand & New, 2015). When a stakeholder takes an action, it is with the intention of achieving some goal. The goal is the end, and the actions taken to achieve that goal are the means. A spectator may perceive that a stakeholder is choosing actions or means that are inappropriate or insufficient for reaching some goal or end that the stakeholder has chosen for himself. Intervening on those grounds would be weak paternalism or means paternalism. If on the other hand the spectator intervenes because the stakeholder’s goals themselves are perceived to be inappropriate, then that is a case of strong paternalism or ends paternalism.

Ambuehl et al. (2020) make a similar distinction in relation to paternalistic motivation, distinguishing between mistakes-projective spectators and ideals-projective spectators. Mistakes projective spectators are motivated to intervene based on an expectation that others will make mistakes similar to those they would make themselves. Ideals-projective spectators on the other hand project their own ideals or preferences onto others. These spectators are more likely to make intervention decisions that lead the stakeholder to receive the outcome that the spectator would have preferred themselves. Ambuehl et al. (2020) show empirically that some spectators are ideals-projective. Bartling et al. (2022) find empirical evidence for a similar type of motivation in what they call authoritative spectators. These are spectators’ that let their own preferences take precedence over the stakeholder’s own preferences, even when the stakeholder’s own preference is known to them. They thus either take action, or refrain from taking action, to ensure that the stakeholder receives the option that the spectator would have chosen for themselves, even when they know that the stakeholder does not share their preferences. For this thesis, I will use the term authoritative to describe spectators let their own preferences guide their intervention decision away from the stated preference of the stakeholder. The experiment conducted for this thesis involves a choice between two options with different degrees of risk. In my analysis, I control for spectator’s self-reported risk preference on a scale from zero to ten. My hope is that this data combined with data on intervention decisions might allow me to determine whether authoritative spectators are present in my sample.

To summarize and highlight how the theory described above ties back to this thesis: For this thesis, I study people's willingness to implement a hard paternalistic intervention. I define the intervention as hard on the grounds that it involves restricting a stakeholder's freedom of choice. Spectators are informed that a stakeholder is acting in a way that is inconsistent with his true preferences, and that they can restrict the stakeholder's choice to ensure that he gets the option he actually prefers. Some spectators may be authoritative paternalists, in that they disregard the stakeholder's preferences in favour of their own, despite being told what the stakeholder actually prefers. I discuss these spectators further when presenting my conceptual framework in chapter 3.

## 2.2 Legitimate Power

Power, legitimacy and authority are three closely related concepts that have been studied extensively both within psychology and sociology. Authority is a term that relates to the possession of power in different forms. Sociologist Max Weber refers to "Herrschaft" (popularly translated to authority) as the possession of "accepted power", i.e. that people obey you voluntarily, independent of expectations of punishments or rewards. (Weber, 1922) In some literature within psychology, Weber's authority is referred to as legitimate authority (French & Raven 1959; Tyler, 2006) as part of discussions of authority as a broader concept. Fehr et al. (2013) use the term authority in relation to decision rights. For the sake of simplicity, I will refrain from using this term going forward. It is useful to know however that this term is sometimes used interchangeably with what I will refer to as "legitimate power".

What is power, and what does it mean for power to be legitimate? Social psychologists John French and Bertram Raven describe power as the "potential influence of some individual or group O over an individual P." (1958b, p.400) They go on to describe five bases of power: (1) Reward power, derived from the capacity of O to administer rewards to P, (2) coercive power, derived from the capacity of O to punish P, (3) legitimate power, derived from the perception that O has the right to prescribe behaviours for P, (4) referent power, derived from P's identification with O and (5) expert power, derived from the perception that O possesses special knowledge or expertise. (French & Raven, 1959) Raven later added a sixth base: (6) informational power, derived from the capacity of O to control the information available to P. (Raven, 1965)

Of particular significance to this thesis is the concept of legitimate power. Psychologist Tom R. Tyler describes legitimacy as “a psychological property of an authority, institution or social arrangement that leads those connected to it to believe that it is appropriate, proper and just.” (Tyler, 2006, p.375) Hence the specification from French and Raven that legitimate power is derived from the *perception* that O has the right to prescribe behaviours for P. They further state that “the individual P may also consider the legitimacy of the attempts to use other types of power by O. (...) P might change in response to coercive power by O, but it will make a considerable difference in his attitude and conformity if O is not seen as having a legitimate right to use such coercion.” (French & Raven, 1959, p.154)

French and Raven list the following three bases for legitimate power: (i) Cultural values, (ii) acceptance of the social structure and (iii) designation by a legitimizing agent. By cultural values as a base of legitimate power, it is meant that O can have characteristics that the culture defines as giving him the right to prescribe behaviours for P. What these characteristics are varies from culture to culture, but examples can include age, sex, caste, education, occupation, physical characteristics, etc.

In the experiment conducted for this thesis, I vary the degree to which spectators have legitimate power over a stakeholder by means of two treatments: One based on cultural values, the other based on designation by a legitimizing agent. As personal freedom is such a central concept in the US, it was a challenge to identify characteristics of cultural value that could give rise to legitimate power. In the end, “Merit” was chosen as a characteristic of cultural value in the US, by which I mean the virtue of having performed well in some domain as a result of one’s talent or hard work. In his most recent book, contemporary philosopher Michael J. Sandel discusses the prevalence of meritocratic attitudes in western cultures, particularly in the US. He claims that within egalitarian and free-market liberalism “talent, and in particular innate or natural talent, [is taken] incredibly seriously” (2020, p. 150), while also referring to a “notion that the meritorious should govern.” (2020, p. 27) Throughout his criticism of meritocracy in the US, Sandel paints a picture of merit in US culture that seems to reflect French and Raven’s (1959) description of legitimate power based on cultural values in that O can have characteristics that the culture defines as giving him the right to prescribe behaviours for P.



### 3. Conceptual Framework

In the experiment conducted for this thesis, spectators are informed that a stakeholder has completed some work. As compensation, the stakeholder will receive one out of two possible payment options: a safe bonus option in the form of a fixed payment and a risky bonus option in the form of a lottery. Spectators are informed that the stakeholder prefers the safe option, but that if given the chance to choose for himself, he is going to choose the risky option due to a calculation error. Spectators are then asked to make a choice: Either do nothing and allow the stakeholder to choose his non-preferred option, or implement a hard paternalistic intervention, restricting the stakeholder's choice to ensure that he gets the option he prefers. This choice design builds on the design developed in Bartling et al. (2022).

#### 3.1 Spectator Utility Function

Bartling et al. (2022) present a utility function through which to consider a spectator's choice of whether to implement a hard paternalistic intervention towards a stakeholder. I use this utility function as the foundation for my discussion on how being granted legitimate power might affect spectators' willingness to restrict a stakeholder's choice.

$$V(\alpha, p, c) = (1 - \alpha) \cdot W(b) + \alpha \cdot F(c) \quad (1)$$

$V$  is here the utility of the spectator.  $W$  is the welfare of the stakeholder, a function of the bonus option received, denoted  $b$ .  $F$  is the freedom of the stakeholder, a function of the stakeholder's choice environment, denoted  $c$ . The variable  $\alpha \in [0,1]$  denotes the weight a spectator places on the stakeholder's freedom relative to the stakeholder's welfare.

The spectator is utility maximizing, and therefore chooses to implement the hard paternalistic intervention so long as the net utility gain from doing so,  $\Delta V$ , is positive.

$$\Delta V = (1 - \alpha) \cdot \Delta W(b) + \alpha \cdot \Delta F(c) \quad (2)$$

If the spectator chooses not to intervene, the stakeholder will choose the risky payment option, denoted  $r$ , subject to an unrestricted choice environment, denoted  $c^0$ . If the spectator instead chooses to implement the hard paternalistic intervention by restricting the stakeholder's choice, the stakeholder will receive the safe option, denoted  $s$ , subject to a restricted choice environment, denoted  $c^-$ .

I assume that all spectators share an understanding that implementing the hard paternalistic intervention leads to a reduction in the stakeholder's freedom,  $F(c^-) - F(c^0) < 0$ .

## 3.2 Types of Spectators

Spectators who choose not to implement the hard paternalistic intervention can be categorized into two groups. The first group, whom I will call group A, consists of authoritative spectators. These are spectators who themselves would prefer the risky option over the safe option, and who disregard the stakeholder's preferences in favour of their own. These spectators thus make their decision of whether to intervene based on their own preference  $W(s) - W(r) < 0$ , rather than the spectator's preference  $W(s) - W(r) > 0$ . Combining this with the assumption that all spectators agree that  $F(c^-) - F(c^0) < 0$ , it follows from equation (2) that for these spectators, implementing the hard intervention would lead to a decrease in utility,  $\Delta V < 0$ . Being utility maximizing, they therefore choose not to restrict the stakeholder's choice.

The second group, whom I will call group B, consists of non-authoritative spectators who perceive that the freedom related cost of the hard paternalistic intervention outweighs the potential welfare gain from said intervention. Being non-authoritative, spectators in group B make their intervention decision based on the stakeholder's preference  $W(s) - W(r) > 0$ . However, because they perceive that the gain in welfare does not justify the cost to freedom, they choose not to intervene, as doing so would lead to a change in utility  $\Delta V < 0$ .

The third and final group of spectators considered here, whom I will call group C, are non-authoritative spectators who perceive that the welfare benefits of intervening outweigh the freedom related costs. Being non-authoritative, these spectators make their intervention decision based on the stakeholder's own preference  $W(s) - W(r) > 0$ . Seeing as they perceive that the welfare gain from intervening justifies the freedom related cost, these spectators would experience a net utility gain,  $\Delta V > 0$ , following an intervention. They therefore choose to implement the hard paternalistic intervention by restricting the stakeholder's choice.

We can say that in a given sample of spectators, a fraction  $\pi$  are non-authoritative and a fraction  $(1 - \pi)$  are authoritative. Out of the non-authoritative spectators, we say that a fraction  $(1 - \theta)$  have a  $\Delta V < 0$  (group B), while a fraction  $\theta$  have a  $\Delta V > 0$  (group C).

Thus, the fraction of spectators choosing to implement the hard intervention, i.e. group C spectators, is  $\theta\pi$ . The fraction of spectators who choose not to implement the hard paternalistic intervention, i.e. the sum of groups A and B, is then  $(1 - \pi) + \pi(1 - \theta) = 1 - \theta\pi$ .

