Norwegian School of Economics

Master's program in economics and administration
Economics profile

Master's thesis

Freedom or Efficiency?
An Experimental Approach to Regulation

Olve Hagen Wold

supervised by
Alexander W. Cappelen

Spring semester, 2017

Abstract

Whether and how to regulate externalities and public goods is one of the most central issues of political economy. It is also one of the areas with the largest discord between economic theory and political reality. This paper examines the issue from a new angle, using an economic experiment to elicit revealed norms about regulation. It shows that regulatory preferences are highly sensitive to potential efficiency gains. It also reveals that norms about regulation are highly heterogeneous, and that a large minority are apparently unwilling to trade reductions in autonomy for any increase in efficiency.

Acknowledgments

Thanks to my supervisor, Alexander W. Cappelen, for invaluable help with designing, interpreting and presenting the results of the experiment, and for steering me back when I had gone too far off the tracks.

Thanks to Sebastian Fest for polishing and executing the experiment; I am sure the experiment would have been less effective without your input.

Thanks to The Choice Lab, for believing in and supporting this project financially.

Thanks to Mom and Dad, for always being there and for relentless moral support.

Thanks to Peder Engesæth, for being a reliable sparring partner, and for pushing me when I doubted if I would be able to finish on time.

Thanks to my current and former roommates, Sunniva, Vilde, Eirik, Ragnhild, Shagha, Mathias and Eirik, for putting up with me these last months.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

An enduring problem of economics is that of externalities. When economic activity causes costs or benefits that are not borne by the party responsible for the activity, a free market is inefficient. This simple insight is the root of many, if not most, of the problems of political economy, such as the (mis)management of natural resources, road congestion, pollution, pharmaceutical research and arts financing. And for as long as the problem has been known, economists have told us that the answer is government regulation.

Still, the problems persist. Regulation often faces massive ideological opposition, and in spite of the advice of economists, regulatory solutions are often not implemented, or not implemented in a way that would actually lead to an efficient outcome. This ideological opposition is mostly grounded in the view that regulations infringe on individual liberty.

The purpose of this paper is not to settle this debate, but to examine experimentally what people's revealed preferences about regulation are. Which trade-offs between efficiency and liberty are people willing to accept? And what accounts for the differences between people?

To answer these questions, I conducted a controlled experiment online using 400 participants located in the US, recruited using Amazon Mechanical Turk. I used a simple cooperation game to simulate a situation where externalities lead to market failure. Specifically, I used a standard public goods game with four players, equal initial endowments and a multiplier of 2. I then gave a group of impartial spectators the opportunity to choose that the players should receive a fixed payment instead of the earnings from the public goods game. I used the same sample for both parts of the experiment, with every participant making decisions for both parts, and roles being allocated later.

The use of spectators builds on previous research in experimental economics. In for instance Erkut, Nosenzo, and Sefton (2015) spectators are used to elicit unbiased fairness considerations in a dictator game. Cappelen, Konow, Sorensen,

and Tungodden (2010) and Mollerstrom, Reme, and Sorensen (2014) use similar designs to examine norms about different forms of inequality. All suggest that spectator choices are a reliable indicator of impartial norms. To my knowledge, such a design has never been used to examine cooperation games previously.

The spectator choice is intended to be analogous to government intervention. The spectators had the opportunity to intervene and choose an outcome that was most likely more efficient, but in order to do this, they had to overturn the decisions made by others, without their prior consent or knowledge.

In reality regulations are not perfect. They are costly to design, implement and enforce, and they are seldom designed in a way that completely avoids distortions and inefficiencies. The experiment used two treatments, intended to represent perfect and imperfect regulation. In the high treatment, spectators could choose a fixed payment equal to the optimal outcome of the game. In the low treatment, the fixed payment was less efficient than the optimal outcome, but more efficient than the Nash equilibrium and the actual average.

The average earnings from the public goods game were 1.35 USD. This means that the fixed payment would lead to an expected efficiency increase in both treatments, but that the expected gain was much larger in the high treatment.

In the high treatment 80 percent chose to intervene, while in the low treatment 52.5 percent chose the fixed payment. Both these numbers are interesting in their own right. The first indicates that a large share of the population exhibit strong anti-regulatory or anti-paternalistic sentiments, choosing not to overturn the choices of others even when this would almost certainly increase their welfare. The second shows that a majority prefer intervention even when there is a substantial potential efficiency loss.

The difference between the treatments was large and highly significant, which provides strong evidence that intervention rates were highly sensitive to differences in expected efficiency gains. There was a significant political effect: self-described conservatives were less likely to intervene in the high treatment, but not in the low treatment.

Near the end of the experiment, the subjects were asked a set of questions about their beliefs about the results. These provided evidence that efficiency is by far the most important factor when determining whether or not to intervene. The overwhelming majority made a decision that was consistent with what they believed would give others the highest return. The results also suggested that differences in beliefs, rather than differences in preferences, was the main explanation for the treatment difference.

Subjects were also asked to give a written explanation for their choice, which most of them did. A qualitative analysis of these gave results were consistent with the quantitative results. Efficiency was the most commonly cited motivation both

for interveners and non-interveners. For interveners, this revealed that equality and uncertainty avoidance were also important motivations. For non-interveners, the responses were consistent with a constant proportion of the population that are totally opposed to interference with the choices of others.

As this is the first experiment of this kind, another important purpose of this article is to serve as a proof of concept. The findings suggest that this is an effective method of eliciting preferences. The choices and responses strongly suggest that participants make active choices that reflect what they believe is in the interest of others, and the findings are consistent both with each other and existing beliefs about regulatory preferences.

The article is structured as follows. Chapter 2 gives a background on the problems of public goods and open access, and their relation to issues of paternalism and autonomy. Chapter 3 gives a detailed description of the experiment design and procedure. Chapter 4 provides a short discussion of the results of the public goods game. The main analysis of the choices made by spectators is found in chapter 5. Chapter 6 provides additional qualitative analysis of these choices. Chapter 7 concludes, discusses policy implications and makes suggestions for further research.

Chapter 2

Background

2.1 The tragedy of the commons

The tragedy of the commons was first described in Lloyd (1980), but popularized by Hardin (1968). Hardin based his argument on the original example used by Lloyd of a pasture which is shared in common by a group of herders. Each herdsman will receive all the proceeds from the sale of the cattle, but only share a fraction of the negative utility from overgrazing. Every rational herder will keep adding cattle until the inevitable consequence: total ecological collapse and, presumably, death by starvation.

Hardin then expands this argument to reach his general Malthusian conclusion: rational action will lead to unlimited consumption and breeding. (Malthus 1798) In a world with limited resources, "[f]reedom in commons brings ruin to all". (Hardin 1968) The only way to avoid disaster is by managing the commons, either through "[...] socialism or the privatism of free enterprise". (Hardin 1998)

Modern economics have refined and moderated the understanding of the tragedy. The updated model no longer necessarily spells doom, but the framework and implications otherwise remain the same.[1] In academic contexts, the problem is normally instead referred to as "the problem of open access", with open access being defined as a situation with no regulation of economic activity or market access.(Conrad 2010)

The problem is typically illustrated using fisheries. Assuming that the marginal cost of fishing falls with the size of the fish population, and that the price is high enough for some fishing to be profitable, fishing will continue until either the price is equal to the marginal cost or the fish become extinct, in either case driving the value of the fish stock to zero. (Conrad 2010) The result is bad from from both

[1]The problem with Hardin's thought experiment is the assumption of a constant marginal cost of raising new cattle.

an economic and an ecological perspective.

The problem has implications for much more than open-access resources. As noted by Hardin (1968), over-pollution is just another version of the same problem. Modern microeconomics brings these problems together under the headlines of "excludability", meaning that others can be excluded from the effects of consumption, and "rivalry", meaning that one party's consumption limits the consumption of others. (Kolstad 2009)

Non-excludability is of course just externalities from a different vantage point. (Cowen 1993) The general problem is that when there are positive or negative externalities, all costs and benefits are not internalized by the decision-maker, and there is consequently too little or too much economic activity. Problems of public goods, open-access resources and pollution are variations over this theme. With public goods, positive externalities lead to underprovision. With both open-access resources and pollution, which is a public bad, negative externalities lead to overconsumption. (Kolstad 2009)

Coase (1960) shows that the market will find an find an efficient solution assuming that well-defined property rights exist and there are no transaction or bargaining costs. Obviously, this way of assuming away the problem is not applicable to many real-world situations, especially not complex global problems such as climate change.[2] The alternatives to self-regulation are direct regulation (eg., non-tradable quotas, provision of public goods through taxation, nature reserves), and market-based measures (eg., Pigouvian taxes and subsidies, tradable quotas), with the latter usually being preferred by economists. (Kolstad 2009)

Ostrom, Burger, Field, Norgaard, and Policansky (1999) and Ostrom (2008) criticize the traditional view of commons, pointing to many examples where communities have found ways to voluntarily manage common property in a way that is mutually beneficial. Ostrom et al. (1999) further discusses factors that are likely to make sustainable management less likely, such as a large number of actors, cultural diversity, complex interlinkages and requirements of unanimity.

Cowen (1993) argues that many of the public goods which are thought of as unattainable through voluntary provision, in fact have been supplied by private enterprise throughout history. This was made profitable by bundling the public good together with private goods.

2.2 Modeling commons

A simple model, knows as the public goods game, captures the key aspects of the problems of externalities. The general form is as follows. n players are given a

[2]Coase did not intend to create a realistic model of self-regulation, but a benchmark to explain why markets fail.

certain amount of money as an initial endowment. They are then simultaneously given the opportunity to contribute any amount from their endowment to a common pool. Whatever is contributed to this pool multiplied by a factor m, where $1 < m < n$, and divided equally among all the players. Whatever they did not contribute, they get to keep. (Wilkinson & Klaes 2012)

The payoff to a player i is then given by

$$E_i + \frac{(m-n)X_i}{n} + \frac{m\sum_j X_j}{n},\ i \neq j, \tag{1}$$

where E_i and X_i are the initial endowment and contributions of player i respectively. Since $m < n$, the factor $\frac{m-n}{n}$ must be negative.

Contributing to the common pool can thus be viewed as an investment with a negative return. It follows intuitively that as long as a player has no way of influencing what the other players contribute, any contribution they make will decrease their earnings. Contributing nothing is the dominant strategy, and the game has a single equilibrium where nobody contributes anything and everybody's payment is equal to their initial endowment.

The total payment to the group is given by

$$\sum_i E_i + (m-1)\sum_i X_i. \tag{2}$$

Since $m > 1$, the factor $m - 1$ must be positive.

From the point of view of a social planner, contributing to the common pool is an investment with a *positive* social return. In other words, there are positive externalities. Any contribution will increase the total welfare of the group. The social optimum must therefore be when everybody contributes their entire endowment, in which case each player earns $\frac{m}{n}\sum_i E_i$. With equal initial endowments, this simplifies to mE.

Under standard assumptions, this means that the unregulated solution is inefficient, and leads to a large efficiency loss compared with the social optimum. It belongs to a larger class of cooperative games where cooperation would make everyone better off, but fails because players have individual incentives to defect and no way to coordinate strategies or make binding commitments.

The game has a similar payoff structure to a prisoner's dilemma, and can essentially be viewed as multiple-player prisoner's dilemma games. (Wilkinson & Klaes 2012) In fact, if you reduce the number of players to two (using a multiplier of for example 1.5), and only allow pure cooperation and defection, a public goods game reduces to a prisoner's dilemma. Equivalent to the temptation and sucker's payoff in a prisoner's dilemma, the best possible outcome of a public goods game comes from contributing nothing when others contribute their entire endowment,

and the worst possible outcomes come from contributing contributing the entire endowment when others do not contribute.

The game is obviously equivalent to real-world public goods problems.[3] However, it describes open-access resources and public bads just as well.

To see this, we can turn the public goods game on its head. Instead of an initial endowment, there is an initial common pool, and instead of contributing, players have the opportunity to take up to a given amount of money out of the pool at a proportional cost. This game is in fact identical to a public good game except for the framing. We can transform the game into a public bads game in a similar way.[4]

That the game is both highly relevant and easy to implement, has made it a staple of experimental economics. (Wilkinson & Klaes 2012) Generally actual contributions tend to be higher than what standard game theory predicts. Ledyard (1995) surveys key experiments and finds contribution rates between 0.4 and 0.6 in one-shot or first iterations of repeated games. (Sally 1995) combines 130 treatments from public goods and prisoner's dilemma games and finds a mean contribution rate of 0.474, with a majority of treatments having a contribution rate between 0.2 and 0.5.

Experiments with repeated games show that contributions tend to start at similarly high levels, but then fall toward a lower equilibrium, in which most people contribute nothing, and a minority continue to contribute.(Andreoni 1988; Gunnthorsdottir, Houser, & McCabe 2007; R. M. Isaac & Walker 1998) R. M. Isaac and Walker (1988) and Isaac, Walker, and Williams (1994) found no evidence for the hypothesis that larger groups decrease cooperation; Lipford (1995) provides further evidence from a non-experimental setting.

There are two types of explanations for the divergence between game theory and evidence. The first is that subjects have prosocial preferences, ie., that they take the interests of others into account when making decisions. The other is that sukjects are selfish, but do not choose the dominant strategy because of mistakes, which is suggested by the falling contribution rates in repeated games.

