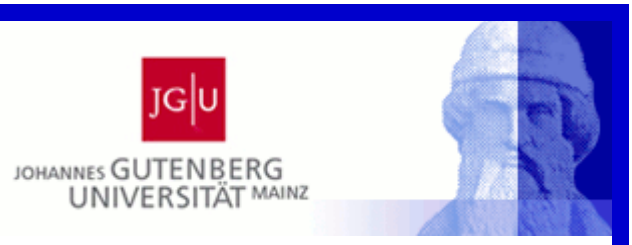


Gutenberg School of Management and Economics

& Research Unit “Interdisciplinary Public Policy”

Discussion Paper Series

Complementary Consumer Responsibility — The Limits to Immoral Delegation in Markets

Mario Scharfbillig

July 30, 2019

Discussion paper number 1909

Johannes Gutenberg University Mainz
Gutenberg School of Management and Economics
Jakob-Welder-Weg 9
55128 Mainz
Germany
https://wiwi.uni-mainz.de/

Contact details

Mario Scharfbillig
Chair of Public and Behavioral Economic
Gutenberg School of Management and Economics
University of Mainz
Jakob-Welder-Weg 9
55128 Mainz
Germany
mario.scharfbillig@uni-mainz.de

COMPLEMENTARY CONSUMER RESPONSIBILITY – THE LIMITS TO IMMORAL DELEGATION IN MARKETS*

Mario Scharfbillig†

July 30, 2019

Abstract

Delegation has been shown to facilitate individual immoral behavior. It is however unclear, if these findings extend to markets, where consumers may punish firms who delegate immoral production decisions. I address this question by employing an experimental market paradigm, involving an unfair product, containing a negative externality, and a fair product without externality. Passive delegation of the production decision, with random matching between an owner and a seller, leads to a lower share of the fair product being traded, consistent with the findings on responsibility diffusion. Active delegation in contrast, where owners have a choice over sellers first, increases the share of the fair product relative to passive delegation. Responsibility norm beliefs support a mechanism of complementary consumer responsibility, which assigns more responsibility to consumers when owners have a choice over sellers and, therefore, over the product type offered. Consumers' buying decisions may therefore limit the possibility for delegating immoral behavior, depending on the specific design of the delegation.

Keywords: Social responsibility, market game, corporate social responsibility, consumer social responsibility, responsibility diffusion

JEL codes: C92, D62, M14, D63

*I would like to thank for numerous helpful comments from Björn Bartling, Charles Bellemare, Florian Hett, Christopher Koch, Michael Kosfeld, Sabine Kröger, Daniel Schunk and the participants of the Brown Bag Seminar at the University of Mainz, the ESA World Meeting 2018, the IfW Behavioral Economics Seminar and the Jeudi-Midi Seminar at the Université Laval. I received funding for my project from the Research Priority Program Interdisciplinary Public Policy (IPP) at the JGU. I received financial support from the research priority program Interdisciplinary Public Policy (IPP) at the University of Mainz.

†Johannes Gutenberg University of Mainz, Jakob-Welder-Weg 4, 55128 Mainz, Germany, Phone: +49 6131 39 22918, Fax: +49 6131 39 27695, mario.scharfbillig@uni-mainz.de

1 INTRODUCTION

Delegating decisions in corporate settings is ubiquitous within and between firms. Firms themselves almost always consist of multiple levels of delegation (Appelbaum and Batt, 2014; Berle and Means, 1932; Jensen and Meckling, 1976), while, in recent decades, we also have seen an unprecedented level of outsourcing along global value-chains, which are ever increasing in depth. In the US alone, for example, the share of people working in firms in-between producers and end consumers of the total workforce in 2011 is estimated to be at 34%, up from 25% in 1999 (Krakovsky, 2015; Spulber, 1999).

One potential motivation for delegation on the seller side, as proposed by Hamman et al. (2010), is that it allows firms to engage indirectly in immoral but profitable behavior, because principals may prefer to let someone else do the "dirty work" for them. Indeed, many firms seem to engage in strategic delegation of perceived immoral behavior. Examples are dramatically raising prices on life saving drugs through subsidiaries (Berenson, 2006), violent displacement of local farmers through "Coyotes" (Brodzinky, 2013; Moskin, 2004), poaching of employees through headhunters (Finlay and Coverdill, 2007) or firing employees through "termination assistance" (Taub, 2016).

In contrast to the seller side argument, there is evidence that consumers are willing to punish immoral behavior of firms by actively boycotting them or simply buying products from competitors, whom they perceive to not act immorally (e.g. Herkenhoff and Krautheim, 2018; Harrison and Scorse, 2010). Experimental evidence on markets with externalities supports this finding, namely a significant willingness to trade-off personal gains by consumers for socially preferred production, deciding not to harm others (Bartling et al., 2017a, 2015; Kirchler et al., 2015; Pigors and Rockenbach, 2016a,b).

This raises the question about the effectiveness of delegation of immoral behavior in market settings, where consumers may potentially punish it. In this article, I shed light on the question how delegation on the seller side influences market outcomes. More specifically, I investigate whether delegation reduces the socially responsible purchase behavior in a market with products, which differ in their moral impact. The main concern is, whether delegation indeed increases immoral behavior or if consumers potentially offset

the effect. Research on this important question is so far lacking.

It is not clear how consumers react to delegation, given the prevalence for it on the one hand, and the evidence on consumer protests and boycotts on the other hand. Potentially, consumers could only look at the final product on the market, and, therefore, their choice should not be affected by the delegation decision at all. This would mean, that the possibility of delegation would only lead to more immoral behavior through the seller side as long as it is profitable (Hamman et al., 2010). Alternatively, consumers could also feel less responsible themselves, because overall one more person or entity is involved. When more people are involved in an immoral act, immoral behavior is more likely (Behnk et al., 2017; Wiltermuth, 2011). This finding is usually attributed to diffusion of responsibility (Darley and Latane, 1968). This would then lead to an additional rationale for delegating immoral behavior, on top of the seller side argument. This argument is also in line with the idea, that not buying from a firm is considered as a sort of punishment (or reward for the competitor). For example, Bartling and Fischbacher (2012) find that by delegating away a choice to harm others, an individual can deflect punishment by third parties (see also Oexl and Grossman (2013)). In a similar vein, in delegation in an ultimatum game setup, Fershtman and Gneezy (2001) propose that responders may shy away from rejecting offers, because they would also hurt delegates. If consumers feel that their decision might hurt delegates in the market setup, they would potentially shy away from punishing immoral product offers and therefore not switch to the competitor. However, the findings on the possibility to avert punishment in Bartling and Fischbacher (2012) are from an individuals perspective, i.e. individuals may avoid punishment to a certain extent by delegating a decision away. In contrast, the *joint* punishment of delegator and delegate, when the delegate chooses the unfair option, was the same or even numerically larger than for an individual alone making the unfair decision. Because when it comes to buying a product, the decision to reward a specific offer can only be made in total – buying or not buying – consumers may decide more often to buy fair. Consumers may be also exploitation averse. When the firm chooses a delegate strategically in order to act immorally, consumers may be averse to being an accomplice to the immoral action in

buying from this firm, because they could feel being exploited, leading them to buy the moral product (Carpenter and Dolifka, 2017). In sum, consumer behavior may very well worsen the consequences of delegation on immoral behavior, or, in contrast, discipline it.