### 3.3 Potential Effects of Legitimate Power on Paternalistic Preferences

Recall that legitimate power according to French and Raven (1959) is derived from the perception that O has a right to prescribe behaviours for P. In this section, I discuss potential ways in which being granted legitimate power over the stakeholder might lead to a change in paternalistic preferences for the three different types of spectators described above, using the spectator utility function, described in equation (1), as a foundation.

The authoritative spectators in group A prefer that the stakeholder receives the risky option. As this is the option that the stakeholder is about to choose for himself, these spectators have no incentive to intervene. It seems unlikely therefore that giving these spectators a reason to regard their own power as more legitimate will lead them to change their decision.

The non-authoritative spectators who choose not to implement the hard paternalistic intervention (group B), do so because they believe that the potential welfare gain from an intervention does not justify the associated cost to stakeholder freedom. Contrary to spectators in group A, spectators in group B do have an incentive to intervene, since they want to see the stakeholder receive the safe option. My prediction is therefore that giving these spectators additional reason to regard their own power over the stakeholder as legitimate could change their decision of whether to implement the hard paternalistic intervention.

Perhaps the most intuitive potential effect mechanism from perceiving that one has the right to prescribe behaviours for the stakeholder is that the importance of stakeholder freedom is reduced relative to the importance of stakeholder welfare. This could be illustrated as a reduction to the freedom term based on the degree of self-perceived legitimacy,  $L \in [0,1]$ . Here, I say that  $L = 0$  means having no legitimacy, while  $L = 1$  means having complete legitimacy, i.e. that a spectator thinks they have an indisputable right to prescribe behaviours for others.

$$V(\alpha, p, c) = (1 - \alpha) \cdot W(b) + \alpha \cdot (1 - L) \cdot F(c) \quad (3)$$

Based on this proposed mechanism, one would expect to see an increase in  $\theta$  when spectators are given additional reason to regard their power as legitimate. This change would be driven by non-authoritative spectators from group B changing their intervention decision due to a reevaluation of the freedom related cost of intervening, thereby transferring over to group C. I will refer to this as the “guardian effect”.

A second effect mechanism may also come into play however, if being granted a right to prescribe behaviours for others prompts non-authoritative spectators to become authoritative. A portion of the non-authoritative spectators in group C are likely risk tolerant, in that they themselves would prefer the risky option over the safe option. So long as these spectators remain non-authoritative, their own risk preference does not affect their decision. If these spectators are given reason to believe that they have a right to prescribe behaviours for the stakeholder however, they may choose to disregard the stakeholder’s risk preference in favour of their own,  $W(s) - W(r) > 0 \rightarrow W(s) - W(r) < 0$ . This would remove their initial incentive to intervene, whereby they would transfer over from group C to group A. This in turn would lead to a decrease in  $\theta\pi$ , through a decrease in  $\pi$ . I will refer to this potential mechanism of non-authoritative spectators becoming authoritative as the “emboldening effect”.

Looking at the overall effect of legitimacy inducing treatments on willingness to implement a hard paternalistic intervention may give some indication of which, if any, of these two proposed effect mechanisms is dominant. If the emboldening effect dominates the guardian effect, one expects to see a reduction in the proportion  $\theta\pi$  of spectators choosing to restrict the stakeholder’s choice, as spectators transfer over from group C to group A. If instead the guardian effect dominates the emboldening effect, one would expect to see an increase in the proportion  $\theta\pi$  of spectators choosing to restrict the stakeholder’s choice to ensure that the stakeholder’s receives his preferred bonus option.

To summarize: Spectators in groups A and B choose not to restrict the stakeholder’s choice, while spectators in group C choose to restrict the stakeholder’s choice. An influx of spectators into group C from the other two groups thus increases the overall average willingness to implement the hard paternalistic intervention, while an exodus of spectators from group C to the other two groups decreases the average willingness to implement said intervention.

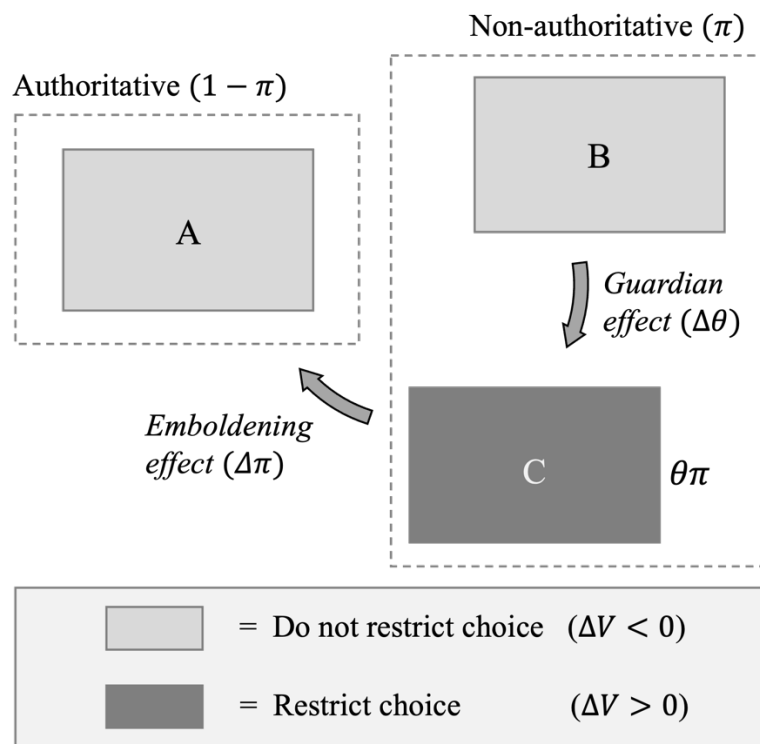


Figure 1: Two proposed effect mechanisms from being granted legitimate power.

## 4. Methodology

The goal of this thesis is to study the effect of legitimate power on people's willingness to implement a hard paternalistic intervention towards another person. To study this empirically, I ran an online survey experiment using a 2×2 factorial, between-subject design involving two types of participants, namely spectators and shareholders. I am interested in spectator behaviour subject to different degrees of legitimate power, as they decide whether to implement a hard paternalistic intervention towards a stakeholder. Spectators and stakeholders were recruited with a ratio of 10:1, so that every spectator was facing a 10% chance that their choice would be implemented towards a real stakeholder. The inclusion of real stakeholders was a way to ensure authentic responses from spectators. Spectators were paid a flat rate of \$0.70 for their participation, but they were subject to moral incentives in that their choices could end up having real consequences for another person.

In this chapter, I present my research methodology. In section 4.1. I present my experimental design. In section 4.2. I present the data collection process and descriptive statistics for my final sample of spectators. In section 4.3. I present my empirical strategy. Instructions given to spectator and stakeholder participants as part of the experiment are presented in appendix B. A project pre-plan was submitted through AsPredicted.org on May 5<sup>th</sup>, 2022.<sup>1</sup>

### 4.1 Experimental Design

The experiment conducted for this thesis had two stages. In the first stage, spectators make a choice of whether to implement a hard paternalistic intervention towards a stakeholder. In the second stage, one in ten randomly selected spectator choices were implemented towards a group of real stakeholders.

#### 4.1.1 First stage: Spectators

The first stage of the experiment, the spectator stage, consisted of three parts. In the first part, spectators were asked to solve five multiple choice mathematical word problems subject to time constraints. In the second part, spectators were assigned to one out of four treatments. In

---

<sup>1</sup> <https://aspredicted.org/uc93z.pdf>

---

the third and final part, spectators were asked to decide whether to implement a hard paternalistic intervention towards a stakeholder. Spectators who answered less than half of the mathematical word problems correctly in the first part of the survey did not proceed to receive a treatment but were instead omitted from the final sample. Thus, my final sample consists only of spectators who answered more than half of the mathematical word problems correctly in the first part of the experiment.

### *Hard Paternalistic Intervention*

In the third and final part of the experiment, spectators were given a choice of whether to implement a hard paternalistic intervention towards a stakeholder. This choice design builds on the design developed in Bartling et al. (2022). Spectators were informed that they would be making a choice that could have real consequences for another person, the other person being the stakeholder. Instructions for this choice revealed the following: After completing some work, a stakeholder is set to receive a bonus payment. The stakeholder will receive one of two payment options: either a safe option of that yields a guaranteed \$4 or a risky option that yields either \$0 or \$10, where either outcome has a .5 probability of occurring. The stakeholder prefers the safe option but will, if allowed to choose for himself, choose the risky option due to a calculation error with respect to the outcome probabilities associated with this option.

After reading these instructions, spectators were presented with a choice: Either restrict the stakeholder's choice or let the stakeholder choose for himself. Restricting the stakeholder's choice means the stakeholder automatically receives the safe option, which is his preferred option. Letting the stakeholder choose freely on the other hand results in the stakeholder choosing the risky option, which is his non-preferred option. Choosing to restrict the stakeholder's choice in this situation constitutes a hard paternalistic intervention. When making this choice, spectators were told, truthfully, that the stakeholder would not be informed about their involvement.

### *Treatments*

Spectators in my final sample received one of four possible treatments prior to making the choice described above. I distribute treatments using a 2×2 factorial, between subject design, where spectators could receive one of four treatments based on two two-level factors. These four treatments were: Base, Merit, Leadership and Combined. The base treatment serves as

my control. The merit treatment was designed to instil a sense of legitimate power based on merit, whereas the leadership treatment was designed to instil a sense of legitimate power based on hierarchical position through a designated leadership role. Spectators who receive the combined treatment essentially receive the merit treatment and the leadership treatment together. Table 1 shows the number of spectators assigned to each treatment.

The decision to omit some spectators relates to the merit treatment. Spectators in the merit treatment group and combined treatment group were informed that they answered the majority of the mathematical word problems in the first part of the experiment correctly. They were also informed that these problems are complex, and that many people get them wrong. In order for the merit treatment to have the desired effect, it was important that spectators were told they performed well relative to some other group. To justify this last statement, it was necessary for the mathematical word problems in the first part of the experiment to have a certain level of complexity, in order to introduce a cut-off that not all spectators would reach. I emphasise again that all spectators in my final sample answered three or more problems correctly. There is thus no difference in actual performance across my treatment groups, only in whether a spectator has been told that they performed well.