Ledyard (1995) argues that results are consistent with subjects making tradeoffs between self-interest, the interest of others and other factors, but that altruism cannot by itself explain the divergence. Kümmerli, Burton-Chellew, Ross-Gillespie, and West (2010) creates a version of the game where the multiplier (5) is larger than the number of players (4), which aligns the interests of individuals and the group. The study finds that contribution rates are still smaller than 1,

[3]More complex models, however, tend to use non-linear utilities, leading to an interior solution where *some* amount of the public good is provided.

[4]While the games are identical from a game-theoretic point of view, experimental results depend heavily on framing. (Sonnemans, Schram, & Offerman 1998; Willinger & Ziegelmeyer 1999)

and argues that aversion to extreme strategies must form part of the explanation. Bayer, Renner, and Sausgruber (2013) finds that confusion cannot by itself explain cooperative behavior, while Andreoni (1995) finds that about half of cooperation can be explained by prosocial preferences.

Sally (1995) finds that the most important factor for explaining differences between contribution rates in different treatments is whether subjects are permitted to communicate before the experiment. This could suggest that the fundamental problem is not misaligned interests but coordination, though there are also other interpretations.

2.3 Collective paternalism

Should governments intervene to mitigate externalities? The question leads back to the most fundamental question of political philosophy: what is the legitimate use of force? One of the most common answers, dating back to Hobbes (1651) and expanded by other social contract theorists, is that the legitimate role of government is to provide security, criminal justice, roads, and other goods that people are unable to voluntarily provide for themselves. The core function of government is in other words to provide public goods. This view is accepted by most political theorists. (Anomaly 2015)

As for negative externalities, some libertarian theorists, such as Nozick (1974) and von Mises (1949), merely postulate that with a system of properly defined property rights, those causing externalities would have to compensate the affected parties for misusing their property, and the problem would solve itself. The opposite view is the one already discussed, represented by Hardin (1968), that negative externalities require a radical reshaping of the economic system.

Outside of these extremes, there is limited discussion about the morality of regulating commons. One exception is Schmidtz (1991), who argues that government intervention to provide public goods represents a form of paternalism, and favors voluntary organization. I would agree with the paternalism part. The problems of paternalism and of externalities share many characteristics, and I would argue that it is a useful perspective to view it as individual and collective versions of the same problem.

Probably the most commonly used definition of paternalism is from Dworkin (2017). Dworkin argues that paternalistic actions have three defining characteristics: they are intended to increase the welfare of another person, they interfere with their autonomy, and they are done without their consent. The role of autonomy dates back to Kant (1785) and Mill (1859), who were both categorically opposed to any form of paternalistic intervention.[5]

[5]They would however have differed with regards to regulation of public goods. Kant was

The traditional premises of the debate were challenged by Thaler and Sunstein (2008). Building on insights from behavioral economics (Ariely 2009; Gneezy, List, & Levitt 2013; Kahneman 2011), Thaler and Sunstein argue that there is no neutral way to present a choice, but that what they term the "choice architecture" will necessarily influence the outcome. Based on this, they claim that choice architects have a moral obligation to present choices in a way that promotes the wellbeing of the decision-maker. They term this "libertarian paternalism", or more colloquially, "nudging".

Based on this new line of thinking, Le Grand and New (2015) argue for an updated definition of paternalism. Instead of being a defining characteristic, coercion is an accidental quality of many, but not all, paternalistic actions. They argue that paternalism is really about what they call "reasoning failure". When an individual is unable to reach their objectives because of failures of rationality, the government is justified in helping them align their actions with their preferences.

If we employ the traditional definition, at first glance, regulating externalities seems to fit. This is the position taken by Schmidtz (1991), but it is rejected by Anomaly (2015), who claims that "[...] if each of us recognizes an end as beneficial, but we lack the power to bring it about without the force of law, a law that compels us to do our part to achieve the end is not paternalistic [...]". To me, it seems like the exact same argument could be made against any form of paternalism. If an alcoholic recognizes the benefit on restrictions on the sale of alcohol, does it then not count as paternalism?[6] In either case, using coercion to promote the welfare of an individual seems very similar to using it to promote the welfare of a group.

This brings us to the second definition. If we follow Le Grand and New (2015), accordance between the ends of individuals and the force of law is not coincidental, but the essential feature of paternalism. Regulatory action that seems to increase the welfare of an entire group then seems to qualify. The difference is that the tragedy of the commons is not a result of reasoning failure, at least not as viewed in behavioral economics. Rather, it is the failure of individual rationality, leading to a form of collective irrationality.

I would argue that regulating externalities does not fit the definition of Le Grand and New. It is however a close analogy. If paternalism is justified on an individual level, I would suggest that the same ethical considerations apply for the regulation of commons. By the same note, I would suggest that the results from this experiment also provide insight into norms about paternalism.

---

strictly opposed to any restriction of autonomy. Mill's view was that the sovereignty of an individual should be unrestricted insofar as an action only affected himself, but not if it affected others, which externalities do.

[6]For a discussion of subsequent gratitude as a possible justification for paternalism, see Kasachkoff (1994).

Chapter 3

Experimental design

The purpose of the experiment is to elicit norms about regulating public goods and common-pool resources. To do this, I use a public goods problem to simulate such a situation. This is described in more detail in section 3.1. I then give a group of spectators the opportunity to regulate the game. After the players had made their choices, but before the results of the game were revealed, the spectators were presented with a simple choice: either they could let the players receive payments based on the results of the game, or they could choose that they instead receive a fixed payment.

The main hypothesis was that the spectators were motivated by outcome considerations, including efficiency, on one hand, and by autonomy concerns on the other. To test this hypothesis, I created two treatment groups that represented different trade-offs between autonomy and efficiency. In the first treatment, the fixed payment option was equal to the optimal outcome of the public goods game. In the second treatment, the fixed payment was significantly smaller, while still representing an efficiency gain compared to the Nash equilibrium.

The fixed payment options in the two treatments thus had different potential efficiency gains, but the same potential loss of autonomy. If the participants make a trade-off between efficiency and autonomy, we would then expect the proportion that chose to intervene to be smaller in the low treatment. The treatments and the motivation behind them are described in more detail in section 3.2.

The motivation for using spectators was to take self-interest out of the consideration. Because the spectators are not affected by the choices they make, their decisions should only reflect what they believe is the right outcome. As the public goods game is analogous to real-world problems, spectator choices should be an unbiased indicator of the population norms about regulating commons.

The public goods game was a real game involving actual money. A much cheaper alternative would have been to describe a fictional game to the spectators and ask what they would have chosen in this situation. However, making the

game real means that the choices the spectators make have actual consequences for others. This has two important advantages.

First, it makes the spectators care more about the result, and make more deliberate choices. Second, if the game was not real, there is a possibility that participants may base their choices on factors unrelated to the outcome of the game, and that their expressed preferences may not reflect their actual preferences. These factors makes the choices from this experiment a much better indicator of actual norms in the population than one using a fictional game.

The two phases were conducted simultaneously and using the same sample. Each participant made a decision in both parts of the experiment, but only one phase was drawn and implemented for each participant. The participants were informed of this at the start of the experiment, but were not given any further information. The primary advantage of this set up was that it made it possible to compare the decisions made by the same participant in the different phases.

The public goods phase was conducted before the spectator phase. One reason was given above, but more importantly this was to avoid that knowledge of the spectator phase would influence choices in the public goods game. This would not in itself be a major issue, the concern was rather strategic interaction with the spectator phase: if players could react to anticipated spectator choices, the spectators could also have to take these reactions into account, making feedback effects possible.

3.1 Public goods phase

I used a public goods game because this has the same structure, and thus represents the same fundamental problem, as real-world public goods and open-access resources problems. A large potential efficiency loss provides strong reason for interference by a social planner. However, with imperfect regulation, there is a possibility that the players might achieve a more efficient outcome if left alone.

The exact set up was as follows. There were four players per group. Each player received an initial endowment of 100 points, where 100 points was equal to 1 USD. They were then told to allocate any number of points between 0 and 100 to a common pool. Whatever was put in this pool was doubled and divided equally between all four players. In order to avoid a default effect, participants had to manually type in the number of points they wished to contribute.

This means that each player would reduce their earnings with 0.5 points per point contributed, but the total earnings for the group would increase with 2 points per point. The dominant strategy for each player was to keep the entire endowment, so the Nash equilibrium was no cooperation, in which players would earn 100 points each. The social optimum was full cooperation, in which players

would earn 200 points each.

The best possible individual outcome was to be the only player defecting, which would earn 250 points, and the worst possible individual outcome was to be the only cooperator, which would earn 50 points. There was a difference of 400 points between the most and the least efficient outcome, ie., there was a potential efficiency gain of 1 USD per player from cooperation.

Using a public goods game for this part of the experiment has several advantages over other types of cooperative games, such as trust games and prisoner's dilemma games. First, it provides a clearer analogy with real-world externality problems, and the results are therefore more translatable to real-world regulatory decisions. Second, there is a virtual continuum of possible strategies, and hence a large and varied set of potential outcomes. This makes the potential gain from intervening in a specific group less certain.

3.2 Spectator phase

The spectator phase was the crucial part of the experiment. In this part, the participants were told that they could decide the payment another group of participants would receive. They were presented with a choice between two options: the other participants could either be paid based on their choices in the previous part, or they could receive a fixed payment. Participants were explicitly told that the choice they made in this part had no impact on their own payment.

I used two treatment variations, a "high" treatment representing perfect regulation, and a "low" treatment representing imperfect regulation. These treatments were identical except for the size of the fixed payment option, which was 200 points in the high treatment and 150 points in the low treatment. These treatments had the same potential autonomy costs, but different potential efficiency gains. The main hypothesis therefore implies that the proportion of interveners should be smaller in the low treatment.

There were important reasons that the treatments had fixed payment options of exactly 200 and 150 points. Both options were designed to be attainable, ie., within the set of possible outcomes of the game. This strengthens the analogy to regulation, as the interventions can be viewed not as choosing payments, but rather deciding contributions. (This, of course, ignores framing effects.)

The high treatment option was exactly equal to the social optimum, ie., the outcome if all players contributed their entire endowment. It was at least as efficient as all possible outcomes, and almost certain to be more efficient than the actual outcome of any given game. It was however within the bounds of the game, and it was possible for individual players to earn more than 200 points, at the expense of other players.

There is thus no possibility that intervention in this treatment will make the group worse off, but there is a possibility that it might hurt individual players. This makes the proportion of non-interveners in this treatment a good estimator for the share of the population that care *only* about autonomy, and not at all about efficiency or equality.

The low treatment option was of 150 points was chosen as a natural midway point between the social optimum, of 200 points per player, and the Nash equilibrium, of 100 points per player. It was also exactly equal to the outcome when every player contributes half of their initial endowment.

Based on previous research on public goods games, (Ledyard 1995; Sally 1995) I could be reasonably certain that the actual average contribution would be between 0 and 50 points. This would mean that intervention in the low treatment would lead to a small expected efficiency gain, but there would still be a high probability that a given group would achieve a more efficient outcome.

To make the effects of autonomy concerns more visible, the experiment was to designed in order to make the cost of intervention as large as possible for the groups that were subjected to intervention. There was no way for the spectator to communicate with the spectated, and hence no way to receive explicit consent. Furthermore, when participants made their choice in the public choice phase, they had no knowledge that their choice might be overturned by a spectator. Finally, spectators did not know the outcome of the game they were spectating. This means they could only guess what the likely outcome would be, and even if they guessed correctly, the outcome might be better in the actual game they were spectating.

A complementary hypothesis is that spectators have a preference for equality. Since both fixed payment options are equal, this would increase the probability that a given spectator chooses to intervene. It is therefore likely that preferences for equality accounts for part of the share of interveners. The treatments are however not designed to test this hypothesis explicitly. Since both fixed payments are perfectly equal, and the distribution of individual outcomes in the unregulated game is independent of the treatment, a preference for equality could not explain a difference between treatments. The hypothesis will however be examined indirectly in other parts of the analysis.

Choosing the payment the spectators would receive presented a conundrum. If they were not given a bonus, there would be a risk that some spectators would "retaliate" by actively trying to decrease payments to the group they were spectating. However, any stated bonus amount would serve as a basis for comparison with the earnings of the players, which could lead to bias.

To mitigate both these problems, spectators were informed that they would receive a bonus, but the amount was not disclosed until the end of the experiment.

The actual amount they received was equal to the average earnings of the players.[1]

With this design, there is substantial risk of a status quo effect. (Wilkinson & Klaes 2012) Because the choices in the public goods game have already been made, there is a possibility that some spectators might choose not to intervene simply because it is the option they perceive to make the least change, independently of what they might think about the options.

I took several measures to reduce this risk. Instead of writing that the players would be paid according to the game *unless* the spectators chose a fixed payment, I used formulations that presented both options as active choices, and avoided formulations that explicitly presented the options as active intervention and passive non-intervention. The options were presented in random order and required active choice.

3.3 Sample and procedure

Sample The experiment was conducted online on 400 participants[2] who were located in the US.[3] The participants were recruited via Amazon Mechanical Turk, an online labor market commonly known as MTurk. The participants were exactly evenly divided between the two treatments.