This study is built on the repeated posted offer market framework of Bartling et al. (2015). In the experimental market, there are two options of producing a good, one costly for the seller, the fair product, and one which destroys resources of an uninvolved third party, the unfair product. Sellers select the product type and set prices. More sellers than buyers compete to sell a product without replacement, leading to buyer market power. In this setting, Bartling et al. (2015) find that even after several rounds of play, the fair product is found to have a stable and considerable market share, but that the unfair product is more profitable for sellers and buyers. These characteristics, a monetary incentive to sell the unfair product and considerable deviation from the pure profit maximizing behavior, make the design useful for studying the impact of delegation on socially responsible purchase behavior.

Using their setup as baseline, I introduce two additional treatments in a between subject design. In the first treatment, delegation is passive, i.e. sellers share the profit with an owner, but the owner is inactive throughout the experiment. This type of delegation can be thought of as in modern publicly listed companies with passively invested shareholders. Typically, shareholder do not actively engage with management and, especially in cases of index funds, often cannot even sell their shares.

In the second treatment, delegation is active. Before sellers make a decision on the product type and the pricing, owners face a selection stage, where they choose a seller based on past behavior. This type of delegation represents a more active investment approach, where shareholders can "fire" non-performing managers (e.g. Parrino et al., 2003). This setting may then induce more immoral behavior on the seller side through two channels. First, the selection stage allows for the selection of those sellers, who are willing to act immorally (Hamman et al., 2010). Second, the selection allows for an additional layer of the replacement logic, i.e. sellers can tell themselves "if I do not do it, someone else will" (Bartling and Özdemir, 2017). For buyers, relative to the passive delegation

case, monetarily nothing changes. That means, the payoff structure for everyone involved in a trade remains exactly the same. Therefore, any changes in purchase behavior are a direct result from the fact that owners can choose a seller. For example, buyers may feel more important in the exchange between seller and buyer, as the sellers' offering power is limited by owner selection.

Further, I provide evidence to elucidate why delegation may affect the market outcome. A key mechanism of proliferation of immoral behavior, which has been proposed in connection with delegation, is responsibility diffusion (Falk and Szech, 2014; Paharia et al., 2009, 2012). Responsibility diffusion is a prime suspect for the potential effect of delegation, because it may enable individuals to maintain a positive self-image, while letting others do the "dirty work" (Benabou and Tirole, 2010; Dana et al., 2007). A novelty in my study is, that I propose a direct test of responsibility diffusion using a norm elicitation game introduced by Krupka and Weber (2013). Before proceeding to the market stage of the game, subjects need to provide their beliefs on who is responsible for the product type being traded. This allows me to directly observe responsibility norm changes in response to different market structures.

I find that passive delegation leads to a higher share of the unfair product being traded relative to the baseline without delegation. Over all periods, the share of the fair product in the passive treatment is consistently and significantly below the fair product share in the baseline by 16 percentage points on average. My finding therefore confirms the findings in non-market settings, where the presence of others, who potentially benefit from immoral actions, increases immoral behavior. In contrast, when owners select sellers in the active treatment, the fair product share is not significantly different from the baseline treatment and consistently above the passive treatment by 21 percentage points on average. This means that active selection of sellers leads to a relatively fairer outcome in the market with the highest overall payoff for third parties.

The results from the responsibility norm elicitation reveal a surprising result. Buyers are seen as significantly less responsible for the product type being traded in the passive treatment compared to the baseline. This means that in the passive treatment, when

simply more individuals share the proceeds of the trade, buyer's responsibility is indeed reduced. Turning to the active delegation treatment, the responsibility assignment of owners increases relative to the passive treatment, presumably because they are seen to have influence on the product type through their choice of sellers. Given the findings of Hamman et al. (2010), the influence of owners is likely to select sellers based on profit, which would pressure sellers to offer more unfair products. I confirm this tendency in separate tests. However, at the same time, buyers are seen as more responsible for the product type being traded. When owners have at least indirectly influence on the product type decision, buyers seem to be assigned similarly increased responsibility, potentially to limit owners influence. The latter finding is in line with the joint punishment of delegator and delegate when delegating the decision to harm others away (Bartling and Fischbacher, 2012). I call this joint increase in responsibility for the buyer side 'complementary consumer responsibility'.

Investigating individual behavior suggests, that the answers from the responsibility norm elicitation are indeed meaningful and support the interpretation of complementary consumer responsibility. Controlling for availability of both types of products and price differences buyers are facing in their purchase decisions, they are less likely to buy a fair product in the passive treatment compared to the active treatment and baseline. This means, that for buyers, active delegation leads to a change in purchase behavior, offsetting potential monetary benefits from delegating immoral behavior on the seller side.

Overall, I interpret the findings from the responsibility norm elicitation and buyer purchase behavior as complementary consumer responsibility in response to the structure of delegation. When the seller side shares profits in the passive treatment and therefore more people profit from the immoral act, the market outcome is less fair then in the baseline. As soon as owners select sellers, however, buyers are seen as more responsible for the product being sold in the market. This means that the possibility of influence of owners triggers the complementary responsibility norm on buyers, which matches their buying decisions. The purchase behavior is important, because I find that owners select sellers solely based on profit. Since the profit from selling an unfair product is higher, it

is essential that sellers care more about the product type they buy.

The findings are important for the rationale of delegation. While other authors find a significant increase in immoral behavior from allowing delegation, for example in higher shares for a dictator in a dictator game (Hamman et al., 2010) and benefits through a lower punishment by third parties for unfair allocations (Bartling and Fischbacher, 2012), when used in a market setting, benefits are not clear. Consumers may very well limit the immoral proliferation of delegation if they know that owners of firms have a choice. This implies that, in enforcing socially responsible behavior in markets, transparency regarding the fact that firms do have a choice over which firms or individuals they delegate to, may have a positive impact with regards to externality reduction.

The findings could be interpreted in terms of self-image concerns. When buyers know that sellers are selected, they may receive more self-image utility from buying fair, e.g. because they can now reward two individuals who are both involved in offering a fair product or punish the opposing side. The results are robust to a number of alternative specifications and are not driven by mistaken understanding of the experiment or other social preference considerations.

The remainder of the paper is structured as follows. Section 2 gives an overview on the related literature. Section 3 describes the experimental design and predictions. Section 4 reports my results and section 5 concludes.

2 RELATED LITERATURE

This paper contributes to three strands of literature. The first and most important one is research on delegation. Research on delegation has a long history in economics (Berle and Means, 1932; Fehr and Gächter, 2000; Stigler and Friedland, 1983), and has been studying mostly optimal incentive structures under imperfect contracts (Tirole, 2002). Hamman et al. (2010) are the first to explicitly include intended immorality into the analysis and show that the moral dimension provides an additional rationale for delegation by facilitating immoral behavior from which principals can profit. They look at the sharing decision in a dictator game, where the sharing decision can be delegated to competing

agents. They find that people prefer letting other people take their immoral decisions (giving a higher share to themselves) and that agents are willing to do the immoral task when incentivized. They interpret their results as evidence for an additional rationale for firms to delegate.