The second factor in my experiment relates to the degree of hierarchy in the relationship between the spectator and the stakeholder. Spectators in the leadership treatment group were informed that they had been assigned the role of leader prior to deciding whether to implement the paternalistic intervention in the final part of the experiment. This design was intended to invoke a sense of legitimate power through designation by a legitimizing agent. (Raven & French, 1958a; 1958b; French & Raven, 1959)

#### **4.1.2 Second stage: Stakeholders**

After collecting responses from spectators, I randomly selected 10% of spectators' choices to be implemented towards real stakeholders. To ensure that instructions given to spectators about the preferences of the stakeholder were truthful, I elicited stakeholder preferences with respect to the safe option and the risky option before implementing the spectators' choices. Only stakeholders who preferred the safe option over the risky option, but who still chose the risky option due to calculation error subject to untransparent information were used for implementing spectator choices. For elaboration on how stakeholder preferences were elicited, I refer to appendix B.



Table 1: Observations per Treatment

---

		Leadership	
		No	Yes
Merit	No	$n = 250$	$n = 254$
	Yes	$n = 256$	$n = 257$

---

*Notes:* Overview of the number of spectators per treatment group. “Merit” denotes having been informed that one performed well on a set of problem-solving tasks relative to others. “Leadership” denotes having been assigned a leadership role.

## 4.2 Sample and Data Collection

Spectators were recruited from the US through the online platform Amazon Mechanical Turk on the 11<sup>th</sup> of May 2022. Out of 2403 participants starting the experiment, 179 did not pass an initial attention check. Out of the remaining 2224 respondents, 1026 failed to answer more than half the mathematical word problems correctly in part one. In my project pre-plan, I had specified that spectators who spent less than ten seconds reading instructions for the final choice in the survey would be excluded from analysis on suspicion of inattention. This was the case for 181 participants. Thus, my final sample consists of 1017 spectators.

Spectator choices were implemented towards stakeholders on the 24<sup>th</sup> and the 26<sup>th</sup> of May 2022, also through Amazon Mechanical Turk. Stakeholder preferences were elicited to determine whether participants' preference profile matched the stakeholder preferences described in instructions to spectators.

### 4.2.1 Sample Descriptives

Spectators were asked about their gender, age, education, annual income, region, risk preferences, marital status, number of children and political preferences, for use as demographic controls in regressions. Descriptive statistics for the final sample of spectators can be seen in Table 2. Some US census statistics have been included in this table for comparison. It was never an expectation that the final sample would be nationally representative, since the probability of exclusion after the five mathematical word problems was deemed likely to be heterogeneous across the general population. Another likely reason why sample statistics here differ from statistics for the US general population is that some demographic groups have been found to be underrepresented on Amazon Mechanical Turk, particularly older persons and those with less education. (Chandler et al., 2019) This is reflected also in my data, as my final sample is younger than the general population. People above the age of 55 are underrepresented (16% in sample vs. 36% in the general population), whereas people below the age of 44 are overrepresented (69% in sample vs. 36% in the general population). People above the age of 55 were underrepresented also among the 2224 spectators who started the mathematical word problems, with 13% being 55 or older in this group. Thus, my younger sample seems to be a result of Amazon Mechanical Turk demographic characteristics rather than any exclusion criteria.

The situation is similar for educational attainment. 8% of spectators in my final sample have no college experience, whereas the corresponding statistic is 38% for the general US population. Inversely, 73% of spectators in my sample hold a degree, compared to 45% on the national level. 18% of spectators in my final sample have some college experience but no degree. Intuitively, one might think that this skewed sampling with respect to education would be a consequence of using mathematical word problems as an exclusion criterion. Comparing the sample of spectators who started the mathematical word problems to my final sample however, I find no indication that this is the case. The initial proportion of spectators with no college experience was 5%. For those with some college but no degree, the proportion was 12%. Thus, the high proportion of spectators with degrees in my final sample relative to the general population seems to be a consequence of skewed initial sampling rather than exclusion criteria.

My sample also differs from the general population in terms of region in that the Northeast is overrepresented while the West and to some extent the Midwest are underrepresented. In terms of annual income, there is an underrepresentation of spectators on the higher end of the income distribution, as well as an overrepresentation of spectators in the category \$30 000 to \$60 000.

In the final sample, 57% of spectators are married and 58% have children. Data on risk preferences was gathered by asking spectators to rate their willingness to take risk on a scale from 0 to 10, with 0 being “Completely unwilling to take risk” and 10 being “Very willing to take risk”. The average answer was 5.7, with 6 being the median answer. “High Risk” is here an indicator for whether a person answered 6 or higher, which was the case for 53% of spectators. In terms of political orientation, 53% of spectators identify as Democrat, 24% as Republican and 21% as Independent or Third Party. These results should be seen in relation to abovementioned trends in the sample, particularly with respect to age and education. (Pew Research Center, 2018)

For an overview of demographic distribution across treatment groups in the experiment, see appendix C.

Table 2: Sample Descriptives

	Mean	SD	US
Female	0.45	0.50	0.51
Male	0.54	0.50	0.49
Age: 18 to 34	0.37	0.48	0.29
Age: 35 to 44	0.32	0.47	0.17
Age: 45 to 54	0.15	0.36	0.16
Age: 55 to 64	0.11	0.31	0.17
Age: 65 +	0.05	0.22	0.21
Education: Up to High School	0.08	0.27	0.38
Education: Some College	0.18	0.39	0.17
Education: Bachelor's or Associate	0.57	0.49	0.32
Education: Master's or Above	0.16	0.37	0.13
Income: < \$30 000	0.25	0.43	0.25
Income: \$30 000 to \$60 000	0.41	0.49	0.25
Income: \$60 000 to \$100 000	0.24	0.43	0.22
Income: \$100 000 to \$150 000	0.07	0.26	0.14
Income: > \$150 000	0.02	0.14	0.14
Region: West	0.18	0.39	0.24
Region: Midwest	0.18	0.39	0.21
Region: Northeast	0.26	0.44	0.17
Region: South	0.38	0.48	0.38
High Risk	0.53	0.50	
Married	0.53	0.50	
Has Children	0.53	0.50	
Democrat	0.53	0.50	
Republican	0.24	0.43	
Independent	0.21	0.41	
<i>N</i>	1017	1017	

*Note:* Table showing sample descriptives for spectators, and some census data for the US general population. US census data on education from 2021, on population per region from 2020, on age and gender from 2019, and census data on income from 2018. Retrieved from <https://data.census.gov/> on May 11<sup>th</sup> 2022.

---

## 4.3 Empirical Strategy

### 4.3.1 Empirical Strategy: Main Regressions

I am interested in studying spectators' willingness to implement the hard paternalistic intervention subject to the two factors in my design, both when applied separately in the merit and leadership treatments and together in the combined treatment. To test for effects of my treatments, I use the following empirical specification:

$$R_i = \beta_0 + \beta_1 M_i + \beta_2 L_i + \beta_3 M_i L_i + \gamma X_i + \varepsilon_i \quad (4)$$

$R_i$  is the dependent variable, taking the value of one if a spectator chooses to restrict the stakeholder's choice.  $M_i$  is an indicator variable taking the value of one in treatments where a spectator is informed about their performance on the mathematical word problems.  $L_i$  is an indicator variable taking the value of one in treatments where the spectator is assigned a leader role in relation to the stakeholder.  $M_i L_i$  is the interaction between  $M_i$  and  $L_i$ , thus taking the value of one if a spectator received the combined treatment.  $X_i$  is a vector containing the indicator variables for the following demographic characteristics with respect to the spectator: Gender, age, education, income, region (US), risk preference, marital status, number of children and political orientation. These control variables were specified in the project pre-plan.  $\varepsilon_i$  is the error term.

Of particular interest are the coefficients  $\beta_1$  and  $\beta_2$ , as they denote the effects of the merit treatment and leadership treatment on average willingness to implement the paternalistic intervention. I am also interested in estimates of  $\beta_3$ , capturing interaction effects between these two factors, comparing paternalistic preferences of spectators who received both factors as part of the combined treatment relative to those who received them separately as part of the merit treatment or leadership treatment. I will run regressions both with and without control variables. In the model where  $M_i$ ,  $L_i$  and  $M_i L_i$  are included without demographic controls,  $\beta_0$  can be interpreted as the average willingness to implement the hard paternalistic intervention in the control group.

### 4.3.2 Hypotheses

I here specify the three main hypotheses that I aim to test using the empirical specification described above.

First hypothesis: Receiving the merit treatment changes spectator willingness to implement the hard paternalistic intervention towards the stakeholder.

$$H_0: \beta_1 = 0 \quad , \quad H_A: \beta_1 \neq 0$$

Second hypothesis: Receiving the leadership treatment changes spectator willingness to implement the hard paternalistic intervention towards the stakeholder.

$$H_0: \beta_2 = 0 \quad , \quad H_A: \beta_2 \neq 0$$

Third hypothesis: The effect of receiving one of the treatments is conditional on whether the spectator also receives the other treatment.

$$H_0: \beta_3 = 0 \quad , \quad H_A: \beta_3 \neq 0$$

### 4.3.3 Empirical Strategy: Heterogeneity Analysis

I will also conduct an exploratory heterogeneity analysis to study effects of treatments conditional on demographic control variables. I do this by running regressions using the specification presented in equation (4) on subsets of my sample, such as only female spectators, only below median risk preferences, only spectators from the Northeast etc. I present results of these regressions together for treatment effect comparisons between demographic groups. This provides a more detailed view of treatment effects across my sample and may also reveal interesting sub-group characteristics.

In cases where it is interesting to test the statistical significance of observed treatment effect differences between sub-groups of my sample, I will run regressions using the following empirical specification:

$$R_i = \beta_0 + \beta_1 M_i + \beta_2 L_i + \beta_3 M_i L_i + \beta_4 x_i + \beta_5 M_i x_i + \beta_6 L_i x_i + \beta_7 M_i L_i x_i + \gamma X_i + \varepsilon_i \quad (5)$$

---

$R_i$ ,  $M_i$ ,  $L_i$  and  $M_iL_i$  are the same variables as before.  $x_i$  is a single control variable of interest, tested against the omitted baseline variable in its category. The vector  $X_i$  now contains the same control variables as in equation (4), except for  $x_i$ .  $M_ix_i$ ,  $L_ix_i$  and  $M_iL_ix_i$  are interaction terms between my two factors and the control variable in question, as well as an interaction term between the original interaction between my two factors and the control variable in question. Using this specification, I can test the significance of observed effect differences of  $M_i$ ,  $L_i$  and  $M_iL_i$  conditional on the control variable  $x_i$ , relative to the chosen baseline in the category to which  $x_i$  belongs. These estimated effect differences will be reflected in estimates of  $\beta_5$ ,  $\beta_6$  and  $\beta_7$ .