I chose to run the experiment exclusively on US subjects for three reasons. First, the US is the largest national group on MTurk (Ipeirotis 2010), so this let me ensure that the experiment was conducted as quickly as possible. Second, it minimized the risk of confusion due to language barriers. Third, the US is more culturally similar to Europeans than for instance workers located in India, the second largest national group.

All participants were asked to state their gender, age, highest attained level of education, and political identification. Figure 1 shows an overview of the distributions of these attributes.[4] More detailed information is included in appendix B.

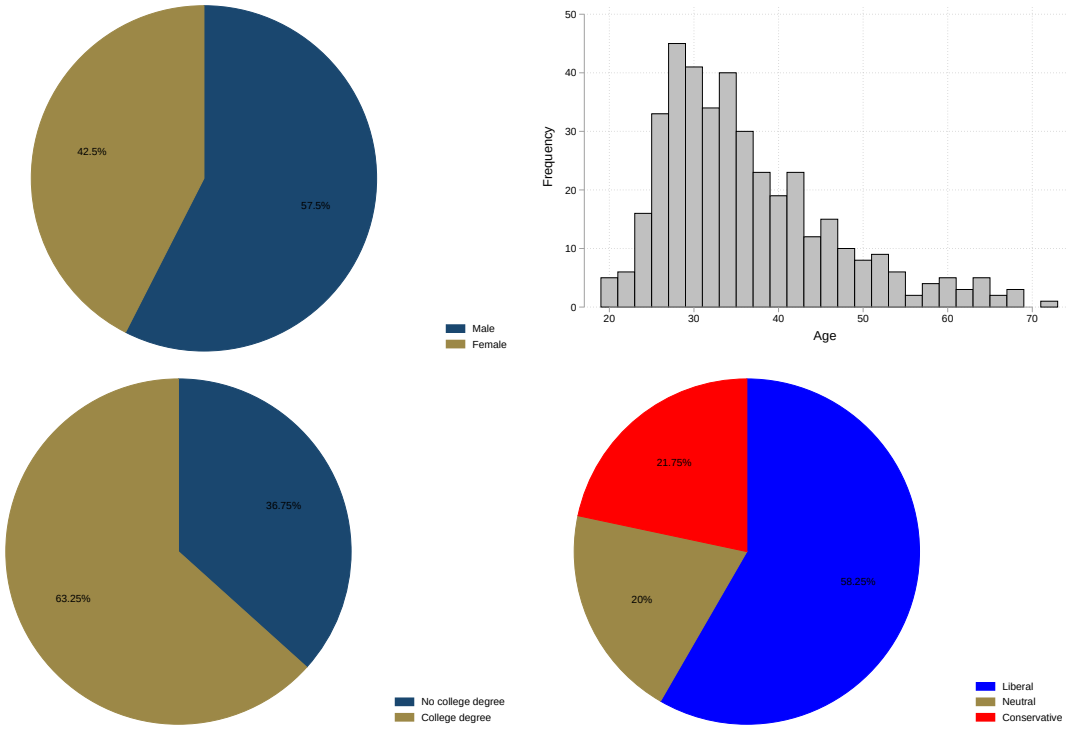
[1]This does create a small incentive for spectators to try to increase total earnings for the players, but since the amount was not disclosed before after they had made their choice, this should be unproblematic.

[2]250 additional workers started the experiment but did not finish it. This could in theory create a selection bias, however, none of these made it through the first set of control questions, and thus dropped out before the randomization happened.

[3]The experiment was restricted to workers located in the US, but this restriction was based on the country reported in worker's profiles, not on IP addresses. This means that it was possible for workers to lie about their location. 15 subjects in fact had IP addresses from outside the US. There is no way to know if any of the other subjects used proxies to change their IP so it appeared as if they were in the US, hence there is no way to know the true proportion of Americans in the sample. This could reduce the applicability of the results to the American population, but it does not reduce the general credibility of the results.

[4]Graphical scheme created by Daniel Bischof. Documentation (unpublished):

Figure 1: Demographic characteristics of the sample



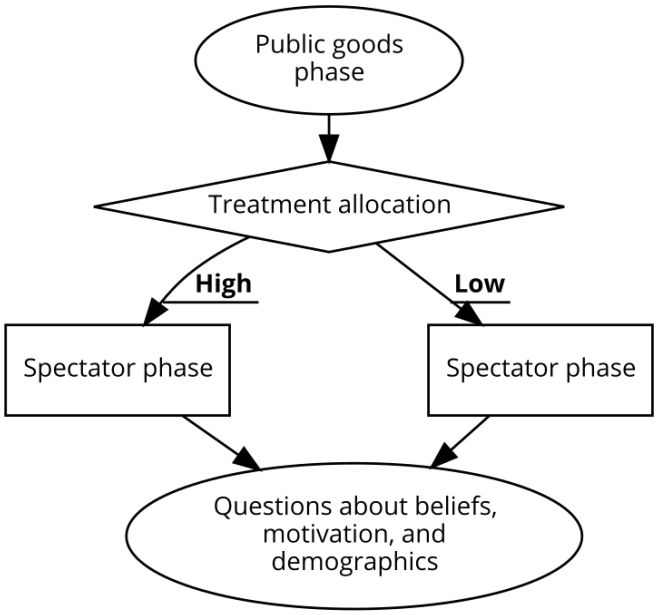
Note: The graphs show demographic characteristics of the sample. The first graph shows the gender distribution. The second is a histogram of age in years, with each bar representing two years. The third shows education level, with "College degree" including everyone with at least a two-year college degree. The final graph shows the distribution of political identification based on ratings on a five-point scale, with 1 and 2 counted as "liberal", 3 as "neutral" and 4 and 5 as "conservative".

There were slightly more males than females. The sample was a bit younger than the general population, with a median age of 34 years, and more than half of the participants between 25 and 35 years old. The youngest participant was 19 pears old and the oldest was 72. Education rates are comparable to that of the general population, except for a significantly smaller proportion having only general education. To make the groups more even, and for simplicity of interpretation, I have in the following analysis combined these into two groups: one for those with a college or professional degree, which contains 63 percent of the participants, and one for those with no degree.

Participants were also asked to state describe themselves politically, using a 5 point scale, where 1 was described as "very liberal", 3 was "neutral" and 5 was

https://www.dropbox.com/s/9yy6uit93flfpeh/FigureSchemeSTJO.pdf?dl=0

Figure 2: Flowchart of experiment procedure



Note: The figure is a flowchart of the entire experimental process. Each subject made decisions in both phases of the experiment; one of these was then randomly implemented.

"very conservative". The majority rated themselves as either 1 or 2. To make the groups larger and more equal, and for ease of interpretation, I have grouped those selecting 1 and 2 as "liberal", kept 3 as "neutral", and combined 4 and 5 as "conservative".

The treatments were randomized, so there is no reason to suspect selection bias or serious imbalances. There is a theoretical possibility that subjects lied about their education level, but this is mostly irrelevant. More serious is the possibility that treatment influenced political identification through the choices made; the choices were political in nature and could plausibly prime some subjects to describe themselves differently. I found no evidence for a difference in political identifications between the treatments, neither when using the aggregated categories (Wilcoxon rank-sum test, $p = 0.180$) or the full set of categories ($p = 0.093$).

Procedure The experiment was conducted on May 11 2017, and completed in a few hours. The participants were redirected from MTurk to the experiment form, which was created using the survey provider Qualtrics. The survey is included in its entirety, including both treatment variations, in appendix C.

The survey automatically retrieved each subject's unique MTurk identification code, which made it possible to pay the correct bonus to each subject and to ensure

that no one participated several times.[5] At the end of the survey, participants retrieved a random code which they needed to input to MTurk to complete the assignment. This allowed me to test that the codes from Qualtrics and MTurk matched, and thus if a participant had completed the experiment.

To decrease the risk of subjects misunderstanding or not paying attention, the instructions for each part of the experiment were followed by a set of control questions that participants had to answer correctly before being able to proceed to the actual decision-making. After completing the two parts of the experiment, subjects were presented with a set of statements about the results, and asked to rate these as true or false. They were also requested to make write a short description of why they made the choice they made in the spectator phase.

After the experiment was completed, subjects were randomly assigned to groups of five. The grouping was independent of treatment. The first participant in each group was chosen to be the spectator, while the remaining four were the players. If the spectator had chosen to intervene, the experiment ended and the players received the fixed payment. If the spectator chose not intervene, the results of the public goods game between the players was implemented. The spectator received a bonus equal to the average earnings of all players.

There were 80 groups in total. The average earnings from all the public goods games was 136.14 bonus points. The spectator chose the fixed payment in 54 of the games, which increased average earnings to 163.59 bonus points. In total 654.82 USD was paid out as bonuses, as well as 400 USD in show-up fees. Including overhead fees charged by Amazon and converting to the local currency, the experiment ended up costing 11,888.97 NOK.

Web-based experiment The participants were recruited via MTurk, which is an online labor marketplace that primarily is used for outsourcing tasks that are simple and time-consuming, but too complex for a computer to solve. Examples are audio and data transcription, image recognition and product categorization. These tasks typically pay a few cents, with average hourly wages ranging between 1 and 5 USD. (Ipeirotis 2010; Mayyasi 2013) For comparison the US federal minimum wage rate is 7.25 USD per hour.

The main benefits of using MTurk were price and simplicity. A large number of available workers means that an experiment with a large number of participants could be conducted quickly and easily. The low wage level means that it was cheap to provide compensation that was large enough to be economically significant.

[5]There is a theoretical possibility that some workers used several accounts to participate in the experiment more than once. This would be difficult, as each account requires a unique bank account, but not impossible. There were 10 instances of respondents using the same IP address, with one instance of the same address being used four times. This could indicate cheaters, but the easier explanation is that some workers shared the same network.

This creates a stronger incentive for participants to care about and think carefully through their decision.

On average it took participants 11.18 minutes to complete this experiment. The average payment was 2.64 USD, which includes a 1 USD show-up fee. This translates to an average hourly wage of 14.17 USD. This is much higher than both average wages on MTurk and the US federal minimum wage, meaning that participants had a strong incentive to make the correct decision.

Using MTurk for conducting experiments has rapidly become an established practice within economics and social sciences. This practice is examined in Paolacci, Chandler, and Ipeirotis (2010), Berinsky, Huber, and Lenz (2017), Buhrmester, Kwang, and Gosling (2011) and Horton, Rand, and Zeckhauser (2011), which all reach similar conclusions: Experiments using MTurk are just as valid and reliable as ones using traditional methods. Workers differ from the general population in some aspects, but are on the whole not very unrepresentative, and most likely more representative than the most commonly used subjects for economic experiments, ie., business students.

However, conducting an experiment online means giving up control over some factors that could be more closely monitored in a physical environment. This presents some challenges to external validity, which are discussed in the subsection. Most of these are inconsequential, and none pose a serious threat to the validity of the study.

Since the experiment was conducted using a web survey, there was no administrator present who could clarify points participants were uncertain of. This means that there is a higher risk of misunderstandings than with equivalent instructions in a physical environment. Some misunderstandings would only increase noise, but others could create biases. The control questions should to a large degree mitigate this problem, but do not rule it out.

There is no way to know if any participants communicated with each other. The large sample and random grouping made it impossible to coordinate strategies, but there is a theoretical possibility that some may have shared information that influenced the choices of others. It is also a possible that some participants shared answer keys to the control questions, guiding confused participants through the safeguards. During and after the experiment, I monitored the most popular MTurk forums. The experiment was only posted once, in a forum called MTurk Forum, and the post did not include any information about the experiment. This does of course not preclude private communication.

Chapter 4

The public goods game

This chapter gives an overview of the results from the public goods phase, and its implications for spectator choices. The results were generally in line with previous research, surveyed in section 2.2. The average contribution was 35.425 and the median was 25 points. A histogram of all contributions is shown in fig. 3. A short regression analysis is relegated to appendix A.

The contributions spiked at 0, 50 and 100 points; 91.5 percent of participants chose one of these three amounts. This suggests a strong anchoring effect from the control questions, which all involved computing payouts for different combinations of these contributions. Some participants may even have believed that these were the only possible contribution levels.

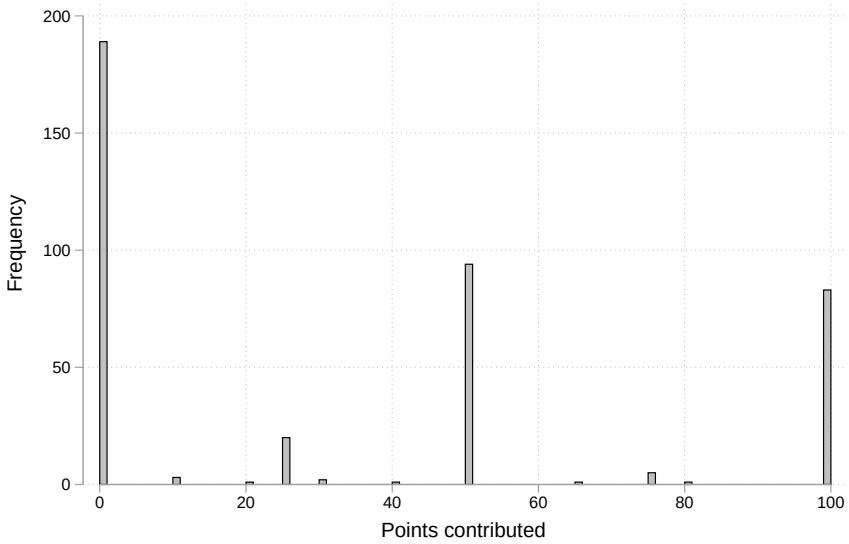
However, we would expect clustering on these numbers even without any anchoring, simply because people prefer round numbers. Even if there is an anchoring effect, there is no reason to believe it has caused any bias, as the control questions were perfectly balanced.