However, firms usually need to sell their products to buyers, who may limit the possibility of immoral delegation on the seller side (e.g. Pigors and Rockenbach, 2016a). I do find that market interaction is indeed crucial for the outcome. Depending on the structure of delegation within the market, namely whether owners have a choice to select sellers or not, consumers feel more responsible and essentially negate the rationale for delegation as a means to facilitate immoral behavior.

Another important difference to the study of Hamman et al. (2010), is that a skewed distribution decision in a dictator game can be easily rationalized as not being immoral, as the agent has the full amount of money at free disposal and the setting may simply change the perception of deservedness (List, 2007; Cappelen et al., 2013). In my treatment however, the externality means, that subjects actively choose to destroy resources from a third party and therefore, the setting has more relevance to the context of corporate social responsibility, for example for situations like worker abuse in clothing production (Clean Clothing Campaign, 2017).

A second strand in the literature links responsibility diffusion to utility, trying to show why responsibility diffusion enables individuals to act more selfishly. Dana et al. (2007) introduced the idea that people do exploit elements of the choice architecture, termed the "moral wiggle room", in order to act more selfishly. This idea has been tied to the concept of self-image, where part of the evidence for social preferences (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000) is interpreted as a wish to maintain a positive perspective on one-self (Bénabou and Tirole, 2006; Bauer et al., 2005; Fiske, 2013). Responsibility diffusion through delegation is mentioned as a potential driver of the finding of Hamman et al. (2010), enabling individuals to maintain a positive self-image by letting someone else do the immoral act. Because in their study delegation is costless and dictators have full control over the choice set, they can interpret the choice of delegation as a pure

preference for delegation. Similarly, Behnk et al. (2017) show that even without the actions of the delegate, the delegation setting may decrease responsibility. In their study, the mere presence of a second person responsible for the action without any interaction leads to a more antisocial outcome and less feeling of inappropriateness of the action by the sending, as well as the receiving side in a sender-receiver deception game. They define responsibility diffusion as "reduction in the intrinsic disutility individuals incur from acting antisocially because more people are involved in the decision-making" (Behnk et al., 2017, p.1). Wiltermuth (2011) shows that, on top of the effect of the presence of another individual, splitting the profit with someone else makes people again more likely to cheat. Again, these studies look only at cases where one individual enacts the outcome.

I contribute to this literature by showing that in market interactions with a product involving a loss for someone else, responsibility assignments for buyers and sellers differ. For sellers there is no objective responsibility diffusion. Participants in the control and both treatments see sellers equally responsible. This points to the fact that delegation does not universally lead to responsibility diffusion. On the contrary, delegation on the seller side has an influence on the responsibility of buyers. Merely adding an individual, in this case a passive owner, to the market context seems to reduce the burden of responsibility of buyers. However, when owners are involved in the product choice through selecting sellers, buyers are seen as more responsible, which I interpret as complementary consumer responsibility. This finding relates to findings on indirect measures of responsibility through punishment. People can avoid individual punishment by delegating (Bartling and Fischbacher, 2012; Oexl and Grossman, 2013; Hill, 2015), but, in case of unfair behavior, punishment for delegator and delegate together is the same or higher than for an individual taking the decision alone. I therefore show that this observation holds in a more complex environment, where buyers can be seen as the evaluating third party. Additionally, market interactions comprise day-to-day interactions, which are constantly present in peoples' lives (Engelmann et al., 2018), which makes them an important facet of life to study for responsibility diffusion.

A third strand of literature is that of socially responsible firm behavior. In a market

setting, Bartling et al. (2015) show that, when firms can produce at a lower cost but with an externality, roughly 40% of participants trade the product without externality. This result is surprisingly robust to increased competition and to making the information on the production technology costly for consumers. In a follow-up paper, Bartling et al. (2017b) show that the result is also robust to placing the externality upon one or several individuals. Other studies extend their design or use a similar one and find that socially responsible products survive in different settings on the market (Friedrichsen, 2017; Kirchler et al., 2015; Pigors and Rockenbach, 2016a,b). Interestingly, Pigors and Rockenbach (2016b) show that consumers are sensitive to the market structure. They compare the willingness to pay a social premium in a monopoly and a duopoly. They find that in a duopoly, consumers are more willing to pay for the social product, potentially because it has a signaling value to consumers. An interesting paper of Jakob et al. (2017) takes another perspective and shows that social responsibility (or moral responsibility as they call it) may be a hurdle when trying to use market instruments to increase efficiency. In my case however, the moral behavior is clearly also efficient in overall terms. I extend this strand of literature, by studying if (corporate) social responsibility is robust to delegation. Delegation is a universal feature of the modern business world, and therefore an important aspect in the context of corporate social responsibility. In my setting, delegation influences the market structure indirectly through the seller side only. My findings show that the concern for the externality is only partially robust to delegation. Only when owners do have a choice over sellers do consumers feel responsible enough to negate the negative effect of delegation on the fair market share.

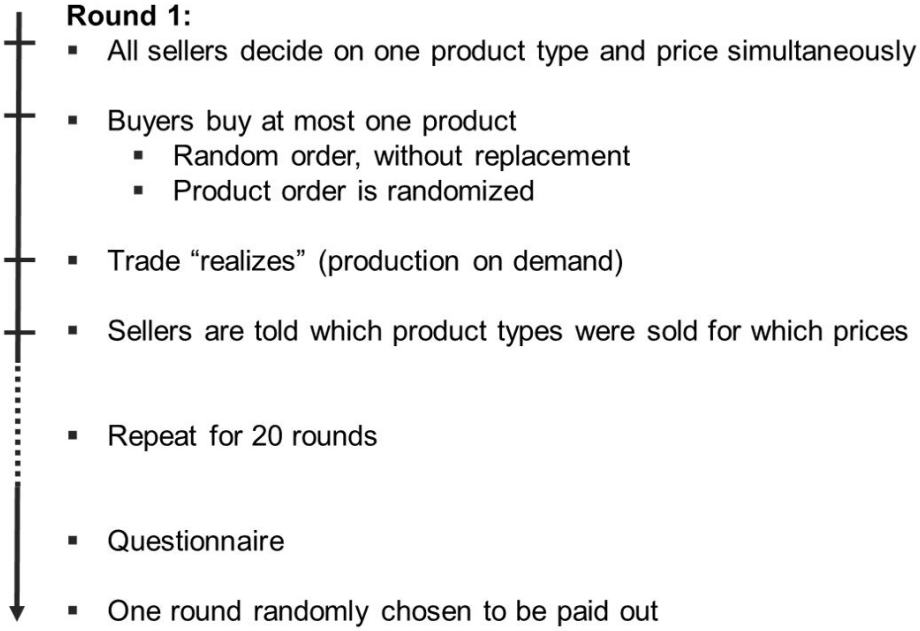
3 STUDY DESIGN

3.1 Experimental Design

The experiment extends the setting from Bartling et al. (2015) by the dimension of delegation. The baseline corresponds, with minor changes, to their original design. It consists of 5 sellers, 4 buyers and 4 third parties within one group, whose roles remain fix through-

out the experiment. The market interaction consists of 21 identical rounds, one of which is randomly chosen to be paid out at the end.