## 5. Results

In this chapter, I present the results from my experiment. In section 5.1. I offer a short justification for my decision to use the linear probability model (OLS) over logistic regression. In section 5.2. I present descriptive results relating to the dependent variable  $R_i$ . In section 5.3. I present regression results relating to my main hypotheses, as well as some explorative results. In section 5.4. I present exploratory results from a heterogeneity analysis conducted with respect to the treatments given to spectators in the experiment.

My main regression results are presented in table 3. Results from the heterogeneity analysis are presented in tables 4 and 5.

### 5.1 Choice of Regression Model

When running regressions using a binary dependent variable, it is common to use either logistic regression or the linear probability model (OLS). In most use cases, these two model types have been found to yield results of similar significance. (Hellevik, 2007) Throughout my analyses, I use the linear probability model for its more intuitive coefficient estimates. One issue with using the linear probability model with binary data is that it violates the fifth Gauss Markov assumption of homoscedasticity. I mitigate this by using robust standard errors throughout my analysis. I report R squared in my tables for transparency, despite most of its relevance being lost due to dependent variable outcomes being discreet. Though not meaningful in isolation, reported R squared values may still be of some use when making comparisons across models.

*«The results for linear and logistic significance probabilities as we have seen turn out to be nearly identical, even with small samples and skewed distributions on the dependent variable. The problem of impossible predictions may be avoided by including interactions terms in the regression equation and/or by transforming continuous variables into dummy variables.*

*This means that we are not restricted to use logistic regression with a binary dependent variable. This is fortunate, since there in many cases are compelling substantive arguments for preferring the linear approach.» (Hellevik, 2007, p.73)*



Mean of R over Treatments

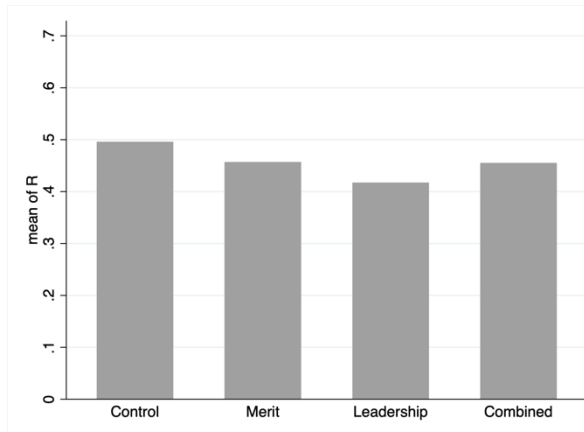


Figure 2: Fraction of spectators choosing to restrict stakeholder choice per treatment group.

Mean of R over Risk Preferences

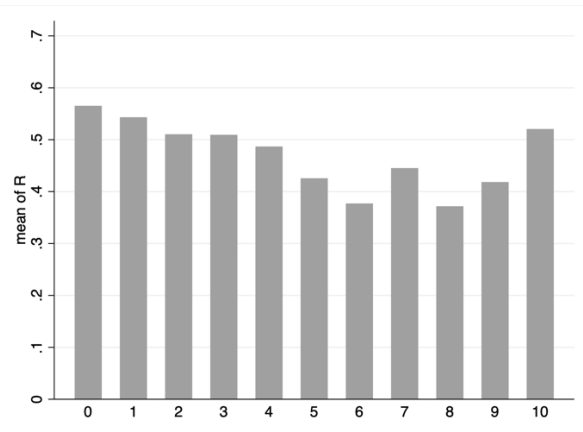


Figure 3: Fraction of spectators choosing to restrict stakeholder choice over spectator risk preferences, with 0 being “Completely unwilling to take risk” and

## 5.2 Descriptive Results for R

Figure 2 shows the fraction of spectators who choose to implement the hard paternalistic intervention per treatment group. We see that the control group has the highest rate of intervention, with 49,6% of spectators choosing to restrict the stakeholder’s choice in this group. The leadership treatment group has the lowest rate of intervention, with 41,7% of spectators choosing to restrict the stakeholder’s choice. Corresponding intervention rates are 45,7% in the merit treatment group and 45,5% in the combined treatment group.

Figure 3 shows the fraction of spectators who choose to restrict the stakeholder’s choice over spectators’ self-reported risk preference on a scale from 0 to 10. From 0, which is the most risk averse, to 6, which is the median risk preference in my sample, we see a steady decline in the proportion of spectators choosing to implement the hard paternalistic intervention to ensure that the stakeholder receives the safe option. This could be an indication that risk preferences play a role in the intervention decisions of some spectators, which would mean that there are authoritative spectators in my sample. Further analysis is required however to determine whether these trends are significant. For spectators with an above median risk preference, the overall trend less consistent. Recall from the conceptual framework that risk tolerant authoritative spectators are expected to not restrict the stakeholder’s choice. It is therefore somewhat surprising that we see a higher willingness to intervene amongst the spectators who report the highest possible score of 10 in terms of willingness to take risk.

## 5.3 Main Regression Results

In this section, I present results from my main regressions, seen in table 3. Results reported in column (6) of table 3 were estimated using the full empirical specification presented in section 4.3.1. I here start by presenting results relating to my three main hypotheses. I then go on to present some additional exploratory results of interest from these regressions. Please note that due to the exploitative nature of my analysis, I set weaker requirements for significance stars than what is standard, reporting estimates significant on the ten percent level ( $p < 0.10$ ) in tables using one star.

**Result 1: No statistically significant effect of the merit treatment on spectators' willingness to implement the hard paternalistic intervention towards the stakeholder.**

Effect estimates for the merit treatment can be seen in the row marked "Merit" in table 3. These are estimates  $\beta_1$  from the empirical specification presented in section 4.3.1. We see that in column (6) of table 3, the merit treatment effect on spectator willingness to intervene towards the stakeholder is estimated to -0.0347. We interpret this estimate as a negative 3,47 percentage point change in the proportion of spectators choosing to restrict the stakeholder's choice associated with having received the merit treatment. This result is not statistically significant ( $p = 0.436$ ).

In conclusion, I fail reject  $H_0: \beta_1 = 0$ .

**Result 2: Weakly significant effect ( $p < 0.10$ ) of the leadership treatment on spectators' willingness to implement the hard paternalistic intervention towards the stakeholder.**

Estimates of the leadership treatment effect can be seen in the row marked "Leadership" in table 3. These estimates correspond to  $\beta_2$  from the empirical specification presented in section 4.3.1. Once again looking at column (6) of table 3, we see that the estimated effect of the leadership treatment on willingness to implement the hard paternalistic intervention is -0.0788, which we interpret as a 7,88 percentage point decrease in average willingness to restrict the stakeholder's choice associated with being assigned a leader role. This estimate has an associated p-value of 0.076, which I here consider to be a weakly significant result.

In conclusion, I can reject  $H_0: \beta_2 = 0$  subject to a significance parameter  $\alpha = 0.10$ .

---

**Result 3: No statistically significant interaction effect between the merit treatment and the leadership treatment.**

Estimates of the interaction effect between merit and leadership can be seen in the row labelled “Merit×Leadership” in table 3. These are estimates of  $\beta_3$  from the empirical specification presented in section 4.3.1. Interpreting  $\beta_3$  is not as straightforward as interpreting  $\beta_1$  and  $\beta_2$ , since it denotes the effect change from the two treatments being applied together as opposed to separately. If  $\beta_3$  has the same direction as  $\beta_1$  and  $\beta_2$ , that is an indication that one treatment amplifies the effect of the other. If  $\beta_3$  is instead estimated to have the opposite direction to  $\beta_1$  and  $\beta_2$ , that is an indication that having already received one treatment reduces the potency of receiving the other. To simplify reading, I included a row marked “lincom” in table 3. This row shows the estimated sum of treatment effects  $\beta_1 + \beta_2 + \beta_3$ . The lincom estimate in column (6) of table 3 can thus be read as an estimated -4,56 percentage point decrease in average willingness to implement the hard paternalistic intervention associated with receiving the merit treatment and the leadership treatment together, relative to the control group. The p-value associated with the  $\beta_3$  estimate in the regression using my full empirical specification is 0.278, which means that the observed interaction effect is not statistically significant.

In conclusion, I fail to reject  $H_0: \beta_3 = 0$ .

**Result 4 (Exploratory): Significant negative correlation between spectators’ risk preference and willingness to implement the hard paternalistic intervention.**

The row marked “HighRisk” in table 3 contains estimates of the effect of a spectator’s own risk preferences on willingness to implement the hard paternalistic intervention. “HighRisk” is a dummy variable, taking the value of 1 if a spectator’s self-reported risk preference on a scale from zero to ten is equal to or greater than the sample median of six. We interpret the negative 0.0789 coefficient estimate seen in column (6) as a 7,89 percentage point decrease in willingness to restrict the stakeholder’s choice associated with having a risk preference equal to or above the sample median, as opposed to lower than the sample median. This result is statistically significant ( $p = 0.019$ ) and can be seen as an indication that the trend we observe in figure 3 is in fact driven by spectators’ risk preferences. As an exploratory result however, this should not be used to draw conclusions about the population parameter without further research.

**Result 5 (Exploratory): Democrat spectators more willing to restrict the stakeholder's choice than republican and third party or independent spectators.**

In regressions where categorical variables for political orientation are included, I use Democrat as the baseline category, as this was the most frequent answer given for political orientation throughout my sample. We see that coefficient estimates in the rows marked “Republican” and “Independent” are statistically significant in all models where they are included. Looking at column (6), we interpret coefficient estimates for republican as an estimated -9,21 percentage point difference between republican and democrat spectators in terms of their average willingness to restrict the stakeholder's choice ( $p = 0.018$ ). Similarly, we see a -10,3 percentage point estimated difference between third party or independent spectators and democrat spectators ( $p = 0.012$ ) in average willingness to intervene. We interpret this as democrat spectators being significantly more willing to implement the hard paternalistic intervention than both republican and independent spectators.