Zero contributions were by far the most common choice, and was made by 47.3 percent of the sample. 20.8 percent of participants chose to contribute their entire contribution, meaning that almost 70 percent chose one of the two extreme options.

In some of the later analyses, I have divided the sample into two categories based on their contribution level. I have labeled the players who contributed at least 50 points "cooperators", and the rest "defectors". This definition relates to the low treatment. 150 points is what all players earn if the each contribute 50 points, so any group consisting entirely of cooperators would do better than the fixed payment, and any group consisting entirely of defectors would do worse. The share of cooperators was 46 percent.

There are two main insights to be gained from these results. First, there is considerable cooperation failure, and hence a large expected gain from regulation. Second, because the average payment is close to the fixed payment in the low

Figure 3: Histogram of contributions



Note: The graph shows a histogram of all contributions in the public goods game. Each bar represents a single number of points.

treatment, the efficiency gain is much bigger in the high treatment. In a group consisting of average players, the expected earnings were 1.35 USD per player. Regulation would lead to an increase of 0.15 USD (11 percent) in the low treatment, and an increase of 0.65 USD (48 percent) in the high treatment.

The expected earnings for any player sans intervention is $153.18 - 0.5X$, where X is the number of points contributed. There are two points to be taken from this. The first is that a player has a net expected payment increase relative to the initial endowment for *any* contribution. While not intervening means players are expected to lose out relative to the fixed payment, they are not expected to be worse off than what they started with.

The second point is while that a completely selfish individual is expected to earn more than the fixed payment in the low treatment, they are not expected to earn more than the fixed payment in the high treatment. Thus even a spectator who somewhat bizarrely only cares about the payments to the highest earner should prefer the intervention in the high treatment.

Chapter 5

The spectator phase

A note on estimation technique I have relied on OLS to estimate linear probability models (LPM) instead of using a non-linear model such as logit or probit. The main advantage is that this produces coefficients that can be interpreted as probabilities. For further support for this decision, I rely on arguments from Pischke (2012) and Angrist and Pischke (2009).

Logit and probit are more efficient if the underlying conditional expectancy function (CEF) has a sigmoid shape, however I have no particular reason to believe this. Not knowing the shape of the CEF, a linear model is as good an approximation as the non-linear alternatives; furthermore, a linear model is a good approximation of a sigmoid shape for anything but the extreme parts of the probability range. Paraphrasing Pischke (2012), I am interested in marginal effects, not the shape of the CEF function, and LPM does a pretty good job of estimating those.

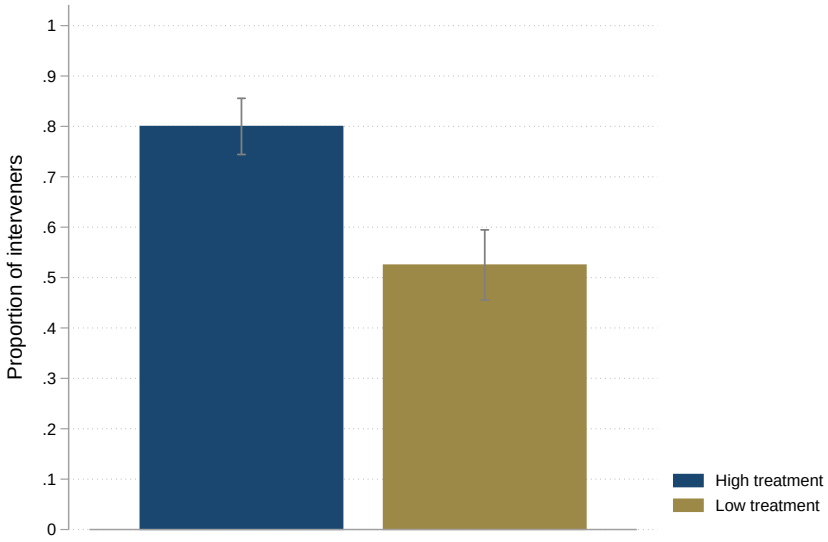
There are two further arguments that are used against LPMs. First, they produce heteroskedastic residuals. Robust standard errors however takes care of this. Second, they may predict nonsensical probabilities that are not between 0 and 1. In this authors view, this simply means that the estimates are wrong, not that there is necessarily anything wrong with the estimation technique.

5.1 Main results

The proportion of total spectators that chose the fixed payment was 0.663. The proportion was 0.8 in the high treatment and 0.525 in the low treatment. The difference between the treatments was strongly significant ($p = 0.000$). The regression is reported in the first row of table 1. The results are illustrated in fig. 4.

These three findings deserve separate treatments. I look first at the high treatment. This shows that a large majority chooses to intervene in the situation where

Figure 4: Share of interveners by treatment



Note: The graph shows the proportion of spectators that chose the fixed payment option in each treatment, including 95 percent confidence intervals. The fixed payment option was 2 USD in the high treatment and 1.50 USD in the low treatment.

there is a large expected efficiency gain and no potential material downside.

What is more exciting is the 20 percent that chose *not* to intervene. This group rejected an option that was virtually certain to make the others better off. This proportion is too large to be discarded as noise or mistakes. The hypothesis that this proportion is zero, ie., that the intercept in model 1 in table 1 is equal to 1, is strongly rejected ($p = 0.000$).

There are really only three possible explanations. The first is that this is evidence for anti-social preferences. But that 20 percent of the population actively wish to make others worse off squares poorly with both intuition and previous behavioral research. The second is confusion, but the segment is large enough that this explanation would require a causal mechanism that leads to systematic errors.

The third and most convincing explanation is that this segment did not intervene because they have a strong preference against overriding the decisions of others. This provides strong evidence that a fairly large share of the population exhibit strong anti-interventionary, or anti-paternalistic norms. This norm so strong that they are unwilling to act paternalistically even in the face of a large and certain welfare increase, possibly any welfare increase. Based on the 95 percent confidence intervals, this should apply to between 15 and 25 percent of the

Table 1: Direct effects of treatment and demographics

	(1)	(2)	(3)	(4)	(5)	(6)
Low treatment	-0.275***					-0.268***
	(0.0454)					(0.0456)
Female		0.0243				0.0251
		(0.0478)				(0.0457)
Age			-0.00414*			-0.00304
			(0.00233)			(0.00233)
Neutral				0.0848		0.0505
				(0.0577)		(0.0544)
Conservative				-0.0905		-0.0902
				(0.0616)		(0.0619)
College degree					0.0257	0.00877
					(0.0494)	(0.0479)
Constant	0.800***	0.652***	0.811***	0.665***	0.646***	0.899***
	(0.0284)	(0.0315)	(0.0856)	(0.0310)	(0.0395)	(0.0950)
Adjusted R^2	0.082	-0.002	0.006	0.009	-0.002	0.087
Observations	400	400	400	400	400	400

Heteroscedasticity-consistent standard errors in parantheses.

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

Note: The table reports OLS-estimated LPMs of the share of spectators that chose the fixed payment. "Low treatment" is an indicator for being in the treatment group where the fixed payment option was 1.50 USD; in the high treatment it was 2 USD. "Female" is an indicator for being female. "Age" is age in years. "Neutral" and "Conservative" are indictors for political identification based on a 5-point scale, with 1 being "very liberal" and 5 being "very conservative". "Neutral" indicates that the participant answered 3, "Conservative" that she answered 4 or 5. "College degree" is an indicator for having completed at least a 2-year college or professional degree.

population.

In response to the worsening trade-off between freedom and welfare between the treatments, a further 32 percentage points of spectators chose not to intervene in the low treatment. But a majority did. The expected gains from intervention was only 0.15 USD per player, compared to 0.65 USD per player in the high treatment, and the was a relatively high probability that players would do better on their own.

There are several possible explanations for why so many chose the intervention. They could have a strong preference for efficiency relative to autonomy, so that

only a small efficiency increase was enough to make intervention preferable. They could have believed that the efficiency gain was larger than it actually was. They could have a strong preference for equality. Or they could have a strong preference against uncertainty. Obviously, several or all of these could apply simultaneously, or they could apply to different segments.

Last, we look at the difference itself. This difference is highly significant, and because the experiment was randomized, it can only be explained by the treatment variation. It must therefore be explained by the decreased efficiency of the fixed payment option, since this was the only thing that changed between treatments.

Based on the results we can postulate that the population can be divided into three. First we have the never-interveners. This segment only cares about autonomy (within the bounds of the attainable outcomes), and does not intervene for any of the presented trade-offs. From the high treatment, we estimated that this accounts for around 20 percent of the population.

Then we have the never-interveners. This segment only cares about other factors, such as equality and certainty (again within the bounds of the attainable outcomes), and so will intervene for any of the presented trade-offs. The size of this segment cannot be estimated from these results, but it cannot be larger the population share that chooses intervention with the low treatment. Nor can it be excluded that the size of this segment is zero.

Finally we have the conditional interveners. Only members of this segment will respond differently to the different treatments. The entire difference between the treatments must therefore be due to changes in this segment.

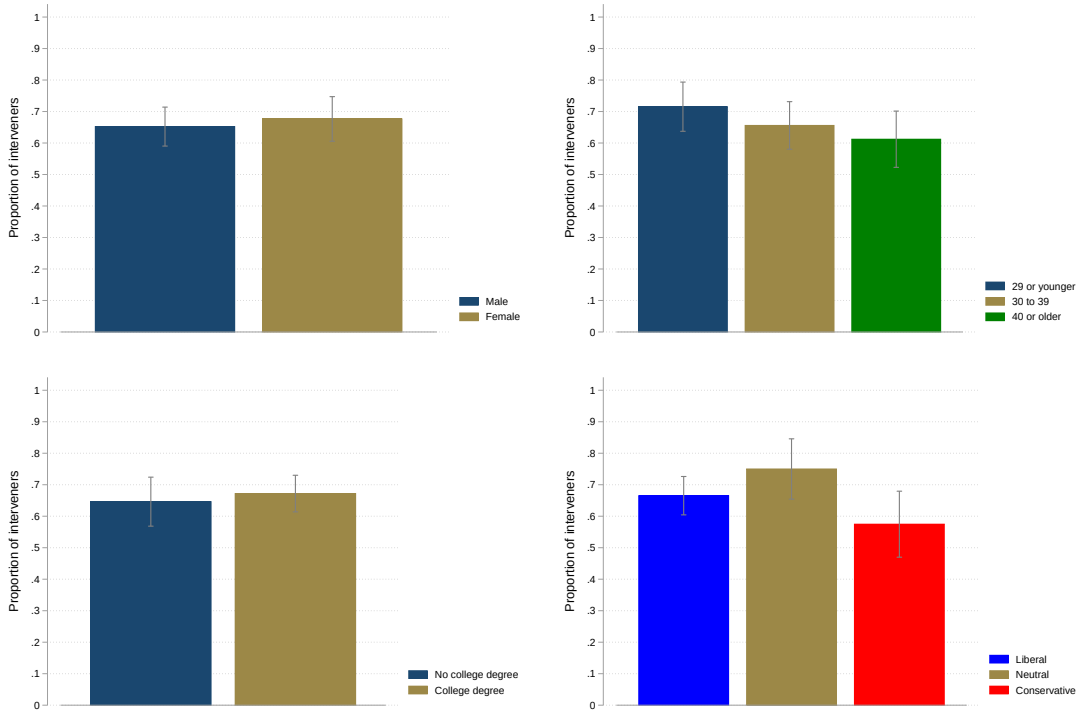
However, the treatments give only a cursory understanding of what motivates this segment. All we know is that this segment has a preference for efficiency. We do not know if they care only about efficiency, or if they also take other factors into consideration. Nor do we know the size of this segment, or of the proportion of the segment that chooses to intervene in the low treatment.

What we do know is that the 0.275 share of the sample that chooses to intervene in the high treatment but not in the low, must belong to this segment. There are two possible explanations for this change in behavior. First, it could be a direct response to the reduced efficiency gain relative to the restriction of autonomy. Second, it could indicate that they believe the players would earn more if left alone, ie., that there would be no efficiency gain.

5.2 Demographic and political differences

This section examines whether there were any systematic differences in spectator choices between the demographic and political groups in the sample. An overview of the intervention shares in the different groups is shown in fig. 5. Regression

Figure 5: Share of interveners by demographic groups



Note: The graphs shows the proportion of spectators that chose the fixed payment in different demographic groups, including 95 percent confidence intervals. "College degree" refers to anyone with at least a 2-year college or professional degree. The political identifications are based on ratings on a five-point scale, with 1 and 2 counted as "liberal", 3 as "neutral" and 4 and 5 as "conservative".