Figure 1: Overview of Experiment Stages



In each round, all sellers simultaneously choose one out of two product types to offer. The fair product costs c=20 to produce, whereas the unfair product costs c=0, but causes an externality e=1, which deducts 60 from an unrelated third party. Both products are worth v=50 to buyers. After sellers made the offer decision, the buyers decide to buy at most one product sequentially without replacement. Products are presented in a random order, such that sellers cannot build up reputation with buyers. The order in which buyers can buy the product, is randomly drawn each round. The third party is randomly matched each round with a buyer, for whom they face the consequence of the trade. This means that production is on demand. If an unfair product is not bought, no negative externality realizes for third parties. Each participant has an endowment of 100 each round.

After each round, sellers learn about all product types being sold for which price, buyers only see their own payoff and the consequences on their matched third party. The market is then repeated for 21 rounds. Afterwards, all participants answer a survey on a

range of socio-demographic questions. The order of the experiment is shown in figure 1.

In order to eliminate all inequality concerns unrelated to the trade decision, each player type starts with an equal endowment of 100. A summary of payoffs in the baseline is given by:

$$\Pi^{Seller}_{Baseline} = \begin{cases} 100 + p - (1-e)c & \text{if seller sells a product} \\ 100 & \text{otherwise} \end{cases} \tag{1}$$

$$\Pi^{Buyer}_{Baseline} = \begin{cases} 100 + 50 - p & \text{if buyer buys a product} \\ 100 & \text{otherwise} \end{cases} \tag{2}$$

$$\Pi^{ThirdParty}_{Baseline} = \begin{cases} 100 - 60e & \text{if buyer buys a product} \\ 100 & \text{otherwise} \end{cases} \tag{3}$$

In the new active treatment, delegation on the seller side in the spirit of Hamman et al. (2010) is introduced. Before sellers make a decision on the product type and its price, the new player type owner selects consecutively at most one seller to take this decision for him. Owners must delegate. In order to induce competition among sellers, in this treatment there are 5 owners and 6 sellers to choose from, one of which will eventually be left out. In all rounds following the first, owners receive a summary of last rounds transaction history. Furthermore, sellers keep their ID, such that they can built up reputation with owners over time. Sellers are incentivized to act in the interest of the owners by receiving half of the profit, the other half goes to the owner. Sellers also see, in which order they are selected by owners. Other than the selection of the seller, there is no interaction between owner and seller.

The presence of the owner and profit sharing between owner and seller itself can have an effect on the willingness to offer a fair product. Therefore, the active delegation treatment will also be compared to a passive delegation treatment, where owners and sellers are randomly matched to each other and still share the profit, but otherwise do not have any interaction.

A summary of payoffs for sellers and owners in the two delegation treatments is given

by equation (4) and (5). For third parties the payoff situation remains unchanged.

$$\Pi^{Seller}_{Delegation} = \begin{cases} 100 + \frac{1}{2}\left(p - (1-e)c\right) & \text{if seller sells a product} \\ 100 & \text{otherwise} \end{cases} \tag{4}$$

$$\Pi^{Owner}_{Delegation} = \begin{cases} 100 + \frac{1}{2}\left(p - (1-e)c\right) & \text{if matched seller sells a product} \\ 100 & \text{otherwise} \end{cases} \tag{5}$$

Additionally, I test for responsibility diffusion directly. Simply asking subjects on how responsible they feel however, may result in players using the elicited beliefs as ex-ante "justification" of their own behavior (Epley and Gilovich, 2016; Gino et al., 2016). Therefore, I include a responsibility assignment question incentivized in the spirit of a norm elicitation game by Krupka and Weber (2013). Before the market interaction starts, all subjects in each treatment were asked about their beliefs on the level of responsibility of all participant types involved in a trade for the product type being traded. The subjects play the elicitation game after they learned their player type.[1] The exact wording of the question is: "In the following, we will ask you for your assessment on the question, who is responsible in the following market for the type of product that is ultimately being traded. This means, when a product is being traded, who is responsible which type of product (with or without loss to the third party) is traded?" Participants can then rate for each of the player types involved (buyer, seller, third party and potentially owner) on a scale from "Not at all responsible", "Rather not responsible", "Rather responsible" to "Very responsible". The task is incentivized by a payout of additional 25 points, if the given answer matches the answer of the mode of all answers in a session. In order to decrease experimenter demand effects for the product type, subjects were also asked about their beliefs of responsibility for the final price of the product being traded. The order of the question was randomized. One of the questions is drawn at random to be paid out at the

[1]Because subjects learn their role before they make their choice on who they think is responsible, they may assign less responsibility to themselves and/or more to others. I do not find any evidence for this when comparing responsibility ratings of their own type with that of others for the same type.

end of the experiment. The incentivization makes the question a coordination game in the nature of a beauty-contest, with the only incentive for choosing an answer based on the beliefs about what the others believe.

3.2 Predictions

Bartling et al. (2015) find a constant share of around 40% of the fair product, but for production costs of 10. The production cost burden of the fair product is equally shared by buyers and sellers, and both gain from trading with each other. Buyers are found to be price sensitive, i.e. they are more likely to buy a fair product when the price of the fair product is lower or the price of the unfair product is higher. This makes their setting ideal for investigating my research question, namely if delegation induces sellers to be more willing to sell an unfair product. Doing so is more profitable than offering the fair product and buyers may be swayed to buy the unfair product depending on price differences.

Note however that, if all subjects are maximizing own profits without consideration for fairness or other subjects' earnings, all treatments should have the same outcome, a market share of zero for the fair product and a prize equal to the marginal cost of each product type. Buyers will reap all benefits of the market, irrespective of the treatment.

This prediction differs, when subjects have other-regarding preferences. The most prevalent ones are outcome dependent social preferences, like inequality aversion (Bolton and Ockenfels, 2000; Fehr and Schmidt, 1999).[2] In their analysis, Bartling et al. (2015) assume that buyers as well as sellers do not care about the utility of the other market participants (as they are traded off according to their own behavior in the market), but about the third party, because their decision directly affects them without the third party being able to act. They also employ the distribution of inequality averse types from

[2]Alternatively, efficiency seeking and maxmin preferences (Charness and Rabin, 2002) may be relevant. However, in my design the unfair product is overall welfare destroying (-60 vs. a total valuation of 50). If all participants exhibit efficiency seeking preferences, there will be no differences between the baseline and any of the treatments, because the overall amount of resources is always the same. The same holds for maxmin preferences, as in all cases, the maximum of the minimum payment will be zero, which is always the best outcome for the third party anyway. If either all or a share of the participants exhibit these two types of preferences, then there should not be any difference between the markets, because randomization ensures the percentage of these preferences to be distributed equally between the roles.