**Result 6 (Exploratory): Significantly higher willingness to implement the hard paternalistic intervention in my control group compared to Bartling et al. (2022).**

Estimates of  $\beta_0$  can be seen in the row marked “\_cons” in table 3. In column (5),  $\beta_0$  is estimated to 0.496. This denotes fraction of spectators in my control group who chose to restrict the stakeholder's choice. As stated previously, the final choice in my experiment builds on the design developed in Bartling et al. (2022). In their large-scale, nationally representative study, the fraction of spectators choosing to implement a hard paternalistic intervention subject to a similar design to that employed here was 0.335. Test the statistical significance of the difference between the  $\beta_0$  estimate in column (5) and the 0.335 mean from Bartling et al. (2022) using a student's t-test returns a p-value of 0.000. The difference in willingness to restrict the stakeholder's choice between these two groups is in other words highly significant. Though I cannot say for certain what causes this effect, potential explaining factors could be differences in data collection tools and sampling criteria. I discuss this further in chapter 6.

Table 3: Regression Results

	(1)	(2)	(3)	(4)	(5)	(6)
Merit	-0.000209 (0.0313)	-0.00114 (0.0313)			-0.0390 (0.0445)	-0.0347 (0.0445)
Leadership			-0.0399 (0.0312)	-0.0444 (0.0314)	-0.0787* (0.0443)	-0.0788* (0.0444)
Merit×Leadership					0.0769 (0.0625)	0.0679 (0.0626)
lincom					-0.0407 (0.0444)	-0.0456 (0.0442)
Male		-0.0226 (0.0330)		-0.0227 (0.0330)		-0.0214 (0.0330)
HighEducation		-0.0246 (0.0379)		-0.0251 (0.0377)		-0.0223 (0.0380)
HighIncome		-0.0291 (0.0354)		-0.0316 (0.0352)		-0.0324 (0.0354)
HighAge		0.0221 (0.0350)		0.0230 (0.0350)		0.0238 (0.0350)
HighRisk		-0.0779** (0.0336)		-0.0785** (0.0335)		-0.0789** (0.0335)
Republican		-0.0920** (0.0387)		-0.0937** (0.0388)		-0.0921** (0.0389)
Independent		-0.104** (0.0409)		-0.104** (0.0408)		-0.103** (0.0408)
_cons	0.456*** (0.0222)	0.542*** (0.0500)	0.476*** (0.0222)	0.565*** (0.0509)	0.496*** (0.0317)	0.579*** (0.0541)
<i>N</i>	1017	1017	1017	1017	1017	1017
<i>R</i> <sup>2</sup>	0.000	0.025	0.002	0.027	0.003	0.028

Notes: OLS regressions on the dependent indicator variable *R*, taking the value of 1 if a spectator chose to restrict the choice of the stakeholder, 0 if not. Included but not reported in models (2), (4) and (6) are the variables “Married”, “Children”, “Prefer not to say” (politics), “Other” (gender), “West”, “Midwest” and “Northeast” (regions). Robust standard errors in parentheses.

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

## 5.4 Heterogeneity Analysis

A heterogeneity analysis was conducted with respect to treatments from the experiment using the empirical strategy described in chapter 4, section 4.3.3. Results from this analysis can be seen in tables 4 and 5. In table 4, I report treatment effects in regressions run on different subsets of my data. In table 5, I report between-group treatment effect differences with respect to risk preference and political orientation. I only report estimates of treatment coefficients for regressions in table four, and only estimates of the coefficients relating to interactions between the treatments and respective control variables in table 5. Demographic controls are included in all regressions here but not reported in the tables. The extensive reporting of heterogeneity results in table 4 is intended to make explorative results available in case they may serve to inform further research. For the purpose of this thesis however, I restrict myself to looking at heterogeneity results relating to risk and political orientation however.

**Result 8 (Exploratory): Leadership treatment effect negative and significant when only studying spectators with a median risk preference or higher, but leadership treatment effect difference between LowRisk and HighRisk is not statistically significant.**

Results involving spectator's risk preferences are of particular interest because they could give some indication of whether authoritative spectators are present in my sample, and how these spectators react to the treatments in my experiment. We see from the column marked "HighRisk" in table 4 that the leadership treatment coefficient estimate is marginally significant in regressions using only those spectators in my sample whose self-reported risk preference was six or higher ( $p = 0.075$ ). However, looking at the row marked "Leadership×HighRisk" in table 5, we do not see a significant leadership treatment effect difference between HighRisk and LowRisk spectators.

**Result 9 (Exploratory): Significant treatment effect differences between Democrats and Republicans in my sample.**

We see treatment effect differences between democrats and republicans in table 4. Results in column (2) of table 5 show that these differences are statistically significant both with respect to the merit treatment and the leadership treatment. These results should be seen in relation to the baseline differences between these two groups in terms of willingness to intervene, reported in table 3.

Table 4: Heterogeneity

	(Male)	(Female)	(LowAge)	(HighAge)	(LowEduc)	(HighEduc)	(LowInc)	(HighInc)	(LowRisk)	(HighRisk)
Merit	-0.0275 (0.0606)	-0.0412 (0.0675)	-0.0663 (0.0532)	0.0429 (0.0843)	0.00336 (0.0906)	-0.0446 (0.0515)	-0.0162 (0.0559)	-0.0534 (0.0751)	-0.101 (0.0652)	0.0235 (0.0615)
Leadership	-0.0865 (0.0606)	-0.0626 (0.0675)	-0.0991* (0.0532)	-0.0298 (0.0824)	-0.110 (0.0855)	-0.0643 (0.0526)	-0.0884* (0.0530)	-0.0508 (0.0834)	-0.0484 (0.0639)	-0.110* (0.0617)
Merit×Leadership	0.0372 (0.0850)	0.0755 (0.0960)	0.110 (0.0756)	-0.0298 (0.116)	-0.0724 (0.122)	0.106 (0.0733)	0.0496 (0.0779)	0.0800 (0.111)	0.0511 (0.0914)	0.0837 (0.0864)
_cons	0.616*** (0.0721)	0.538*** (0.0775)	0.587*** (0.0639)	0.616*** (0.103)	0.618*** (0.0912)	0.530*** (0.0597)	0.563*** (0.0625)	0.561*** (0.129)	0.668*** (0.0737)	0.384*** (0.0820)
N	554	454	699	318	273	744	674	343	483	534
R <sup>2</sup>	0.050	0.012	0.034	0.038	0.083	0.034	0.032	0.037	0.040	0.033

Notes: OLS regressions on the dependent indicator variable R, taking the value of 1 if a spectator chose to restrict the stakeholder's choice, 0 if not. Column labels indicate the defining characteristic for the sample described in a given column. Male and Female are indicators for gender. LowAge and HighAge are indicators for being above and below the age of 45. LowEduc and HighEduc are indicators of not having or having a bachelor's degree equivalent or higher. LowInc and HighInc are indicator variables for having an income below or above \$60 000. LowRisk is an indicator for having a risk preference below the sample median, while HighRisk indicates a median risk preference or higher. Robust standard errors in parentheses  
\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table 4: Heterogeneity (contd.)

	(West)	(Midwest)	(Northeast)	(South)	(Democrat)	(Republican)	(Independent)
Merit	-0.146 (0.122)	-0.0287 (0.101)	-0.133 (0.0836)	0.0950 (0.0714)	-0.0968 (0.0631)	0.111 (0.0874)	-0.0475 (0.101)
Leadership	-0.213** (0.105)	-0.120 (0.101)	-0.0433 (0.0950)	0.00869 (0.0719)	-0.154** (0.0617)	0.0664 (0.0914)	-0.0806 (0.0973)
Merit×Leadership	0.231 (0.154)	0.000261 (0.147)	-0.0193 (0.125)	0.00340 (0.104)	0.189** (0.0869)	-0.0500 (0.128)	-0.0795 (0.137)
_cons	0.596*** (0.129)	0.729*** (0.108)	0.678*** (0.101)	0.482*** (0.0824)	0.545*** (0.0706)	0.345*** (0.110)	0.751*** (0.111)
N	188	186	261	382	539	249	213
R <sup>2</sup>	0.076	0.087	0.083	0.046	0.027	0.046	0.079

Notes: OLS regressions on the dependent indicator variable R, taking the value of 1 if a spectator chose to restrict the stakeholder's choice, 0 if not. Column labels indicate the defining characteristic for the sample described in a given column. West, Midwest, Northeast and South are indicators for US region. Democrat, Republican and Independent are indicators for political orientation. Included but not reported are all demographic control variables. Robust standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$



Table 4: Heterogeneity (contd.)

	(NoChildren)	(Children)	(NotMarried)	(Married)
Merit	-0.0937 (0.0667)	0.00582 (0.0603)	-0.0384 (0.0665)	-0.0318 (0.0607)
Leadership	-0.161*** (0.0614)	0.00256 (0.0642)	-0.0930 (0.0640)	-0.0619 (0.0629)
MeritXLeadership	0.0889 (0.0912)	0.0271 (0.0874)	-0.0334 (0.0909)	0.143 (0.0869)
_cons	0.721*** (0.0735)	0.453*** (0.0811)	0.665*** (0.0722)	0.473*** (0.0930)
N	472	545	476	541
R <sup>2</sup>	0.072	0.019	0.060	0.031

Notes: OLS regressions on the dependent indicator variable R, taking the value of 1 if a spectator chose to restrict the stakeholder's choice, 0 if not. Column labels indicate the defining characteristic for the sample described in a given column. Children and NoChildren are indicators for having and not having children. Married or NotMarried are indicators for whether a spectator is married. Included but not reported are all demographic control variables. Robust standard errors in parentheses  
\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table 5: Heterogeneity - Risk, Republican and Independent

	(1)	(2)	(3)
Merit×HighRisk	0.113 (0.0753)		
Leadership×HighRisk	-0.0743 (0.0751)		
Merit×Leadership×HighRisk	0.0104 (0.0169)		
Merit×Republican		0.192* (0.0995)	
Leadership×Republican		0.202** (0.101)	
Merit×Leadership×Republican		-0.166 (0.144)	
Merit×Independent			-0.0538 (0.108)
Leadership×Independent			-0.00876 (0.107)
Merit×Leadership×Independent			-0.133 (0.151)
_cons	0.598*** (0.0591)	0.616*** (0.0559)	0.565*** (0.0559)
<i>N</i>	1017	1017	1017
<i>R</i> <sup>2</sup>	0.033	0.034	0.032

*Notes:* OLS regressions on the dependent indicator variable *R*, taking the value of 1 if a spectator chose to restrict the choice of the stakeholder, 0 if not. Regressions reported in this table follow the empirical specification outlined in section 4.3.3. Included but not reported are controls for gender, education, income, age, risk preference, region and political orientation, as well as non-interacted treatment variables. Robust standard errors in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

---

## 6. Discussion

The goal of the experiment conducted for this thesis was to study whether people's willingness to implement a hard paternalistic intervention towards another person is affected by the extent to which they perceive their power over that other person as legitimate. In this chapter I discuss the results presented in chapter 5. In section 6.1., I discuss results relating to my main hypotheses. In section 6.2., I discuss a selection of results from the exploratory part of my analysis and offer suggestions as to how these may inform further research. Where appropriate throughout this chapter, I offer criticisms of decisions made for this study, as well as some suggestions on how this study can be improved upon in future research.