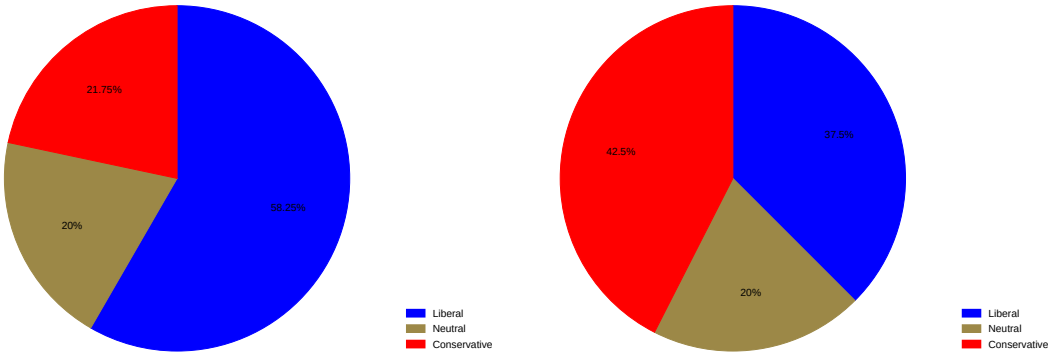
results are reported in table 1 and table 2. A detailed table of the characteristics of interveners and non-interveners is included in appendix B.

While women were more likely to intervene both on average and in both treatments, there was no significant effect of gender in any of the specifications. I also found no effect of age, using a simple linear specification. To account for possible non-linearities, I re-estimated the model using a quadratic specification, but still found no support for a relationship.

Those with a college degree were on average more likely to intervene. They were more likely to intervene in the high treatment, and less likely to intervene in the low treatment. None of the effects were significant.

On average, conservatives were less likely to intervene than liberals, and neutrals were more likely to intervene than both. Neither conservatives or neutrals were significantly different from liberals. They were however significantly different from each other: Conservatives were about 18 percentage points less likely to

Figure 6: Political identification of never-interveners



Note: The graphs show the distribution of political identification based on ratings on a five-point scale, with 1 and 2 counted as "liberal", 3 as "neutral" and 4 and 5 as "conservative". On the left is the distribution for the entire sample. On the right is the distribution among non-interveners in the high treatment, ie., with a fixed payment option of 2 USD.

intervene ($p = 0.015$).

The difference becomes more apparent when looking at treatment-specific effects. Conservatives in the high treatment were about 25 percentage points less likely to intervene than liberals ($p = 0.002$) and about 22 percentage points less likely than neutrals ($p = 0.017$). There was however no significant difference between conservatives and liberals or between conservatives and neutrals in the low treatment. Nor was there any difference between liberals and neutrals in either treatment.

To summarize, when presented with a relatively bad trade-off between autonomy and efficiency, conservatives, liberals and neutrals behave similarly. When presented with a much better trade-off, conservatives differ wildly from the rest. Why?

We can look at this in terms of the framework from the previous section. There, I labeled the segment that did not intervene in the high treatment never-interveners. Conservatives are overrepresented among never-interveners: the proportion of conservatives in this group is 0.425, compared with 0.2175 in the entire sample. The inverse relationship is even more telling: from the high treatment, I estimate that 38.6 percent of conservatives in the population are never-interveners. The political composition of never-interveners and the entire sample is illustrated in fig. 6.

The estimated proportion of conservative never-interveners in the total population is however only 8.5 percent. My hypothesis is that this is large enough to make a splash in the high treatment, where opposition to paternalism is the only reason not to intervene, but not in the low treatment, where there were many

Table 2: Interacted effects of treatment and demographics

	(1)	(2)	(3)	(4)	(5)
Low treatment	-0.284***	-0.169	-0.365***	-0.202***	-0.194
	(0.0604)	(0.170)	(0.0562)	(0.0767)	(0.186)
Female	0.0164				0.00227
	(0.0573)				(0.0563)
Low treatment × Female	0.0213				0.0305
	(0.0917)				(0.0916)
Age		-0.00211			-0.00201
		(0.00312)			(0.00301)
Low treatment × Age		-0.00288			-0.00355
		(0.00460)			(0.00462)
Neutral			-0.0278		-0.0242
			(0.0637)		(0.0630)
Conservative			-0.247***		-0.242***
			(0.0812)		(0.0824)
Low treatment × Neutral			0.157		0.151
			(0.116)		(0.114)
Low treatment × Conservative			0.286**		0.320***
			(0.120)		(0.123)
College degree				0.0633	0.0494
				(0.0626)	(0.0619)
Low treatment × College degree				-0.115	-0.0946
				(0.0955)	(0.0963)
Constant	0.793***	0.875***	0.861***	0.758***	0.896***
	(0.0378)	(0.112)	(0.0335)	(0.0530)	(0.118)
Adjusted R^2	0.079	0.085	0.100	0.081	0.097
Observations	400	400	400	400	400

Heteroscedasticity-consistent standard errors in parantheses.

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

Note: The table reports OLS-estimated LPMs of the share of spectators that chose the fixed payment. "Low treatment" is an indicator for being in the treatment group where the fixed payment option was 1.50 USD; in the high treatment it was 2 USD. "Female" is an indicator for being female. "Age" is age in years. "Neutral" and "Conservative" are indictors for political identification based on a 5-point scale, with 1 being "very liberal" and 5 being "very conservative". "Neutral" indicates that the participant answered 3, "Conservative" that she answered 4 or 5. "College degree" is an indicator for having completed at least a 2-year college or professional degree. × indicates an interaction term.

possible reasons. If this interpretation is correct, it means that this effect does not disappear in the low treatment, and so it should reappear with a larger sample or more targeted treatments.

5.3 Beliefs and previous choices

Contribution levels The experiment was conducted such that each subject made decisions in both parts of the experiment. Each player was then randomly assigned a role, and the decisions made in that part were implemented for her. This subsection compares the choices made by the same subject in the spectator and the public goods game. The regression results are reported in table 3.

I used a linear specification of the amount of points a player contributed in the public goods phase. I estimated the models both with and without using demographic characteristics as controls, and found almost identical coefficients. I also re-estimated the model using two categories (with a cut-off of 50 points), with similar results.

The probability of intervention decreased on average with 0.0019 per point contributed, which implies a difference of 19 percentage points between full cooperators and defectors. This effect was strongly significant ($p = 0.002$).

When looking at treatment-specific effects, we find a large difference between the groups. The contribution level had no significant effect on the high treatment, and a stronger effect in the low treatment. In the latter, the probability of intervention is estimated to decrease with 0.0034 per point contributed, translating to a difference of 34 percentage points between high and low contributors. This effect was strongly significant ($p = 0.000$).

What can account for this effect? My hypothesis is that this results from people believing others are more similar to themselves than they are. High contributors would therefore tend to believe others also contributed. This should have no effect in the high treatment, where intervention maximized efficiency, but a large effect in the low treatment. Further evidence for this hypothesis is presented in the next subsection.

Beliefs After completing the two main parts of the experiment, the subjects were presented with four statements about different aspects of the experiment. The subjects were required to rate these statements as true, false or uncertain. The questions were intended to measure their beliefs about what would be the result, which could then be compared with their actions in order to provide further insight into their motivation. The results of the regression analyses are reported in table 4 and table 5. Detailed information about the distribution of answers is included in appendix B.

Table 3: Effect of public goods game choices

	(1)	(2)	(3)	(4)
Low treatment			-0.147**	-0.143**
			(0.0605)	(0.0602)
Points contributed	-0.00190***	-0.00177***	0.0000134	0.0000638
	(0.000614)	(0.000629)	(0.000749)	(0.000732)
Female		0.0364		0.0413
		(0.0471)		(0.0448)
Age		-0.00339		0.00133
		(0.00234)		(0.0135)
College degree		0.0178		-0.00439
		(0.0491)		(0.0471)
Neutral		0.0700		0.0399
		(0.0561)		(0.0511)
Conservative		-0.0735		-0.0884
		(0.0612)		(0.0628)
Low treatment × Points contributed			-0.00346***	-0.00347***
			(0.00111)	(0.00111)
Age^2				-0.0000453
				(0.000162)
Constant	0.730***	0.822***	0.800***	0.808***
	(0.0307)	(0.0975)	(0.0381)	(0.267)
Adjusted R^2	0.023	0.028	0.120	0.121
Observations	400	400	400	400

Heteroscedasticity-consistent standard errors in parantheses.

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

Note: The table reports OLS-estimated LPMs of the share of spectators that chose the fixed payment. "Points contributed" is how many points the participant contributed in the public goods game. "Low treatment" is an indicator for being in the treatment group where the fixed payment option was 1.50 USD; in the high treatment it was 2 USD. "Female" is an indicator for being female. "Age" is age in years. "Neutral" and "Conservative" are indictors for political identification based on a 5-point scale, with 1 being "very liberal" and 5 being "very conservative". "Neutral" indicates that the participant answered 3, "Conservative" that she answered 4 or 5. "College degree" is an indicator for having completed at least a 2-year college or professional degree. × indicates an interaction term.

The first statement was: "Participants on average contributed more than 50 points to the common pool." The results are in the first two models of table 4.

This question was intended to measure beliefs about how efficient the outcome of the public goods game was likely to be. It is directly related to the low treatment: if average contributions were 50 points or more, the average group would earn more

Table 4: Beliefs about efficiency

	(1)	(2)	(3)
Low treatment		-0.155***	-0.0766
		(0.0598)	(0.0687)
Q1: True	-0.220***	-0.0917	-0.142
	(0.0587)	(0.0759)	(0.0999)
Q1: True × Low treatment		-0.239**	-0.0596
		(0.110)	(0.142)
Points contributed			0.00113
			(0.00116)
Low treatment × Points contributed			-0.00362**
			(0.00166)
Constant	0.725***	0.800***	0.776***
	(0.0301)	(0.0375)	(0.0428)
Adjusted R^2	0.043	0.109	0.121
Observations	321	321	321

Heteroscedasticity-consistent standard errors in parantheses.

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

Note: The table reports OLS-estimated LPMs of the share of spectators that chose the fixed payment. It reports the effects of the answers to the question "Participants on average contributed more than 50 points to the common pool.". "Q1: True" is an indicator for whether the participants answered "true". Participants that answered "I don't know" are excluded. "Low treatment" is an indicator for being in the treatment group where the fixed payment option was 1.50 USD; in the high treatment it was 2 USD. "Points contributed" is how many points the participant contributed in the public goods game. × indicates an interaction term.

from the public goods game than the fixed payment. If a large share of spectators had a strong preference for efficiency, we would expect the answer to this question to have a large effect in the low treatment, and none at all in the high treatment.

This is exactly what I found. The answer had no significant effect in the high treatment. In the low treatment, answering affirmatively to this question reduced the probability of intervention by 0.3311, which was strongly significant ($p = 0.000$).

It is worth taking a look at the predicted probabilities as well. The probability that someone from the low treatment who gave a positive answer would intervene was 0.313. The probability that someone who answered negatively would intervene was 0.645. These indicate that while there is strong evidence for a preference for efficiency, it cannot be the only motive.

This finding also provides strong support for the hypothesis stated above. The regression on the beliefs about efficiency gives exactly the same picture as the regression on contribution levels. Furthermore, when we regress this belief on the

Table 5: Beliefs about equality and the spectator phase

	(1)	(2)	(3)	(4)	(5)	(6)
Low treatment		-0.207**		-0.198**		-0.353***
		(0.0810)		(0.0987)		(0.127)
Q2: True	-0.0294	0.00465				
	(0.0542)	(0.0673)				
Q2: True × Low treatment		-0.0999				
		(0.105)				
Q3: True			0.661***	0.544***		
			(0.0465)	(0.0924)		
Q3: True × Low treatment				0.142		
				(0.109)		
Q4: True					0.527***	0.315***
					(0.0582)	(0.119)
Q4: True × Low treatment						0.239*
						(0.137)
Constant	0.667***	0.774***	0.206***	0.345***	0.268***	0.526***
	(0.0416)	(0.0534)	(0.0412)	(0.0888)	(0.0527)	(0.115)
Adjusted R^2	-0.002	0.073	0.405	0.415	0.211	0.242
Observations	322	322	346	346	339	339

Heteroscedasticity-consistent standard errors in parantheses.

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

Note: The table reports OLS-estimated LPMs of the share of spectators that chose the fixed payment. The first two columns report the effects of the answers to the question "In most groups, participants contributed a roughly equal number of points to the common pool.". Models 3 and 4 report the effects of the answers to the question "The majority of participants chose Option 1, ie., that the other participants receive a fixed payment." The last two columns report the effects of the answers to the question "Most participants would be better off with Option 1, ie., a fixed payment, than Option 2, ie., payments based on the choices they made in Part 1.". "Q2/Q3/Q4: True" are indicators for whether the participants answered "true" to the respective questions. Participants that answered "I don't know" are excluded from the models involving that question. "Low treatment" is an indicator for being in the treatment group where the fixed payment option was 1.50 USD; in the high treatment it was 2 USD. × indicates an interaction term.

contribution level directly, we find that each point contributed increases the probability of answering affirmatively by 0.0044 ($p = 0.000$). The predicted probability that someone who contributed 100 points answered positively is 0.927.