Fehr and Schmidt (1999) to derive their equilibrium. I follow their assumptions as well and further assume that sellers and buyers consider the utility of the owners involved in the trade, but not of the uninvolved owners. As a result, as shown in the appendix, sellers in both the passive and the active treatments are equally or even less willing to offer the unfair product to buyers in both treatments compared to the baseline. The intuition behind this result is, that sellers personally gain less from selling the unfair product compared to the baseline, even assuming they receive positive profits (which in equilibrium they should not). The same holds true for owners, who earn exactly the same due to the equal profit sharing, which should reflect in their choice on sellers. There is also no difference between the passive and active treatment, because the monetary outcome is exactly the same. For buyers, there is no difference in behavior, because ultimately, it is irrelevant how many people share the producing side surplus, as long as it stays the same in absolute terms.

Therefore, for the above made standard assumptions on outcome dependent preferences, in equilibrium, the share of the fair product should be exactly the same between both delegation treatments. The result seems somewhat surprising, but is a direct result of the fact that monetarily, there is no difference between passive and active delegation. Furthermore, because individuals are randomly distributed into different roles, the selection power of owners also has no effect, because each owner can hire the corresponding type of individual for his preferences. When comparing the delegation treatments with the baseline, the share of the fair product will either stay the same or may even go up.[3] In the baseline, individuals with $\beta = 0.25$ are indifferent between selling or buying the unfair or fair product, while in both delegation treatments, sellers are no longer willing to do so, sellers personally gain less profit from the trade, while incurring the whole utility cost of being ahead vis-à-vis the third party. This means that in the baseline, when all indifferent individuals choose the fair product, the share of the fair product should be exactly the same for all markets, in all other cases the share of the fair product should go

[3]The typical inequality aversion utility model formulates utility for a wealth bundle x for the two player case as $U_i(x) = x_i - \alpha_i max(x_j - x_i) - \beta_i max(x_i - x_j)$, where α_i is aversion to disadvantages inequality and β_i is aversion to advantages inequality.

equally up in the delegation treatments.

In contrast to the predictions from outcome dependent preferences, following the literature on responsibility diffusion through delegation, it seems unlikely that in the delegation treatments, sellers feel the full disutility from the advantageous inequality vis-à-vis the third party, if they are delegates of owners. Responsibility may differ both objectively, when shared by all subjects in the norm, or subjectively, when individuals deviate with their perception from the norm. This may allow all subjects involved to act more selfish (buy/sell unfair) while maintaining a positive self-image. In the passive treatment, responsibility diffusion only works through the presence of another involved individual and profit sharing. In the active treatment, on top of the profit sharing, comes the selection of sellers by the owners. Owners do see seller's actions in the market and receive a summary of the results, when selecting a seller for the next round. If they feel less responsible and therefore do not fully incorporate the loss of the third party into their utility, they can choose sellers, who only offer profit maximizing product types. This in turn may incentivize sellers to offer more unfair products, even if they would, without the selection, do not do so. As described above, the effect of passive and active delegation on buyers is ex ante not clear, depending on if they focus on the product itself or take the seller side into account. If they take the seller side into account, having one more individual involved who profits from the immoral action may reduce their incorporation of the loss to the third party as well. For the active delegation, the effect is less obvious.

Therefore, I base my predictions for the market outcome on the seller side arguments:

$$\theta_{Baseline} > \theta_{Passive} > \theta_{Active} \tag{6}$$

where θ_i= market share of fair product in market i.

The answers from the norm elicitation game should reflect the findings from the market for buyers and sellers. For owners, it seems obvious that they are more responsible if they have a choice over sellers. Third parties are passive in all treatments and therefore their responsibility should stay the same at a very low level. Consequently, the prediction for

responsibility is:

$$\gamma^{Seller}_{Baseline} > \gamma^{Seller}_{Passive} > \gamma^{Seller}_{Active} \tag{7}$$

$$\gamma^{Buyer}_{Baseline} > \gamma^{Buyer}_{Passive} = \gamma^{Buyer}_{Active} \tag{8}$$

$$\gamma^{Owner}_{Passive} < \gamma^{Owner}_{Active} \tag{9}$$

$$\gamma^{ThirdParty}_{Baseline} = \gamma^{ThirdParty}_{Passive} = \gamma^{ThirdParty}_{Active} \tag{10}$$

where γ_i = mean responsibility norm assignment in market i for the respective player type.

3.3 General Information and Procedures

The experiment has been conducted in the Mainz Behavioural and Experimental Laboratory (MABELLA) between February and August 2018 with the general subject pool from the Johannes Gutenberg University Mainz. The participants were invited to the lab using ORSEE (Greiner, 2015). Table 1 gives an overview of the number of participants per treatment. The experiment employed a total of 300 subjects in 6 markets per treatment, 78 subjects in the baseline, 108 in the passive delegation treatment and 114 in the active delegation treatment.

Table 1: Session Overview

Treatment	Markets	Buyers	Sellers	Owners	Third parties	Overall
Baseline	6	24	30	0	24	78
Passive Delegation	6	24	30	30	24	108
Active Delegation	6	24	36	30	24	114

The study was conducted using the software oTree (Chen et al., 2016). Seats within the lab were assigned by randomly drawing seat cards. The seats within the lab were randomly assigned the role to be played in the experiment. Each subject learned from the screen which role they had, before reading the written instructions, explaining the experimental

procedures. The instructions also included control questions, which participants were asked to fill out and which were gathered at the end of the experiment. Questions were answered at the seat of the participant. Sessions lasted on average 1.5 hours. Payoffs within the experiment were called "points" and converted into Euro at the rate of 100 points to 9€ (1€= 1.2$ at the time of the study). On average, individuals earned 11.80€.

Passive participants, like third parties or passive owners played a different and unrelated experiment after they read the instructions and answered the norm elicitation game. None of the other participants knew about this fact or doubted it.[4] Additionally, a total of 9 passive individuals (owners and third parties) where not present in the market elicitation at the same time as the active participants and were elicited later.[5]

4 RESULTS

4.1 Fairness in the Market

The main focus of this paper is the comparison of the market share of the fair product being traded between the two delegation treatments and the baseline market. Figure 2 displays the main result of the experiment. The solid line represents the share of fair product purchases. The share is calculated on all completed trades, which represent 97% of the total cases buyers were asked to make a purchase decision (1465 out of 1512).[6] Similar to Bartling et al. (2015), in order to smooth out random variations over time, the data is presented in three period blocks only in the figure. The share of the fair product in the baseline reduces slightly over time from around 60% to 45%. This share is not directly comparable to the original study, because they have production costs of either 10 or 40. Nevertheless, the market share seems to be slightly above their findings, even for c=10, which may be due to the responsibility elicitation question asked directly before