### 6.1 Discussion of Results Relating to the Main Hypotheses

Recall that in the conceptual framework presented in chapter 3, I denoted the fraction of spectators implementing the hard paternalistic intervention as  $\theta\pi$ , which I measure empirically through the dependent variable  $R_i$ . In the conceptual framework, I proposed two potential effect mechanisms with respect to legitimate power, which I referred to as the “guardian effect” and the “emboldening effect”. These two proposed effect mechanisms were thought to take on opposite directions with respect to my dependent variable  $R_i$ , with the guardian effect increasing the average willingness to intervene through an increase in  $\theta$  and the emboldening effect decreasing average willingness to intervene through a decrease in  $\pi$ .

My main results of interest relate to whether the two factors introduced in the experiment, merit and leadership, would have a significant effect on  $\theta\pi$  when applied either together or separately. I do not find a significant effect of being informed that one performed well relative to others on a set of complex tasks. I do however find a significant negative effect of being assigned a leadership role on willingness to intervene paternalistically.

This effect is estimated to -0.0788 in my complete regression model, which we interpret as a 7.88 percentage point decrease in average willingness to restrict the stakeholders choice associated with being assigned a leadership role. Tying this back to my conceptual framework, we can see this result as an indication that the emboldening effect dominates the guardian effect when it comes to the leadership treatment.

As seen in table 4, I find a significant negative effect of the leadership treatment also when running regressions only on the sub-sample of spectators whose self-reported risk preference was median or above. However, the difference in leadership treatment effect between the LowRisk and HighRisk spectators is not statistically significant, as seen in the row marked “Leadership×HighRisk” in column (1) of table 5.

Significant leadership treatment effect differences when comparing spectators with an above average risk tolerance and spectators with a below average risk tolerance could have been seen as an indication that the overall negative effect of the leadership treatment was driven by something akin to the proposed emboldening effect. I am not able to identify such a result in my data, however.

This brings us to a weakness in my design, which is that I have few means of disentangling effects relating to my two proposed effect mechanisms. Thus, I only observe the overall effect of my treatments on willingness to implement the hard paternalistic intervention. I am therefore not in a position to identify whether and to what extent the proposed emboldening effect is driving the negative overall effect of the leadership treatment. I am also not in a position to study if and how the proposed guardian effect mechanism is affecting results. My data suggests however that the former effect dominates the latter with respect to the leadership treatment. A suggestion for further research would be to attempt to test for these hypothesised effect mechanisms separately.

In summary, I find a negative effect of the leadership treatment on spectators’ willingness to implement the hard paternalistic intervention towards a stakeholder. This effect could be explained by an increase in the fraction of spectators disregarding the stakeholder’s preferences in favour of their own. Further research is required to determine whether this proposed mechanism is driving the results seen in this study.

## 6.2 Discussion of Exploratory Results

### 6.2.1 Risk Preferences

I find significant correlation between risk preferences and willingness to implement the hard paternalistic intervention, as seen my main regression model in column (6) of table 3. In this model, spectators with a lower than median self-reported risk preference are estimated to be

---

7.89 percentage points more likely to implement the hard paternalistic intervention than spectators with a median self-reported risk preference or higher.

Recall that spectators who make intervention decisions based on their own preferences rather than the stakeholder's stated preference are defined as authoritative. Significant negative correlation between spectators' risk tolerance and their willingness to implement the hard paternalistic intervention can be seen as an indication that there are authoritative spectators present in my sample.

## 6.2.2 Political Orientation

Looking at coefficient estimates relating to indicator variables for political orientation in the complete regression model presented in column (6) of table 3, I find that democrat spectators have a significantly higher willingness to restrict the stakeholder's choice than both the republican spectators (estimated 9,21 percentage point difference,  $p = 0.018$ ) and independent or third-party spectators (estimated 10,3 percentage point difference,  $p = 0.012$ ) in my sample. These results suggest differences across party lines in terms of weight placed on welfare relative to personal freedom. Tying this back to the conceptual framework presented in chapter 3, we can say that democrat spectators show signs of having a lower average value of  $\alpha$  than republican and independent spectators. I also find significant treatment effect differences between democrat and republican spectators with respect to both the merit treatment and the leadership treatment, as seen in column (2) of table 5.

The connection between politics and paternalistic preferences is particularly interesting, as government policies that restrict the freedom of citizens justified by concerns for welfare are oftentimes controversial. (Le Grand & New, 2015; Thaler & Sunstein, 2020) Paternalistic preferences can therefore play a role not only in politics in terms of elections, but also in determining the political cost of enacting certain policies for those who hold office. Examples of political debates where paternalistic preferences play a role range from public health insurance to gun legislation, from sugar taxation to marijuana legalization, from euthanasia to vaccines.

If paternalistic preferences with respect to one's own intervention decisions is any reflection of paternalistic preferences when it comes to government policy, these findings suggest that US citizens identifying as democrat may be more willing to accept paternalism in public policy than republican and independent voters. These findings also raise interesting questions with

regards to how the factors that determine acceptance of freedom restricting interventions may vary across political parties. A potentially interesting avenue for further research could therefore be whether characteristics seen as appropriate for giving someone a right to govern and prescribe behaviours for others are conditional on political beliefs.

### **6.2.3 Control Group Intercept**

As presented in chapter 5, I find a highly significant difference in willingness to implement the hard paternalistic intervention between my control group and spectators choosing whether to implement a hard paternalistic intervention in Bartling et al. (2022). In light of this result, we must assume that one or more characteristics of my sample or my experiment design is driving this increased willingness to restrict the stakeholder's choice that we observe in my control group.

The two main differences between my sample and the large, nationally representative sample analysed in Bartling et al. is (1) that spectators in my sample were selected based on their aptitude for solving mathematical word problems, and (2) that spectators in my sample were all users of the online work platform Amazon Mechanical Turk. As mentioned in chapter 4, samples obtained through Amazon Mechanical Turk tend to be younger and to have a higher average educational attainment than the general population in the US. (Chandler et al., 2019) As shown in the descriptive statistics in chapter 4, this also applies to my sample. It is important to remember that these observed differences between my sample and the general US population could be accompanied by unobserved differences in the error term.

It is also worth noting that my spectators all solved mathematical word problems just prior to deciding whether to restrict the choice of the stakeholder. I cannot rule out the possibility that this activity might have affected paternalistic choices in my control group, for example that solving mathematical word problems under time pressure could induce a state of mind where decisions are made more quickly and decisively, and that this may carry over into the final part of the survey.

I am not in a position to pinpoint the reason or reasons why spectators in my control group show a significantly higher willingness to implement the hard paternalistic intervention than spectators in Bartling et al. (2022). One factor which may have contributed to these differences, is that I select spectators to be included in my sample based on their aptitude for solving mathematical word problems. Intercept differences between my project and Bartling

---

et al. (2022), when seen in relation to differences in sampling methods, point to interesting topics for further research, such as whether underlying traits that are correlated with mathematical word problem performance are also correlated with a higher willingness to implement a hard paternalistic intervention towards a stakeholder.

## 7. Conclusion

In this thesis, I have studied the effect of legitimate power on people's willingness to implement a hard paternalistic intervention towards another person. In an online experiment, participants in the role of spectators decided whether to implement a hard paternalistic intervention towards a stakeholder by restricting his freedom of choice to prevent him from making a mistake. Employing a 2×2 factorial, between subject design, I varied spectators' legitimacy along two dimensions: Whether they had received positive feedback on their performance on a set of cognitive tasks, and whether they had been assigned a leadership role in relation to the stakeholder prior to making their intervention decision.

I find that being assigned a leadership role leads fewer spectators to implement the hard paternalistic intervention towards the stakeholder. I do not find a significant effect of receiving positive feedback on performance on willingness to intervene, nor do I find a significant interaction effect between the two factors in my experiment.

My design does not allow for a causal interpretation with regards to why being assigned a leadership role decreases the average willingness to implement the hard paternalistic intervention by restricting the stakeholder's choice. That said, the observed negative effect of the leadership treatment is consistent with the "emboldening effect" described in my conceptual framework, where receiving a leadership role in relation to a stakeholder leads some spectators to disregard the stakeholder's preferences, instead making their intervention decision based on which outcome they themselves would prefer. In light of my results with respect to the leadership treatment, I encourage further research into potential connections between hierarchical position and authoritativeness in paternalistic decision making.

In summary, my findings suggest that seeing legitimacy as "right to restrict" a stakeholder's freedom of choice may not be sufficient in explaining how people respond to increased legitimacy when deciding whether to implement a hard paternalistic intervention towards a stakeholder. It could be that for some people acting as spectators, legitimacy is just as much a "right to prescribe" some outcome other than that which is the stakeholder's preference.