This implies that most of the effect from contributions on choices can be explained as an indirect effect through changed beliefs. However, when controlling for contributions, shown in model 3 in table 4, I find evidence of independent effects of both on spectator choices in the low treatment ($p_{Contribution} = 0.038$,

$p_{Q1} = 0.048$).

The second statement was: "In most groups, participants contributed a roughly equal number of points to the common pool." The results from the regressions are reported in the last two columns of table 5 This question was intended as a measure about how much unequal participants expected the outcomes from the public goods games to be.

I found no evidence that the answer to this question had any impact on choices, neither on average or in the individual treatments. That there were no treatment-specific effects was as expected, as there was no difference between the treatments with regards to equality. That there was no average effect is more surprising. As will be shown in chapter 6, many interveners cited equality as the primary motive for intervening, so this seems like a poor fit.

Assuming that the question accurately captures beliefs about equality, there are two possible explanations for this disparity. The first is that there are some subjects who care *only* about equality. Since the fixed payment maximizes equality in both treatments, neither beliefs nor treatment should make any difference. (They would in other words be always-interveners.) The second is that interveners and non-interveners do not differ in their beliefs about the distribution of the outcome, only in the relative weight they ascribe equality.

Alternatively, the participants could overstate the actual importance of equality, or there could be an effect that the regression does not capture, due to the sample size or the way the question was phrased.

The third statement was: "The majority of participants chose Option 1, ie., that the other participants receive a fixed payment." This was intended mostly as a test of whether people believed others made the same choice as them. I found strong evidence that they did. Interveners were on average 66.1 percentage points more likely to believe others also intervened ($p = 0.000$). 86.7 percent of interveners thought others also intervened, while 79.4 percent of non-interveners believed the opposite. I found no evidence for any treatment-specific effects.

The fourth statement was: "Most participants would be better off with Option 1, ie., a fixed payment, than Option 2, ie., payments based on the choices they made in Part 1." This was intended to test whether subjects intended to increase welfare. If they were, we would expect them to answer consistently with their choice.

I did find a high degree of consistency: subjects who believed intervention would be beneficial were also 52.7 percentage points more likely to intervene ($p = 0.000$). There was only weak support that this effect was different between treatments ($p = 0.082$).

While the coefficient is high, it is far from 1. 78.2 percent of the participants who answered this question, gave an answer consistent with their choice, mean-

ing that more than a fifth of the spectators made the choice they believed to be least beneficial. This effect is highly asymmetrical. Only 8.2 percent of interveners believed their choice to be unbeneficial, compared with 51.4 percent of non-interveners. The last number was 41.1 percent in the low treatment and 73.5 percent in the high treatment.[1]

While the answers to this question reveal that a large majority try to maximize welfare, it also provides further support that a large minority are not willing to overrule the choices of others even when they believe it will lead to increased welfare.

[1]It could be argued that the fact that anyone at all in the high treatment believed intervention would not be beneficial indicates a high level of misunderstanding. However, it could also be explained by some subjects interpreting the question in a broader way, accounting for autonomy loss as well.

Chapter 6

Qualitative analysis of stated spectator motivations

Near the end of the experiment, the participants were requested to write a short explanation for why they made the choice they made in the spectator phase. Almost every participant did, with most writing thorough explanations In total I have 376 usable responses.[1][2] In this chapter I use qualitative analysis of these to complement the quantitative analysis in the last chapter.

In the first part I define categories based on the main hypotheses, ie., efficiency and autonomy, in addition to other possible motivations. I then examine the distribution of these motivations. In the second part I look closer at some of the actual responses in order to cast further light on the mechanisms that led subjects to choose intervention or not. In the third part I look at some misunderstandings and errors revealed by the responses and examine their implications for the other results.

6.1 Main results

The first category included the responses that explicitly cited increasing payments, or avoiding lower payments, as a reason for intervening or not intervening. The second category, labeled autonomy, only applied to non-interveners, and included the responses that in some way cited not to overrule decisions made by others as

[1]Three answered blank, and two complained about the experiment. 19 were excluded for other reasons, and will be discussed later.

[2]I have changed the choices made by 17 participants whose response explicitly stated that they believed they had made the opposite choice. This change does not carry over to or effect the quantitative analysis in any way. However, for the qualitative part, I believe there is valuable information to be gained from these responses, and including them in the "wrong" subgroup would lead to nonsensical results.

Table 6: Stated motivations

	High treatment	Low treatment	Total
Interveners			
Efficiency	0.5669	0.3100	0.4669
Equality	0.3057	0.1900	0.2607
Certainty	0.3248	0.3300	0.3268
Total	157	100	257
Non-interveners			
Efficiency	0.2813	0.4713	0.4202
Autonomy	0.6250	0.3678	0.4370
Total	32	87	119

Note: The table gives an overview of the motivations stated in the written responses of subjects. "Efficiency" means they stated a preference for increased payments. "Equality" means they stated a preference for equal payments. "Certainty" means they stated a preference for certain payments. "Automony" means they expressed a preference against intervening in the choices of others. "Total" lists the number of included responses in each subgroup. Categories overlap, and some responses did not fall into either category, so the proportions do not add up to 1.

a reason for not intervening. The last two only applied to interveners: the third cited equal or fair payments as a reason, and the fourth cited a preference for certainty of payments.

The share of responses that were included in each category are shown in table 6. Note that several of the categories overlap, and that some responses were not included in any of the categories, so the proportions do not add up to 1.

The findings provide further support for the interpretations given for the results of the quantitative analysis. Efficiency is the most commonly cited reason for intervention, cited by 46.7 percent of interveners. There is a clear treatment difference: the share that stated efficiency as a reason for intervening was almost twice as large in the high treatment, where the efficiency gain from intervention was much larger (two-tailed equal proportions test, $p = 0.000$). Efficiency was also the most important reason among non-interveners in the low treatment, being cited by 47.1 percent.[3]

A preference for autonomy was stated by 43.7 percent of interveners. This was also strongly treatment-specific: a much larger share of non-interveners in the high treatment cited autonomy ($p = 0.012$). Since there were more interveners in total

[3] 9 non-interveners in the high treatment gave similar reasons. This clearly indicates misunderstanding, and is discussed in section 6.3.

in the low treatment, the frequency that cited autonomy was a little higher here. This is consistent with the interpretation of the main results of the quantitative analysis: the share that intervenes because they care only about autonomy remains constant, but the total number that cites autonomy increases, because some of those that care about autonomy among other things now choose to intervene.

Finally, the responses provide support that both preferences for equality and preferences against uncertainty are important motivations for intervention, and that they are not strongly affected by the treatment variation.

Of the remaining responses, many would probably have been included in some of the aforementioned categories had they been more fleshed out, but were too vague to fit any of the specified criteria.[4]

There are some limitations to this approach. First, while I have tried to use objective criteria to categorize responses, some amount of subjective interpretation is unavoidable. Second, all the estimated proportions are likely strongly downward-biased, as the responses were brief, and it is likely that many subjects did not state all the reasons that influenced their choice, but no one was likely to state motivations that did not influence their choice. Third, that not all responses were included is an obvious source of potential selection bias.

In spite of these limitations, the findings are consistent with and provide further support for the main results of the experiment.

6.2 A closer look

The responses are the only aspect of the experiment that provides a window into the minds and thoughts of the participants. While responses were often brief and ambiguous, and may not accurately represent their actual decision-making process, I believe valuable information can be obtained by taking a closer look at some of the responses.

This subsection will not discuss all the types of responses, but instead concentrate on two qualitative observations. The first is about interveners. As stated, most participants stated at least one of increasing average/total payments, making the outcome equal, and decreasing uncertainty as a reason for intervention.

Some of these suggest a simple decision-making rule (such as "maximize total payments", "choose the most equal option", etc.). Others seem to imply decision-making rules that take all of these into account. These responses typically underlined an intention to reward cooperators and protect them from being "screwed over" by greedy participants, as put by several respondents.

[4]Miscellaneous reasons include curiosity about the results of the public goods game, a preference for uncertainty, one participant who did not want others to earn more than them, and one who plainly stated they did not believe the other participants were real.

There are two things to take notice of. First, these responses imply a wish to hold others accountable for their actions. Defectors are seen as perpetrators, not as victims of a rigged system that favors selfishness. Second, some of these responses seem to be consistent with a more Rawlsian decision-making rule, where the goal is to maximize the payments to the ones who are likely to make the least. (Rawls 1971)

The second observation concerns autonomy. As stated, a large share of non-interveners in total, and a majority in the high treatment, stated a preference against interfering in the choices of others. There were however large differences in the way this preference was stated.

In some responses, non-interference was stated as a positive value: these responses stressed the independent value of letting others make decisions for themselves. Others used more negative phrasings. Rather than extolling the virtues of self-determination, these respondents spoke of responsibility and consequences. The message was not as much of preserving liberty as of not saving others from the consequences of their actions. Some examples are included below.

> The choice was already made in part 1 so I think people should live with the decision they make the first time around and accept the consequences of that decision.

> Because it seems the most fair, if they wanted a different outcome they should have played differently.

> I decided to let the payment be based on actual choices because I wanted people to see the impact of their own decisions.

Some of these seem almost vindictive. Others express an explicit intention of punishing defectors and rewarding cooperators.

The latter is telling. When other participants are seen as individuals, with no possibility of communicating or coordinating strategies, this is illogical. On an individual level, the game punishes cooperation and rewards defection. These responses only make sense if the players are seen as a group that acts collectively.

What can be be learned from both of these observations is that views about fairness are a lot more complex than a simple reflection about the outcome distribution. A preference for rewarding cooperation and punishing defectors seems to play a large role, possibly even blinding some subjects to the consequences of their actions.

6.3 Misunderstandings

The responses revealed that some participants had made errors or misunderstood key aspects of the experiment. With a large sample size, some amount of confusion among subjects is almost unavoidable. Most of these are minor, and should be negligible in the larger picture. The concern is if a mistake is made by a large share of subjects, and is likely to have a directional impact on outcomes. The responses indicate two types of misunderstandings that are the cause of such concern.

The first type is that some participants evidently believed they were making decisions about their own payment in the spectator phase. 19 participants responded in a way that unequivocally indicated that they were trying to maximize their own payment. I have excluded these responses from the qualitative analysis, as they provide little information about third-party views about interference.

Around 30 more responded in a way that made it ambiguous if they had misunderstood: these subjects could be talking about their own literal outcome, but they could also be talking about what they would have preferred if they were selected for the public goods phase, in effect putting themselves in others' shoes.

It is uncertain how widespread this misunderstanding was. It is also uncertain in which direction this would bias decisions, if any, as both options could be consistent with self-interest, depending on individual contributions, beliefs about others and risk preference.

The second type is believing that it was possible for a group to make more than 2 USD per player in the public goods game. In total 13 subjects from the high treatment responded in a way that clearly shows this is what they believed. As with the previous type of misunderstanding, other responses make it ambiguous whether subjects had misunderstood or not, so the true extent is uncertain.

This type of misunderstanding has a clear directionality. For a high-treatment spectator who believed this and still preferred the fixed payment, this should make no difference. It is however likely that some of those who chose not to intervene, believing that this had the potential to increase group payments, would have made the opposite choice if this misunderstanding had been cleared up. Consequently, this is likely to have introduced a small downward bias to the intervention rate of the high treatment.

Chapter 7

Conclusion

The main hypothesis for this study was that people are motivated both by autonomy and efficiency considerations when making decisions about regulation. I found strong evidence for this. There was a 27.5 percentage points between the treatments, meaning that subjects were much more likely to intervene when the efficiency gain was large and certain, than when it was small and uncertain. The subsequent analyses highly consistent with this result and each other.

However, there was only weak support that *individuals* make trade-offs between autonomy and efficiency. Rather, the results were indicative of highly heterogeneous preferences about regulation: people seemed to care either very much or very little about autonomy, with only a few indicators of intermediacy.

20 percent of spectators in the high treatment chose not to intervene, despite having the opportunity to choose the optimal outcome of the public goods game. Further analyses support the interpretation that there is a share of the population, which I termed never-interveners, that are not willing to make any concessions to the autonomy of others in order to increase efficiency. I estimated this share to be between 15 and 25 percent. There was some indication that this is upward-biased, but still when taking this into account, this remains a significant share.

I also found strong evidence that total opposition to intervention is linked to political conservatism. The results indicate that a large share, estimated at almost 40 percent, of conservatives in the population are never-interveners.

For the rest of the population, the results suggested that expected efficiency is by far the most important factor when deciding whether to intervene or not. While respondents' written statements suggest that a preference for equality and uncertainty aversion played important roles, 78.2 percent made choices consistent with what they believed would lead to the highest payments for others. In the low treatment, where it was possible for a group of cooperators to do better than the fixed payment, beliefs about how much players were likely to contribute was the single strongest predictor of spectator choices.

These results also indicate that most of the treatment effect was due to different beliefs about the likely outcome of the public goods game, rather than different relative preferences for efficiency and autonomy. However, beliefs and choices were far from perfectly correlated, so different beliefs are unlikely to be the only explanation.

In general, the results provide strong support that this type of experiment is an effective way of eliciting revealed preferences about regulation. The large treatment effect, the high degree of consistency between the different parts of the experiment, and the written responses all strongly indicate that spectators care about and think carefully through the consequences of their decisions, even though these decisions have no impact on themselves.