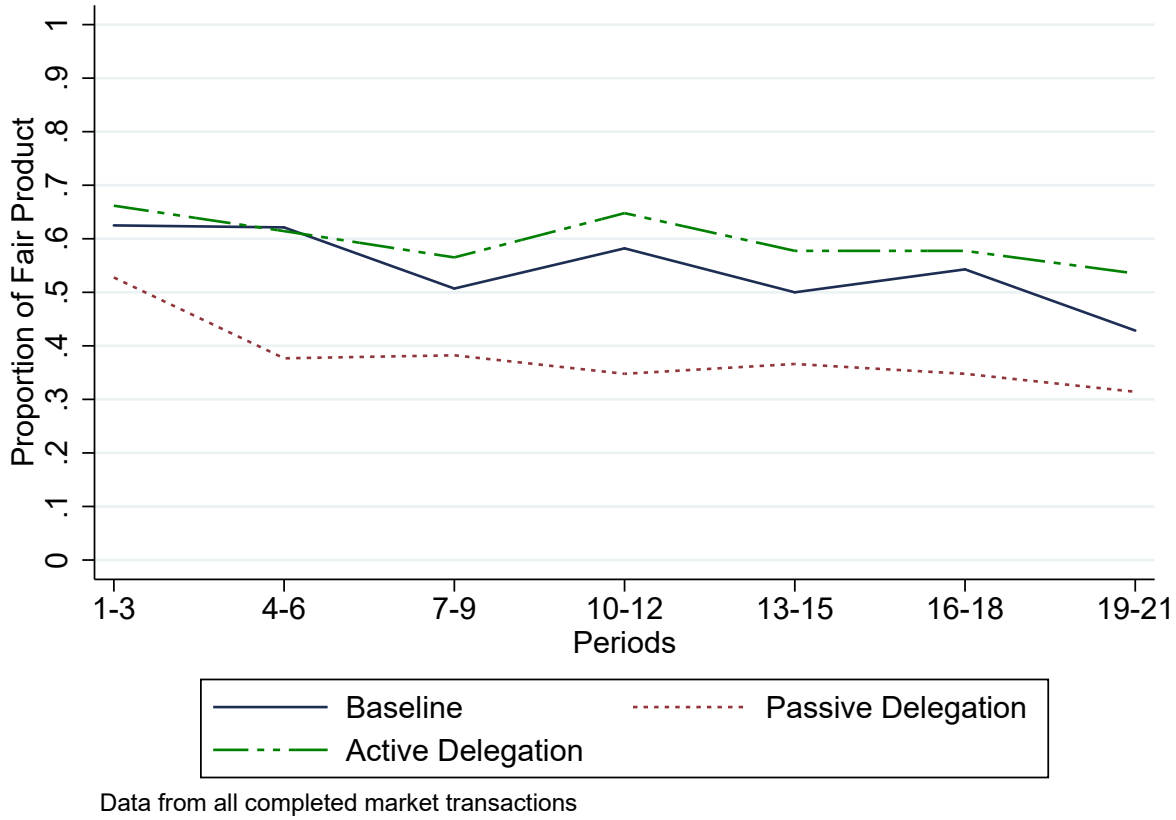
[4]Payouts for this additional experiment are not included in the above mentioned average payout.

[5]Because many subjects needed to be present for the experiment to start, several reserve participants were invited. However, in some sessions (5 overall), there were either 2 or 4 subjects too little or too much (3 times control treatment, 2 times active treatment). There is no significant effect of different presence of less or more third parties or owners on the market outcome.

[6]Note that buyers can also choose to not buy any product. There are no significant differences in not buying between the treatments with p=0.770, 0.538 and 0.695 for baseline vs. passive, baseline vs. active and passive vs. active treatment respectively.

the market interaction.

Figure 2: Market Share of Fair Product



Data from all completed market transactions

The market share of the fair product in the passive treatment is consistently below the baseline market. Having to share the profit on the producer side leads to a lower market share of the fair product, consistent with the idea that profit sharing and the presence of another individual facilitates immoral actions (Behnk et al., 2017; Wiltermuth, 2011). Counter to the predictions though, the market share of the fair product in the active treatment is consistently above the baseline and the passive delegation treatment.

Table 2 reports the corresponding regressions for the treatment effect. In order to correct for multiple observations per individual, I estimate a panel probit model with individual random effects with robust standard errors clustered at the market session level.[7] The results are also robust to clustering on the individual level. In column (1) to (4), the estimates are shown compared to the baseline. TreatPassive and TreatActive

[7]Similar to Bartling et al. (2015), table 2 and 3 report raw probit estimates, due to the interpretability of the interaction effects. All other tables with probit estimations report marginal effects.

Table 2: Main Treatment Effect

	Baseline vs. Passive		Baseline vs. Active		Active vs. Passive	
	(1)	(2)	(3)	(4)	(5)	(6)
TreatPassive	-0.730**	-0.715**			-0.869**	-0.717*
	(0.334)	(0.330)			(0.377)	(0.379)
TreatActive			0.154	0.022		
			(0.346)	(0.338)		
Period		-0.034		-0.033		-0.021***
		(0.022)		(0.022)		(0.005)
TreatPassiveXPeriod		-0.003				-0.016
		(0.026)				(0.014)
TreatActiveXPeriod				0.012		
				(0.022)		
Constant	0.273	0.645***	0.242	0.603***	0.417	0.650**
	(0.235)	(0.223)	(0.220)	(0.210)	(0.297)	(0.295)
Random Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	974	974	977	977	979	979
Unique N	48	48	48	48	48	48

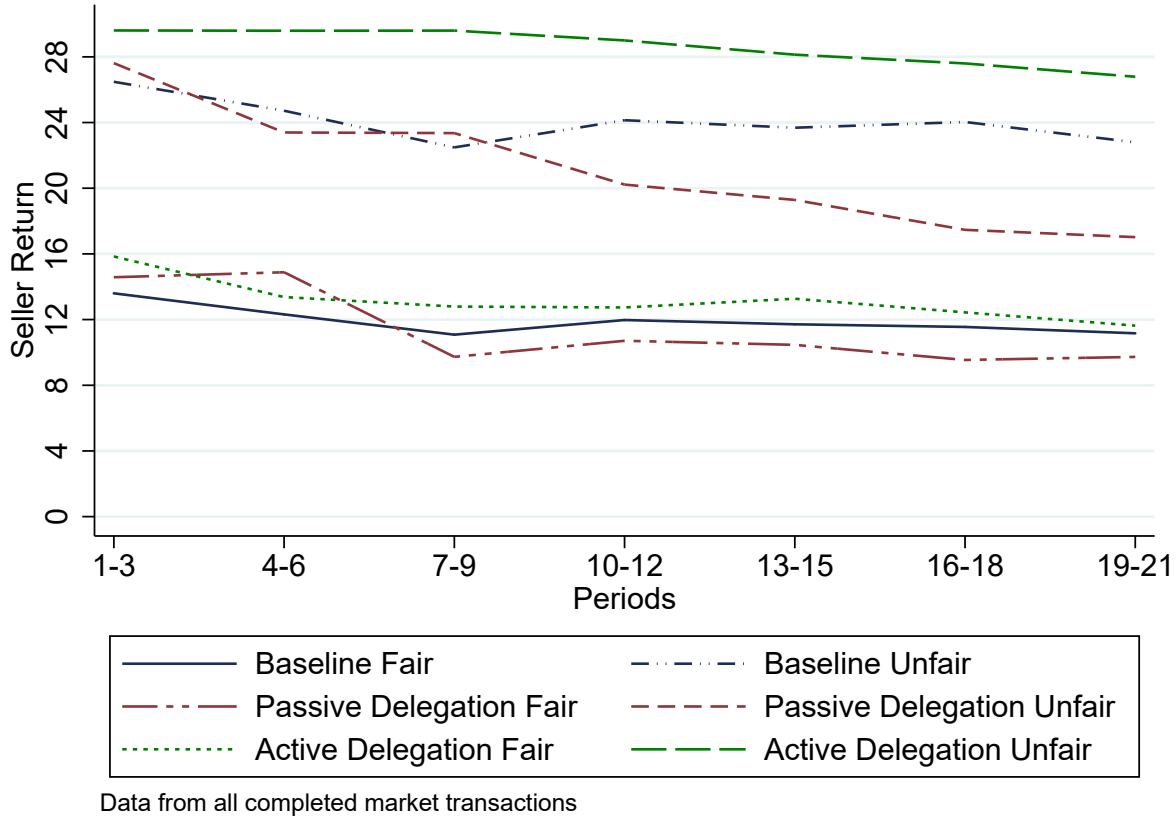
Note: Dependent variable is a dummy equal to one if the product type bought does not cause an externality and zero otherwise. In the first 4 columns, Baseline is the omitted category, in the column (5) and (6) Active is the omitted category. Cases in which no product was bought are omitted. TreatPassive is a dummy equal to one, if the observation is from the passive treatment, TreatActive if it is from the active treatment. Period is an integer taking on the value between 1 and 21. The table reports raw probit coefficients. All standard errors are clustered at the market level.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