---

## References

- Ambuehl, S., Bernheim, B. D., & Ockenfels, A. (2021) What Motivates Paternalism? An Experimental Study, *American Economic Review*, 111(3), 787-830
- Almås, I., Cappelen, A. W., & Tungodden, B. (2020) Cutthroat Capitalism versus Cuddly Socialism: Are Americans More Meritocratic and Efficiency-Seeking than Scandinavians? *Journal of Political Economy*, 128(5) 1753-1788  
<https://doi.org/10.1086/705551>
- Bartling, B., Cappelen, A. W., Hermes, H., Skivenes, M., Tungodden, B. (2022) *Free to Fail? Paternalistic Preferences in the United States*. [Unpublished]
- Bartling, B., Fehr, E., & Herz, H. (2014). The Intrinsic Value of Decision Rights. *Econometrica* 82(6), 2005-2039, <https://doi.org/10.3982/ECTA11573>
- Chandler, J., Rosenzweig, C., Moss, A. J., Robinson, J., & Litman, L. (2019). Online Panels in Social Science Research: Expanding Sampling Methods Beyond Mechanical Turk. *Behavior Research Methods*. 51, 2022-2038, <https://doi.org/10.3758/s13428-019-01273-7>
- DellaVigna, S. (2009) Psychology and Economics: Evidence from the Field. *Journal of Economic Literature* 47(2), 315-372
- DeScioli, P. & Krishna, S. (2013). Giving to whom? Altruism in different types of relationships. *Journal of Economic Psychology*. 34 218-228  
<http://dx.doi.org/10.1016/j.joep.2012.10.003>
- Dworkin, G. (2020) Paternalism. *The Stanford Encyclopedia of Philosophy* (Fall 2020 Edition), Edward N. Zalta (ed.) URL: <https://plato.stanford.edu/entries/paternalism/>
- Dworkin G. (1983) Paternalism: Some Second Thoughts. In Rolf Sartorius (Ed.) *Paternalism*. (pp. 105-112). University of Minnesota Press.
- Fehr, E. & Fischbacher, U. (2003) The Nature of Human Altruism. *Nature*, 425, 785-791  
[10.1038/nature02043](https://doi.org/10.1038/nature02043)
- Fehr, E., Herz, H. & Wilkening, T. (2013) The Lure of Authority: Motivation and Incentive Effects of Power. *American Economic Review*. 103(4), 1325-1359

- French, J. R. P., Jr., & Raven, B. (1959). The bases of social power. In D. Cartwright (Ed.), *Studies in social power* (pp. 150-167). Univer. Michigan.
- Hellevik, O. (2009). Linear versus logistic regression when the dependent variable is a dichotomy. *Quality & Quantity*, 43, 59-74 <https://doi.org/10.1007/s11135-007-9077-3>
- Haworth, L. (1991) Review: Dworkin on Autonomy. *Ethics*, 102(1), 129-139
- Le Grand, J. & New, B. (2015) *Government Paternalism – Nanny State or Helpful Friend?* Princeton University Press.
- Mädche, A. (2019) How Effective is Nudging? A Quantitative Review on the Effect Sizes and Limits of Empirical Nudging Studies. *Journal of Behavioral and Experimental Economics*. 80, 47-58 <https://doi.org/10.1016/j.socec.2019.03.005>
- Mill, J. S. (1859) *On Liberty*. Pelican books, 1974 ed.
- OECD (2017) *Behavioural Insights and Public Policy: Lessons from Around the World*. OECD Publishing, Paris <https://doi.org/10.1787/9789264270480-en>
- Pew Research Center (2018) *Wide Gender Gap, Growing Educational Divide in Voter's Party Affiliations*, Pew Research Center.  
<https://www.pewresearch.org/politics/2018/03/20/1-trends-in-party-affiliation-among-demographic-groups/>
- Pikulina, E. S., & Tergiman, C. (2020) Preferences for Power. *Journal of Public Economics*, 185 <https://doi.org/10.1016/j.jpubeco.2020.104173>
- Pope, T. (2004). Counting the Dragon's Teeth and Claws: The Definition of Hard Paternalism. *Georgia State University Law Review*, 20(3), 659-722
- Raven, B. H. & French, J. R. P., Jr. (1958a). Legitimate Power, Coercive Power, and Observability in Social Influence. *Sociometry*, 21(2), 83-97 <https://doi-org.ezproxy.nhh.no/10.2307/2785895>
- Raven, B. H. & French, J. R. P., Jr. (1958b). Group Support, Legitimate Power and Social Influence. *Journal of Personality* 26(3), 400-409 <https://doi.org/10.1111/j.1467-6494.1958.tb01595.x>

- 
- Read, P. B. (1974) Source of Authority and The Legitimation of Leadership in Small Groups. *Sociometry* 37(2), 189-204 <https://doi.org/10.2307/2786375>
- Sandel, M. J. (2020). *The Tyranny of Merit*. Allen Lane.
- Schurter, K. & Wilson, B. J. (2009) Justice and Fairness in the Dictator Game. *Southern Economic Journal*. 76(1) 130-145 <https://doi.org/10.4284/sej.2009.76.1.130>
- Scoccia, D. (2008) In Defense of Hard Paternalism, *Law and Philosophy* 27(4), 351-385 <https://doi.org/10.1007/s10982-007-9020-8>
- Sen, A. (1976) Liberty, Unanimity and Rights. *Economica*, 43(171) 217-245 <https://doi.org/10.2307/2553122>
- Sen, A. (1983) Liberty and Social Choice. *The Journal of Philosophy*. 80(1), 5-28 <https://doi.org/10.2307/2026284>
- Thaler, R. H. & Sunstein, C. R. (2003) Libertarian Paternalism *American Economic Review*, 93(2), 175-179
- Thaler, R. H. & Sunstein, C. R. (2020) *Nudge – The Final Edition*. Allen Lane.
- Tyler, T. R. (2006) Psychological Perspectives on Legitimacy and Legitimation. *Annual Review of Psychology* 57, 375-400 <https://doi.org/10.1146/annurev.psych.57.102904.190038>
- US Census Bureau (2022, May 12<sup>th</sup>) *CPS Historical Time Series Tables*, United States Census Bureau, <https://www.census.gov/data/tables/time-series/demo/educational-attainment/cps-historical-time-series.html>
- Weber, M. (1922) *Wirtschaft und Gesellschaft: Grundriß der Verstehenden Soziologie*. (5th ed.) Mohr Siebeck, Tübingen.
- Weiß, M., Rodrigues, J., Paelecke, M. & Hewig, J. (2020). We, Them, and It: Dictator Game Offers Depend on Hierarchical Social Status, Artificial Intelligence, and Social Dominance. *Front. Psychol.* 11 <https://doi.org/10.3389/fpsyg.2020.541756>

## **Appendix A. – Ethical Considerations**

In this section, I discuss ethical considerations for this study. The reference document for this discussion are the guidelines from the National Committee for Research Ethics in the Social Sciences and the Humanities (NESH). As new guidelines are scheduled to be published in May of 2022, I clarify that the active guidelines when the survey experiment took place were those published on June 8<sup>th</sup> of 2019.<sup>2</sup>

An NHH Institutional Review Board application was submitted to the compliance officer at the FAIR Center for Experimental Research on Fairness, Inequality and Rationality on April 25<sup>th</sup> of 2022. The application was approved on April 28<sup>th</sup> of 2022 on grounds that the study falls under the category of studies described as non-invasive survey experiments in the framework agreement NHH-IRB 31/21.

### **Data Collection and Privacy**

The software used for running the experiment, Qualtrics, offers the option to anonymize data. This option was enabled so long as the survey was active, meaning that no personal information such as names, emails, IP-addresses, postal codes, etc. was collected. Participants were presented with a consent form before starting the survey, in which they were informed about how their data would be stored and used. Participants who consented to participate in the study were asked to indicate their gender, age group, educational attainment, income, region, marital status, number of children, and risk preference. “Prefer not to answer” was an available option with respect to gender and political orientation. It would not be possible to identify participants using information given for demographic control variables.

### **Decent Pay**

Payment to participants was administered through Amazon Mechanical Turk’s worker ID system. In the consent form, participants were informed that their worker ID would be used to administer payment. This is an identification code used to administer payment in such a way that participants recruited through Amazon Mechanical Turk remained anonymous. Spectators

---

<sup>2</sup> <https://www.forskningsetikk.no/en/guidelines/social-sciences-humanities-law-and-theology/guidelines-for-research-ethics-in-the-social-sciences-humanities-law-and-theology/>

took approximately 5 minutes to complete their stage, for which they were paid \$0.70. This comes out to an hourly rate of \$8.40. Stakeholders took approximately 4 minutes to complete their stage, for which they were paid \$1, with some stakeholders also receiving a bonus payment. Disregarding bonus, this comes out to an hourly rate of \$15.

### **Human Dignity in Research**

The NESH Guidelines for Research Ethics state that “Researchers must protect personal integrity, preserve individual freedom and self-determination, respect privacy and family life, and safeguard against harm or unreasonable strain.” Asking spectators to solve mathematical word problems was necessary to ensure that all information given as part of the merit treatment was true. I tried to be conscious about how problems with right and wrong answers could be stressful for some spectators and mitigate accordingly. Prior to solving the problems, spectators were informed that there were no minus points for wrong answers and that guessing was encouraged if they were not sure about the answer. Only spectators who got more than half of the word problems correct were informed about their performance. Spectators who were omitted from the analysis due to answering fewer than three of the five word problems correctly were not made aware of selection criteria used to determine who moved on to the final part of the survey. This was done to minimize strain on participants.

Preserving individual freedom and self-determination when studying hard paternalism can be difficult, since the intervention is defined by a restriction placed on someone’s liberty. The solution implemented here was to have spectators make their choices first, and then have those choices determine which set of instructions a stakeholder would receive, without the stakeholder being informed about the involvement of the spectator. This avoids a sense of liberty or autonomy loss on the part of the stakeholder.