Obviously, there are large differences between the simple public goods game and real-world externality issues. Most importantly, in real-world situations the sums involved are much higher, individual choices are a lot more important to them than a simple amount of bonus points, and it is possible, at least in theory, to communicate and coordinate strategies that are mutually beneficial. Nevertheless, that the results fit so well with intuition imply a high degree of generality. Especially, that political conservatives, who also in real-world situations are the most hostile to regulations, differed so much from the rest of the sample, supports the generality of the findings.

As the first of its kind, this experiment only scratches the surface of revealed preferences about regulation. It provides clear evidence for the role of efficiency and autonomy, but leaves some questions about the exact mechanisms open to interpretation. It also strongly indicates that equality and uncertainty aversion are important for regulatory preferences, but does not test these hypotheses directly.

In further research, the experiment could be extended or tweaked to include treatment variations for hypotheses about other motivating factors. To examine the importance of equality, the experiments could include treatments that have the same total payment, but different distributions. To examine the role of uncertainty, the experiment could be tweaked to give some spectators information about the outcome of the game. Analogously to how I estimated the proportion of never-interveners by including a treatment that was identical to the social optimum, the proportion of always-interveners, if they exist, could be estimated by a treatment with no potential efficiency gain, ie., with fixed payments equal to the initial endowments of the public goods game.

The results of this and possibly subsequent experiments have major policy implications. Every form of government intervention represents a trade-off between personal autonomy and other values. Understanding what people's values are, and which trade-offs they are willing to accept, is crucial to developing and gaining support for good policy.

These preliminary findings are especially important for those who argue for more regulation. They suggest that a minority are not willing to support regulation under any circumstances. They also suggest that the majority are willing to accept regulation as long as the welfare gains are made clear and visible. Politicians and activists, especially those working in fields where potential benefits of regulation are hard to see, such as climate change, should take notice.

Appendix A

Regression analysis of public goods phase results

This section presents a simple analysis of the results of the public goods game. It is not included in the main part of the article as it has no particular relevance to the regulatory choices made by the spectators. It is however included here for readers who may be interested in the results by themselves. The estimates are reported in table 7.

There was no significant difference between the genders. Age had a significant effect. I estimated a quadratic relationship, where contributions increase until subjects are 44 years old, and then decreases. Subjects with a college degree contributed significantly less. Those who described themselves as politically contributed significantly less than both liberals and conservatives, but liberals and conservatives were not significantly different from each other.

Table 7: Results of the public goods game

	(1)	(2)	(3)	(4)	(5)
Female	6.985*				6.352
	(3.906)				(3.898)
Age		0.154			0.127
		(0.190)			(0.194)
College degree			-7.986*		-8.332**
			(4.108)		(4.119)
Neutral				-10.56**	-10.88**
				(4.866)	(4.853)
Conservative				1.899	1.903
				(5.062)	(5.085)
Constant	32.46***	29.91***	40.48***	37.12***	35.19***
	(2.707)	(7.085)	(3.310)	(2.585)	(8.032)
Adjusted R^2	0.005	-0.001	0.007	0.008	0.019
Observations	400	400	400	400	400

Heteroscedasticity-consistent standard errors in parantheses.

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

Note: The table reports OLS estimates of the number of points contributed in the public goods game. "Low treatment" is an indicator for being in the treatment group where the fixed payment option was 1.50 USD; in the high treatment it was 2 USD. "Female" is an indicator for being female. "Age" is age in years. "Neutral" and "Conservative" are indictors for political identification based on a 5-point scale, with 1 being "very liberal" and 5 being "very conservative". "Neutral" indicates that the participant answered 3, "Conservative" that she answered 4 or 5. "College degree" is an indicator for having completed at least a 2-year college or professional degree.

Appendix B

Additional tables

Table 8: Characteristics of interveners and non-interveners

	Non-intervener	Intervener	Total
High treatment	0.2963	0.6038	0.5000
Low treatment	0.7037	0.3962	0.5000
Male	0.5926	0.5660	0.5750
Female	0.4074	0.4340	0.4250
29 or younger	0.2741	0.3509	0.3250
30 to 39	0.3926	0.3811	0.4125
40 or older	0.3333	0.2679	0.2900
Less than high school	0.0000	0.0038	0.0025
High school/GED	0.1111	0.1132	0.1125
Some college	0.2741	0.2415	0.2525
2-year college degree	0.0889	0.1057	0.1000
4-year college degree	0.4370	0.3849	0.4025
Master's degree	0.0593	0.1245	0.1025
Doctoral degree	0.0074	0.0075	0.0075
Professional degree (JD, MD)	0.0222	0.0189	0.0200
Very liberal	0.2148	0.2226	0.2200
Liberal	0.3630	0.3623	0.3625
Neutral	0.1481	0.2264	0.2000
Conservative	0.1852	0.1321	0.1500
Very conservative	0.0889	0.0566	0.0675
Total	135	265	400

Note: The table reports proportions of demographic groups by the choice made in the spectator phase. "Intervened" means that the subject chose the fixed payment option, while "non-intervener" means they chose the result of the public goods game. The fixed payment was 2 USD in the high treatment and 1.50 USD in the low treatment. The bottom row shows the actual number of participants in each subgroup.

Table 9: Beliefs by treatment and spectator choice

	Non-intervener	Intervener	High treatment	Low treatment	Total
Participants on average contributed more than 50 points to the common pool.					
True	0.4519	0.6075	0.5750	0.5350	0.5550
False	0.3630	0.1887	0.2550	0.4950	0.2475
I don't know	0.1852	0.2038	0.2100	0.3950	0.1975
In most groups, participants contributed a roughly equal number of points to the common pool.					
True	0.3185	0.3245	0.3100	0.3350	0.3225
False	0.5185	0.4642	0.5200	0.4450	0.4825
I don't know	0.1630	0.2113	0.1700	0.2200	0.1950
The majority of participants chose Option 1, ie., that the other participants receive a fixed payment.					
True	0.5704	0.0755	0.1450	0.3400	0.2425
False	0.2444	0.8151	0.7650	0.4800	0.6225
I don't know	0.1852	0.1094	0.0900	0.1800	0.1350
Most participants would be better off with Option 1, ie., a fixed payment, than Option 2, ie., payments based on the choices they made in Part 1.					
True	0.3852	0.0717	0.0950	0.2600	0.1775
False	0.4074	0.8038	0.7900	0.5500	0.6700
I don't know	0.2074	0.1245	0.1150	0.1900	0.1525
Total	135	265	200	200	400

Note: The table reports the proportions that gave each answer to four sets of questions the subject answered at the end of the survey. The exact questions used are shown above each set of statistics. "Intervened" means that the subject chose the fixed payment option. This was 2 USD in the high treatment and 1.50 USD in the low treatment. The bottom row shows the actual number of participants in each subgroup.

Appendix C

Survey text

Welcome!

Please note that your participation will be registered on the following Amazon Mechanical Turk worker ID:

The worker ID was retrieved automatically when you clicked on the link that brought you here. This step is necessary for assigning payments to the right account and to ensure that you only participate in this study once.

>>

Introduction

Welcome to this research project! We very much appreciate your participation.

This is a study about the economics of decision-making. Several research institutions have provided funds for this research.

Payment

Your payment will consist of the participation fee plus the amount of **bonus points** that you accumulate throughout the study. The exact amount of bonus points that you receive will depend on your and/or others' decision. At the end, each bonus points is converted into USD at a rate of **1 cent per bonus point**.

Your bonus will be paid to you using the bonus system within two to three weeks after the completion of this HIT. Your payment for taking the HIT will be sent to you shortly after the completion of this HIT.

Procedures

The study consists of two parts and you will be given instructions on your screen before every single part of the survey. At the end, we will randomly select **one of the two parts** and implement the outcome of this part for you. Please always make sure to read the instructions carefully before you continue.

Participation

Participation in this research study is completely voluntary. You have the right to withdraw at anytime or refuse to participate entirely without jeopardy to future participation in other studies conducted by us.

Confidentiality

All data obtained from you will be kept confidential and will only be reported in an aggregate format (by reporting only combined results and never reporting individual ones). All questionnaires will be concealed, and no one other than the primary investigator will have access to them. The data collected will be stored in the HIPPA-compliant, Qualtrics-secure database until it has been deleted by us.

Verification

At the end of this survey, you will be given a **completion code**. You will need to copy this code to the survey code field on the AMT web page that directed you here at the beginning.

Questions about the Research

If you have questions regarding this study, you may contact

thechoicelab@nhh.no

I have read and understood the above consent form and desire to participate in this study.

- ○ Yes
- ○ No

>>

Part 1

You are placed in a group together with **3 other participants**, drawn at random. Each of you will be in the same situation and have the same set of choices available.

You are each given **100 bonus points**.

You may choose to contribute anything between 0 and 100 bonus points to a common pool. The points you choose NOT to contribute, you will keep.

The sum of all bonus points contributed to the common pool will be multiplied by 2. They will then be equally distributed between all 4 participants (i.e. you and the other three participants). In other words, for each bonus point that is contributed to the common pool, by you or other participants in the group, each participant will receive 0.5 bonus points. You total earnings can therefore be calculated as follows:

My earnings = (100 bonus points - my contribution to the common pool)
+ the sum of all contributions, including mine * 0.5

Before making your choice, please answer these questions to ensure you have understood the instructions. You may use a calculator to help finding the corect answer.

Every other participant in the group chooses to contribute nothing to the common pool.

What are you earnings if you choose to contribute:

0 bonus points
50 bonus points
100 bonus points

Every other participant in the group chooses to contribute 50 to the common pool.

What are you earnings if you choose to contribute:

0 bonus points
50 bonus points
100 bonus points

Every other participant in the group chooses to contribute 100 to the common pool.

What are you earnings if you choose to contribute:

0 bonus points	
50 bonus points	
100 bonus points	

>>

How many bonus points do you wish to contribute to the common pool?

>>

Part 2

For Part 2 of this study you will receive a fixed amount of bonus points that we will announce at the end of this survey. This means that your earnings in this part will not depend on the choices you will make in any way. Your choices will however, with some probability, influence the earnings of other participants in this study.

You are randomly matched with a group consisting of 4 other participants from Part 1, the same part that you partook. You have the opportunity to decide the payment that these 4 participants will receive.

You have 2 options:

Option 1: The participants receive a fixed payment of 200 bonus points instead of the payment they would have received based on their actual choices in Part 1.

Option 2: The participants receive payments based on their actual choices in Part 1.

(Recap of Phase 1**:** Each participant will be given 100 bonus points. They may each choose to contribute anything between 0 and 100 bonus points to a common pool. The bonus points they choose not to contribute, they will keep. The number of bonus points contributed to the common pool will be multiplied by 2 and then divided equally between all 4 participants.)

Before making your choice, please answer these questions to ensure you have understood the instructions.

If you choose Option 1, what will determine the payments of the other participants?

- ◯ They will receive a fixed payment of 200 bonus points
- ◯ They will be paid based on their choices in Part 1

If you choose Option 2, what will determine the payments of the other participants?

- ◯ They will be paid based on their actual choices in Part 1
- ◯ They will receive a fixed payment of 200 bonus points

>>

You must now choose between the two payment options for the other 4 participants.

Option 1: The participants receive a fixed payment of 200 bonus points instead of the payment they would have received based on their actual choices in Part 1.

Option 2: The participants receive payments based on their actual choices in Part 1.

Which option do you choose?

- ○ Option 1
- ○ Option 2

>>

Part 2

For Part 2 of this study you will receive a fixed amount of bonus points that we will announce at the end of this survey. This means that your earnings in this part will not depend on the choices you will make in any way. Your choices will however, with some probability, influence the earnings of other participants in this study.

You are randomly matched with a group consisting of 4 other participants from Part 1, the same part that you partook. You have the opportunity to decide the payment that these 4 participants will receive.

You have 2 options:

Option 1: The participants receive a fixed payment of 150 bonus points instead of the payment they would have received based on their actual choices in Part 1.

Option 2: The participants receive payments based on their actual choices in Part 1.

(Recap of Phase 1**:** Each participant will be given 100 bonus points. They may each choose to contribute anything between 0 and 100 bonus points to a common pool. The bonus points they choose not to contribute, they will keep. The number of bonus points contributed to the common pool will be multiplied by 2 and then divided equally between all 4 participants.)

Before making your choice, please answer these questions to ensure you have understood the instructions.

If you choose Option 1, what will determine the payments of the other participants?

- ◯ They will receive a fixed payment of 150 bonus points
- ◯ They will be paid based on their choices in Part 1

If you choose Option 2, what will determine the payments of the other participants?

- ◯ They will receive a fixed payment of 150 bonus points
- ◯ They will be paid based on their choices in Part 1

>>

You must now choose between the two payment options for the other 4 participants.

Option 1: The participants receive a fixed payment of 150 bonus points instead of the payment they would have received based on their actual choices in Part 1.

Option 2: The participants receive payments based on their actual choices in Part 1.

Which option do you choose?

- ○ Option 2
- ○ Option 1

>>

In this section, you will be asked a few questions about what you believe other participants chose in Part 1.

Please take time to answer these properly.

>>

Participants on average contributed more than 50 bonus points to the common pool.