represent the treatment dummy for the respective treatment. TreatPassive has a significantly negative effect. The marginal effect is 18.9%, which means that compared to the baseline the fair product market share is on average reduced by almost 19 percentage points in the passive treatment. In column (2), I include period, which is a count variable taking on values between 1 and 21. Looking at the period treatment interactions, the treatment effect is stable over time. The next two columns compare the active treatment to the baseline. TreatActive has no significant effect, which is also numerically small. The marginal effect of the treatment is 4.3%. The final two columns compare the two treatments to each other with active as the omitted category. According to column (5), the difference between the two treatments is comparable to the difference between the passive treatment and the baseline (23% marginal effect) and the effect is also stable over

time.

Figure 3: Seller Profit



Data from all completed market transactions

Next, I turn to the study of who carries the burden of the externality avoidance, by looking at the profit of the market interactions. Figure 3 shows the seller return evolution over all rounds. For comparability, total returns are displayed, which combine profit of owner and seller in the two delegation treatments. Strikingly, combined profit on the seller side for the fair product seems to be the same for each market condition as well as over time. In contrast, for the unfair product profits seem to differ graphically, but turn out to be insignificant in a regression. This also means, due to the treatment design, that seller (and owner) profit is halved in the delegation treatments compared to the baseline. Table 3 reports the corresponding results of a random effects regression on prices. Column (1) regresses the prices as a stable relation, column (2) includes period interactions as controls and column (3) includes treatment interactions with time. I do not find a significant effect in any of the specifications of the treatments on prices of the fair and unfair product. Apparently, the price difference is stable over time. Only the

overall price level is eroding faster for both types of products in the passive delegation treatment.

The results for overall surplus from trading in the market are summarized in figure 4. Each quadrant represents the average surplus from trade of a market side in the different treatments. For comparability, the profit of owners and sellers is displayed together under Seller Side Surplus for the two delegation treatments. Note that third parties can only lose from the trade between buyers and sellers, when they trade the unfair product. In the passive treatment, the share of unfair products being traded is significantly higher, which therefore leads to a lower surplus for third parties. This directly translates into a lower overall surplus as the surplus of seller and buyer together is equal to 50, irrespective of the price and product type. Note, the maximum overall surplus from the trade is 30 (50-20 vs 50-60). When looking at the surplus of buyers and sellers, the difference is pronounced between the treatments. Average buyer surplus increases significantly from baseline to passive delegation (p=0.093), while it decreases significantly in the active delegation in comparison to passive (p=0.003) and even to baseline (p=0.009). These results suggest that buyers may be a driver behind the different product type trading shares, which I investigate further in the following sections.

Overall, the results of the passive delegation treatment are in line with previous research on the effect of the presence of additional individuals and profit sharing on immoral behavior, leading to a lower share of the fair product. The findings in the active delegation treatment, in contrast, are in line with the idea that buyers potentially react negatively to delegation of immoral behavior. This reaction is however, only triggered by active delegation, i.e. the fact that owners have a choice over sellers and thereby over the products offered.

4.2 Responsibility Diffusion

In order to get a better understanding of which type of player is driving the effect, I investigate whether market findings are consistent with the idea of responsibility diffusion in passive delegation and the opposite in active delegation. The results from the