## Appendix B. – Experimental Instructions

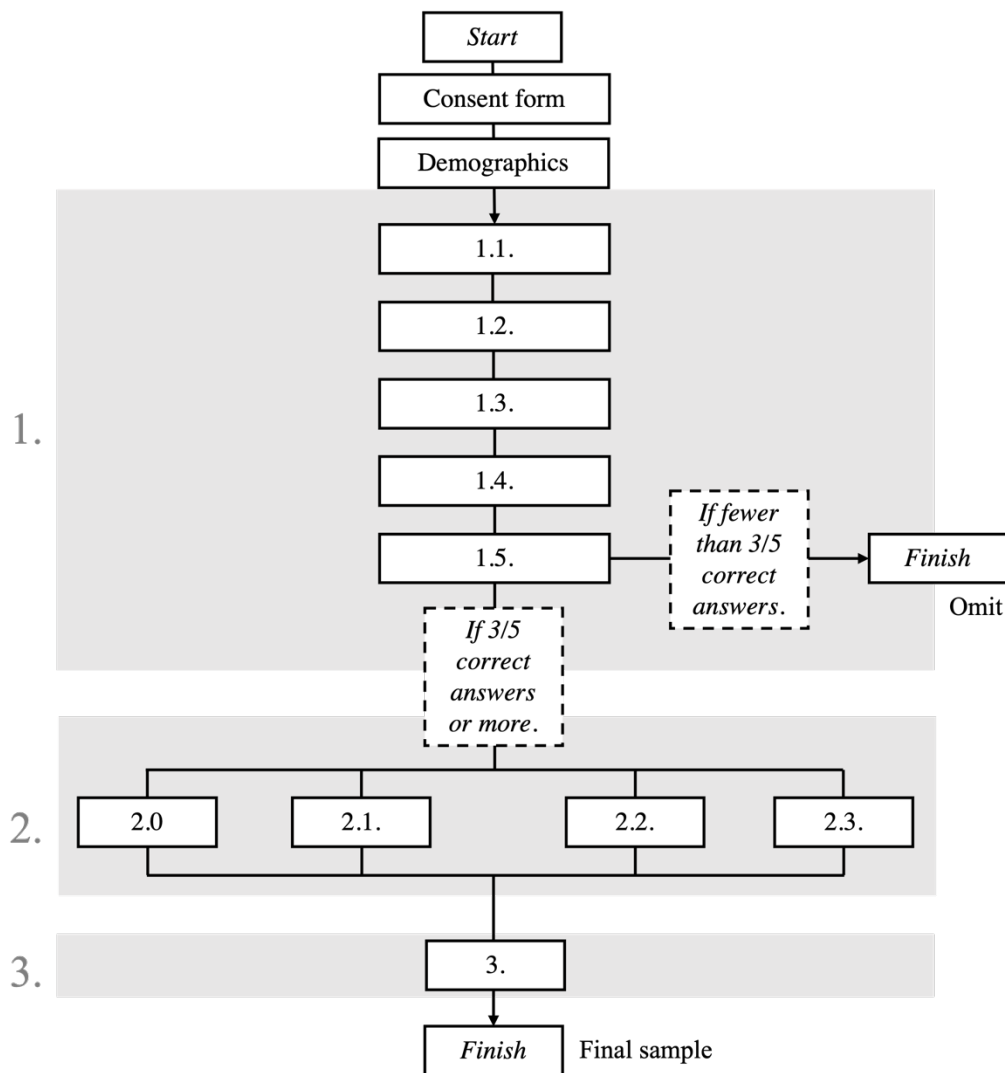


Figure B.1: Structure of spectator stage of experiment.

1. The first part of the experiment consists of five mathematical word problems. (1.1-1.5.) Respondents with 3/5 or more correct answers proceed to the next part of the survey. Respondents with fewer than 3 correct answers are omitted.
2. In the second part, remaining respondents are randomly assigned to one of four treatment conditions. (2.0-2.3.)
3. In the third part, respondents choose whether to implement a hard paternalistic intervention towards a real stakeholder.

**B.1.1. Mathematical Word Problems**

*Alternatives for each question were 1, 2, 3, 4 and 5. Time control was 45 seconds per question.*

*1.1.*

A sunflower is growing in a garden. Its height doubles each year. Last year it was 1 foot tall. This year it is 2 feet tall. How tall will it be next year (in feet)? (Correct answer: 4)

*1.2.*

Three strangers meet for the first time. How many handshakes does it take for everyone to have shaken hands with each other? (Correct answer: 3)

*1.3.*

It takes one machine one minute to make one gadget. How many gadgets do two machines make in two minutes? (Correct answer: 4)

*1.4.*

Hannah and Lucy are daughters of Bob and Mary. Hannah and Lucy have one sister each. How many daughters do Bob and Mary have in total? (Correct answer: 2)

*1.5.*

Alice is one of ten people competing in a bicycle race. Alice finishes the race with six people behind her. How many people finished before Alice? (Correct answer: 3)

## **B.1.2. Treatments**

### *2.0. Base treatment*

You will now be asked to make a decision that may have real consequences for another person.

### *2.1. Merit treatment*

You answered the majority of the questions correctly. Despite seeming simple at first, these questions are in fact quite complex and sometimes counterintuitive, which is why many people get them wrong. Your strong performance on these questions is an indication that you are good at analysing and solving complex problems.

You will now be asked to make a decision that may have real consequences for another person.

### *2.2. Leadership treatment*

You have been assigned the role of leader for the next part of the survey.

As a leader, you will now be asked to make a decision that may have real consequences for another person.

### *2.3. Combined treatment*

You answered the majority of the questions correctly. Despite seeming simple at first, these questions are in fact quite complex and sometimes counterintuitive, which is why many people get them wrong. Your strong performance on these questions is an indication that you are good at analysing and solving complex problems.

Because of your strong performance, you have been assigned the role of leader for the next part of the survey.

As a leader, you will now be asked to make a decision that may have real consequences for another person.



---

### B.1.3. Hard Paternalistic Intervention

*At the end of the experiment, spectators choose whether to implement a hard paternalistic intervention by restricting the stakeholder's choice:*

This other person was hired to do some work. After completing the work, the person was informed that he or she will get a bonus.

There are two bonus options available:

Safe option:	a bonus of 4 USD for sure
Risky option:	either a bonus of 10 USD or nothing, where the two outcomes are equally likely

When the person was informed about the two options, the risky option was not presented as in the table above. Rather, the person had to calculate the likelihoods of the two outcomes of the risky option. The person made a mistake in the calculations.

As a result, the person prefers the risky option. However, had the person calculated the likelihoods correctly, he or she would have preferred the safe option.

The person has not yet made a choice. You can now decide between two alternatives:

- Restrict choice: The person will not have the opportunity to make a choice and will receive the safe option.
- Do not restrict choice: The person will have the opportunity to make a choice between the safe and the risky option.

Remember that your decision is consequential. You and nine other respondents make this decision, and we will randomly select one of you to be the one whose decision will determine whether the person is paid the compensation. Your decision is anonymous and not observed by other respondents.

## B.2. Stakeholder Instructions

*In between other tasks, stakeholders were asked the two questions seen below. These questions are technically identical, but in the second iteration the risky option is presented in a way that makes many people overestimate the likelihood of having a winning ticket. Stakeholders who chose the safe option in the first iteration and the risky option in the second display the type of preference profile described to spectators. Only stakeholders who answered these questions in such a way were used for implementing spectator choices.*

Recall that at the end of this study, you may receive a bonus payment (depending on your choices). Suppose you could choose between the following two bonus options:

Safe option:	a bonus of 4 USD for sure
Risky option:	either a bonus of 10 USD or nothing, where the two outcomes are equally likely

Which of these two options would you prefer?

- Safe option
- Risky option

(...)

Recall that at the end of this study, you may receive a bonus payment (depending on your choices). Suppose you could choose between the following two bonus options:

Safe option:	a bonus of 4 USD for sure
Risky option:	This option gives you a ticket for a lottery. You win a bonus of 10 USD, if you have a winning ticket. A random ticket wins with a probability of 1%. However, your ticket was pre-tested and according to the pre-test it is a winning ticket. The pre-test correctly identifies winning and losing tickets in 99% of the cases.

Which of these two options would you prefer?

- Safe option
- Risky option

## Appendix C. – Demographic Distribution Across Treatment Groups

Table C: Demographic distribution across treatment groups

	Control		Merit treatment		Leadership treatment		Both treatments		Total
	n	%	n	%	n	%	n	%	n
<b>Gender</b>									
Male	130	23%	144	26%	145	26%	135	24%	554
Female	119	26%	109	24%	106	23%	120	26%	454
Non-binary / third gender	0	0%	2	40%	1	20%	2	40%	5
Prefer not to answer	1	25%	1	25%	2	50%	0	0%	4
<b>Age</b>									
18 to 34	96	26%	97	26%	95	25%	87	23%	375
35 to 44	82	25%	79	24%	76	23%	87	27%	324
45 to 54	34	22%	40	25%	47	30%	36	23%	157
55 to 64	22	20%	33	30%	26	24%	29	26%	110
64 +	16	31%	7	14%	10	20%	18	35%	51
<b>Education</b>									
Up to High School	22	26%	14	16%	27	32%	22	26%	85
Some College	50	27%	44	23%	40	21%	54	29%	188
Bachelor or Associate	137	24%	153	26%	149	26%	143	25%	582
Master or above	41	25%	45	28%	38	23%	38	23%	162
<b>Income</b>									
< \$30 000	71	28%	45	18%	68	27%	71	28%	255
\$30 000 to \$60 000	96	23%	109	26%	118	28%	96	23%	419
\$60 000 to \$100 000	64	26%	72	29%	42	17%	69	28%	247
\$100 000 to \$150 000	12	17%	22	31%	19	27%	18	25%	71
> \$150 000	7	28%	8	32%	7	28%	3	12%	25

Table C: Demographic distribution across treatment groups (contd.)

	Control		Merit treatment		Leadership treatment		Both treatments		Total n
	n	%	n	%	n	%	n	%	
<b>Region</b>									
West	36	19%	37	20%	65	35%	50	27%	188
Midwest	44	24%	46	25%	55	30%	41	22%	186
Northeast	68	26%	78	30%	50	19%	65	25%	261
South	102	27%	95	25%	84	22%	101	26%	382
<b>Marital status</b>									
Married	133	25%	140	26%	125	23%	143	26%	541
Not married	117	25%	116	24%	129	27%	114	24%	476
<b>Children</b>									
No children	112	24%	112	24%	137	29%	111	24%	472
1-2 children	117	25%	124	27%	101	22%	125	27%	467
More than 2 children	21	27%	20	26%	16	21%	21	27%	78
<b>Politics</b>									
Republican	62	25%	70	28%	59	24%	58	23%	249
Democrat	135	25%	129	24%	133	25%	142	26%	539
Independent / Third party	52	24%	53	25%	57	27%	51	24%	213
Prefer not to answer / Do not know	1	6%	4	25%	5	31%	6	38%	16