- ○ True
- ○ False
- ○ I don't know

In most groups, participants contributed a roughly equal number of bonus points to the common pool.

- ○ True
- ○ False
- ○ I don't know

>>

In this section, you will be asked a few questions about what you believe about the outcome of Part 2.

Please take time to answer these properly.

>>

The majority of participants chose Option 1, i.e., that the other participants receive a fixed payment.

- ○ True
- ○ False
- ○ I don't know

>>

Most participants would be better off with Option 1, i.e., a fixed payment, than Option 2, i.e., payments based on the actual choices they made in Part 1.

- ○ True
- ○ False
- ○ I don't know

>>

Can you please provide us with a short description of why you made the choice in Part 2.

>>

You have completed both parts of this study and we will explain now how the final amount of bonus points are generated.

The number of bonus points generated for you will depend on whether you are selected for Part 1 or Part 2 of this study.

If you are selected for Part 1, your bonus will depend on the choice made by another participant who was selected for Part 2. This means that your bonus will either be a fixed bonus payment or be determined based on the how much you and the 3 other participants contributed in Part 1.

If you are selected for Part 2, the decision you made will determine the bonus payment for 4 other participants, who were chosen for Part 1. Your bonus will be the average of the bonuses for all participants who were chosen for Part 1.

There is a 4/5 probability that you will be selected for Part 1 and a 1/5 probability that you will be selected for Part 2.

>>

You have completed the second part of the survey. We would now like to ask you five more questions before we conclude this survey.

What is your gender?

- Male
- Female

How old are you?

What is the highest level of education you have completed?

- Less than High School
- High School / GED
- Some College
- 2-year College Degree
- 4-year College Degree
- Masters Degree
- Doctoral Degree
- Professional Degree (JD, MD)

Would you describe yourself as politically on the "left" (eg. a liberal) or on the "right" (eg. a conservative)?

1 - Very liberal	2	Neutral	4	5 - Very conservative

Finally, if you have any comments or suggestions related to this study please write them down in the field below. Your feedback is very important to improve our research.

>>

Thanks for your participation! We will calculate and pay your bonus as soon as possible.

Your completion code is:

4701986

Copy and paste the code above into the survey code field on the AMT web page that directed you here at the beginning.

References

Andreoni, J. (1988, December). Why free ride? *Journal of Public Economics*, *37*(3), 291–304. Retrieved from http://dx.doi.org/10.1016/0047-2727(88)90043-6 doi: 10.1016/0047-2727(88)90043-6

Andreoni, J. (1995). Cooperation in public-goods experiments: Kindness or confusion? *The American Economic Review*, *85*(4), 891-904. Retrieved from http://www.jstor.org/stable/2118238

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics*. Princeton University Press.

Anomaly, J. (2015). Public goods and government action. *Politics, Philosophy & Economics*, *14*(2), 109-128. Retrieved from http://dx.doi.org/10.1177/1470594X13505414 doi: 10.1177/1470594X13505414

Ariely, D. (2009). *Predictably irrational: The hidden forces that shape our decisions*. Harper.

Bayer, R.-C., Renner, E., & Sausgruber, R. (2013). Confusion and learning in the voluntary contributions game. *Experimental Economics*, *16*(4), 478–496. Retrieved from http://dx.doi.org/10.1007/s10683-012-9348-2 doi: 10.1007/s10683-012-9348-2

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2017, January 4). Evaluating online labor markets for experimental research: Amazon.com's mechanical turk. *Political Analysis*, *20*(03), 351–368. Retrieved from http://dx.doi.org/10.1093/pan/mpr057 doi: 10.1093/pan/mpr057

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011, January 1). Amazon's mechanical turk. *Perspectives on Psychological Science*, *6*(1), 3–5. Retrieved from http://dx.doi.org/10.1177/1745691610393980 doi: 10.1177/1745691610393980

Cappelen, A. W., Konow, J., Sorensen, E., & Tungodden, B. (2010). Just luck: An experimental study of risk taking and fairness. *SSRN Electronic Journal*. Retrieved from http://dx.doi.org/10.2139/ssrn.1554965 doi: 10.2139/ssrn.1554965

Coase, R. H. (1960). The problem of social cost. *The Journal of Law & Economics*, *3*, 1-44. Retrieved from http://www.jstor.org/stable/724810

Conrad, J. M. (2010). *Resource economics* (2nd ed.). Cambridge University Press. doi: 10.1017/CBO9780511781087

Cowen, T. (1993). Public goods and externalities. In D. R. Henderson (Ed.), *The concise encyclopedia of economics.* Library of Economics and Liberty. Retrieved from http://www.econlib.org/library/Enc1/PublicGoodsandExternalities.html

Dworkin, G. (2017). Paternalism. In E. N. Zalta (Ed.), *The stanford encyclopedia of philosophy* (Spring 2017 ed.). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2017/entries/paternalism/.

Erkut, H., Nosenzo, D., & Sefton, M. (2015, May). Identifying social norms using coordination games: Spectators vs. stakeholders. *Economics Letters*, *130*, 28–31. Retrieved from http://dx.doi.org/10.1016/j.econlet.2015.02.021 doi: 10.1016/j.econlet.2015.02.021

Gneezy, U., List, J., & Levitt, S. (2013). *The why axis: Hidden motives and the undiscovered economics of everyday life.* PublicAffairs.

Gunnthorsdottir, A., Houser, D., & McCabe, K. (2007, February). Disposition, history and contributions in public goods experiments. *Journal of Economic Behavior & Organization*, *62*(2), 304–315. Retrieved from http://dx.doi.org/10.1016/j.jebo.2005.03.008 doi: 10.1016/j.jebo.2005.03.008

Hardin, G. (1968, December 13). The tragedy of the commons. *Science*, *162*(3859), 1243–1248. Retrieved from http://dx.doi.org/10.1126/science.162.3859.1243 doi: 10.1126/science.162.3859.1243

Hardin, G. (1998, May 1). Extensions of "the tragedy of the commons". *Science*, *280*(5364), 682–683. Retrieved from http://dx.doi.org/10.1126/science.280.5364.682 doi: 10.1126/science.280.5364.682

Hobbes, T. (1651). *Leviathan.* Penguin Books Limited.

Horton, J., Rand, D., & Zeckhauser, R. (2011, September 1). The online laboratory: conducting experiments in a real labor market. *Experimental Economics*, *14*(3), 399–425. Retrieved from http://dx.doi.org/10.1007/s10683-011-9273-9 doi: 10.1007/s10683-011-9273-9

Ipeirotis, P. G. (2010, December). Analyzing the amazon mechanical turk marketplace. *XRDS*, *17*(2), 16–21. Retrieved

from http://dx.doi.org/10.1145/1869086.1869094 doi: 10.1145/1869086.1869094

Isaac, Walker, J. M., & Williams, A. W. (1994, May). Group size and the voluntary provision of public goods. *Journal of Public Economics*, *54*(1), 1–36. Retrieved from http://dx.doi.org/10.1016/0047-2727(94)90068-x doi: 10.1016/0047-2727(94)90068-x

Isaac, R. M., & Walker, J. M. (1988, February). Group size effects in public goods provision: The voluntary contributions mechanism. *The Quarterly Journal of Economics*, *103*(1), 179–199. Retrieved from http://dx.doi.org/10.2307/1882648 doi: 10.2307/1882648

Isaac, R. M., & Walker, J. M. (1998). Nash as an organizing principle in the voluntary provision of public goods: Experimental evidence. *Experimental Economics*, *1*(3), 191–206. Retrieved from http://dx.doi.org/10.1007/BF01669342 doi: 10.1007/BF01669342

Janssen, M. C. W. (2006, October). On the strategic use of focal points in bargaining situations. *Journal of Economic Psychology*, *27*(5), 622–634. Retrieved from http://dx.doi.org/10.1016/j.joep.2006.04.006 doi: 10.1016/j.joep.2006.04.006

Kahneman, D. (2011). *Thinking, fast and slow.* New York: Penguin Books.

Kant, I. (1785). (A. W. Wood, Trans.). Yale University Press.

Kasachkoff, T. (1994). Paternalism: Does gratitude make it okay? *Social Theory and Practice*, *20*(1), 1-23. Retrieved from http://www.jstor.org/stable/23560310

Kolstad, C. (2009). *Environmental economics* (International Second Edition ed.). Oxford University Press.

Kümmerli, R., Burton-Chellew, M. N., Ross-Gillespie, A., & West, S. A. (2010, June 01). Resistance to extreme strategies, rather than prosocial preferences, can explain human cooperation in public goods games. *Proceedings of the National Academy of Sciences*, *107*(22), 10125–10130. Retrieved from http://dx.doi.org/10.1073/pnas.1000829107 doi: 10.1073/pnas.1000829107

Le Grand, J., & New, B. (2015). *Government paternalism: Nanny state or helpful friend?* Princeton University Press.

Ledyard, J. O. (1995). *Public goods: A survey of experimental research. s. 111–194 in: J. kagel/a. roth (hrsg.), handbook of experimental economics.* Princeton: Princeton University Press.

Lipford, J. (1995). Group size and the free-rider hypothesis: An examination of new evidence from churches. *Public Choice*, *83*(3-4), 291–303. Retrieved from http://dx.doi.org/10.1007/bf01047748 doi: 10.1007/bf01047748

Lloyd, W. F. (1980). W. F. Lloyd on the checks to population.

Population and Development Review, *6*(3), 473-496. Retrieved from http://www.jstor.org/stable/1972412

Malthus, T. (1798). *An essay on the principle of population.* Cambridge University Press.

Mayyasi, A. (2013, July). *Who makes below minimum wage in the mechanical turk sweatshop?* Retrieved from https://priceonomics.com/who-makes-below-minimum-wage-in-the-mechanical/

Mehta, J., Starmer, C., & Sugden, R. (1994). The nature of salience: An experimental investigation of pure coordination games. *The American Economic Review*, *84*(3), 658-673. Retrieved from http://www.jstor.org/stable/2118074

Mill, J. S. (1859). *On liberty* (G. Himmelfarb, Ed.). Penguin Books.

Mollerstrom, J., Reme, B.-A., & Sorensen, E. (2014). Luck, choice and responsibility. *SSRN Electronic Journal*. Retrieved from http://dx.doi.org/10.2139/ssrn.2408465 doi: 10.2139/ssrn.2408465

Myerson, R. B. (2009, December). Learning from schelling's strategy of conflict. *Journal of Economic Literature*, *47*(4), 1109–1125. Retrieved from http://dx.doi.org/10.1257/jel.47.4.1109 doi: 10.1257/jel.47.4.1109

Nozick, R. (1974). *Anarchy, state, and utopia.* Basic Books.

Ostrom, E. (2008). The challenge of common-pool resources. *Environment: Science and Policy for Sustainable Development*, *50*(4), 8-21. Retrieved from http://dx.doi.org/10.3200/ENVT.50.4.8-21 doi: 10.3200/ENVT.50.4.8-21

Ostrom, E., Burger, J., Field, C. B., Norgaard, R. B., & Policansky, D. (1999, April 9). Revisiting the commons: Local lessons, global challenges. *Science*, *284*(5412), 278–282. Retrieved from http://dx.doi.org/10.1126/science.284.5412.278 doi: 10.1126/science.284.5412.278

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on amazon mechanical turk. *Judgment and Decision Making*, *5*(5), 411-419. Retrieved from http://EconPapers.repec.org/RePEc:jdm:journl:v:5:y:2010:i:5:p:411-419

Pischke, J.-S. (2012, July). *Probit better than LPM?* Retrieved from http://www.mostlyharmlesseconometrics.com/2012/07/probit-better-than-lpm/

Rawls, J. (1971). *A theory of justice.* Harvard University Press.

Sally, D. (1995, January 01). Conversation and cooperation in social dilemmas: A Meta-Analysis of experiments from 1958

to 1992. *Rationality and Society*, *7*(1), 58–92. Retrieved from http://dx.doi.org/10.1177/1043463195007001004 doi: 10.1177/1043463195007001004

Schmidtz, D. (1991). *The limits of government: An essay on the public goods argument*. Westview Press. Retrieved from https://books.google.no/books?id=OjjGAAAAIAAJ

Sonnemans, J., Schram, A., & Offerman, T. (1998). Public good provision and public bad prevention: The effect of framing. *Journal of Economic Behavior & Organization*, *34*(1), 143 - 161. Retrieved from http://www.sciencedirect.com/science/article/pii/S0167268197000425 doi: http://dx.doi.org/10.1016/S0167-2681(97)00042-5

Sugden, R. (1995). A theory of focal points. *The Economic Journal*, *105*(430), 533-550. Retrieved from http://www.jstor.org/stable/2235016

Thaler, R. H., & Sunstein, C. B. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.

von Mises, L. (1949). *Human action, the scholar's edition*. Ludwig von Mises Institute.

Wilkinson, N., & Klaes, M. (2012). *An introduction to behavioral economics*. Palgrave Macmillan.

Willinger, M., & Ziegelmeyer, A. (1999). Framing and cooperation in public good games: an experiment with an interior solution. *Economics Letters*, *65*(3), 323 - 328. Retrieved from http://www.sciencedirect.com/science/article/pii/S0165176599001779 doi: http://dx.doi.org/10.1016/S0165-1765(99)00177-9