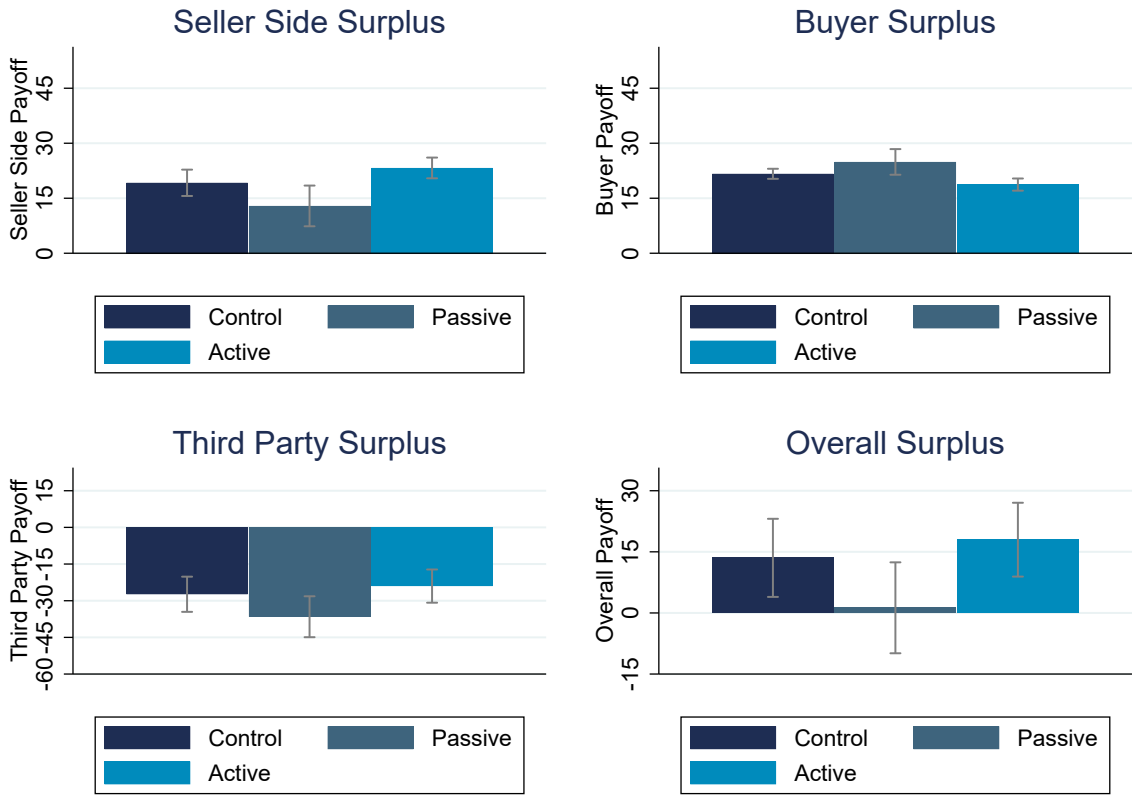
Table 3: Price Evolution

	Baeline vs. Passive		Baseline vs. Active		Active vs. Passive	
	(1)	(2)	(3)	(4)	(5)	(6)
TreatPassive	-2.348	2.575			-5.336	-0.898
	(4.030)	(3.332)			(4.461)	(4.233)
TreatPassiveXFair	0.445	-1.597			1.124	-0.985
	(2.743)	(2.245)			(2.341)	(2.266)
TreatPassiveXPeriod		-0.413***				-0.375**
		(0.101)				(0.166)
TreatPassiveXPeriodXFair		0.117				0.139
		(0.120)				(0.135)
TreatActive			3.170	3.641		
			(2.900)	(3.472)		
TreatActiveXFair			-1.041	-0.873		
			(2.654)	(2.503)		
TreatActiveXPeriod				-0.038		
				(0.163)		
TreatActiveXPeriodXFair				-0.024		
				(0.125)		
FairProduct	5.183***	6.752***	6.917***	6.814***	4.456***	6.098***
	(1.937)	(1.760)	(1.966)	(1.763)	(1.545)	(1.798)
Period	-0.380***	-0.143**	-0.160**	-0.142**	-0.405***	-0.180
	(0.087)	(0.067)	(0.079)	(0.068)	(0.101)	(0.148)
PeriodXFair	0.167**	0.053	0.037	0.052	0.160**	0.030
	(0.082)	(0.077)	(0.060)	(0.075)	(0.079)	(0.100)
Constant	28.751***	25.932***	26.121***	25.897***	32.056***	29.426***
	(1.763)	(1.644)	(1.743)	(1.648)	(2.604)	(3.091)
Random Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	974	974	977	977	979	979
Unique N	48	48	48	48	48	48

Note: The dependent variable is the product price of products sold, which can take the value between 0 and 50. In the first 4 columns, Baseline is the omitted category, in the column (5) and (6) Active is the omitted category. Cases in which no product was bought are omitted. TreatPassive is a dummy equal to one, if the observation is from the passive treatment, TreatActive if it is from the active treatment. FairProduct is a dummy equal to one if the product sold is a fair product. Period is an integer taking on the value between 1 and 21. The table reports the results of an OLS regression. All standard errors are clustered at the market level.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Figure 4: Treatment Effects on Surplus

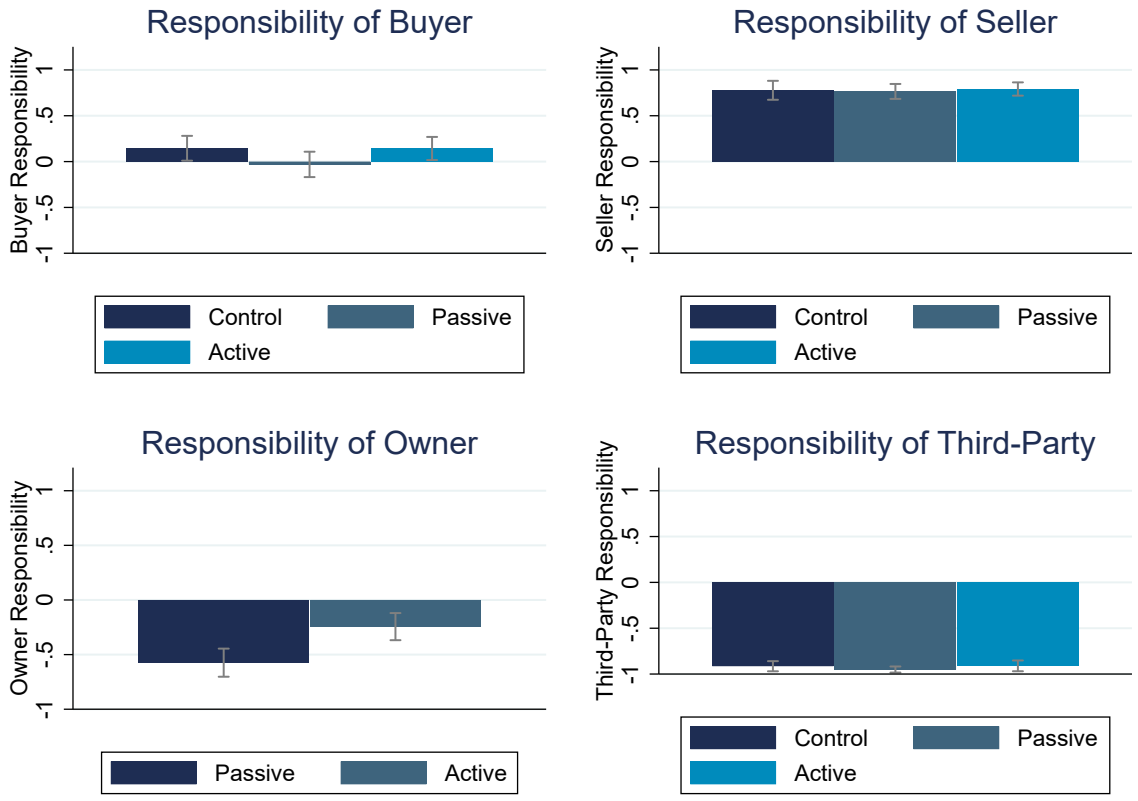


Note: The bars represent the average payoff points per participant over the 21 rounds without the endowment. Seller surplus includes the profit of owner and seller together in passive and active treatment. Overall surplus is the sum of buyer, third party and the sum of seller and owner profit within a treatment. Whiskers represent 95% confidence intervals.

responsibility norm elicitation on product type are displayed in figure 5.The answering options, ranging from "Not responsible at all" to "Very responsible", are normalized to the range from -1 to 1. Overall, sellers are always seen as the most responsible player type, with a mean of 0.78 in the baseline, 0.77 in the passive treatment and 0.79 in the active treatment. Buyers receive a lower mean with 0.15 in the baseline, -0.03 in the passive treatment and 0.14 in the active treatment. This means, that despite the market power of buyers, sellers are seen as the most decisive factor for which product type is traded in the end. This finding corresponds to findings from (Duch et al., 2015), who show that decision makers with proposal power are perceived to have the most influence on the outcome.

The most important question is whether there are treatment differences in responsibility attribution for sellers and buyers. For sellers the pairwise differences are far from

Figure 5: Responsibility Assignment



Note: The bars represent the average answers to the question 'In the following, we will ask you for your assessment on the question, who is responsible in the following market for the type of product that is ultimately being traded. This means, when a product is being traded, who is responsible which type of product (with or without loss to the third party) is traded?' Answers were: 'Not at all responsible', 'Rather not responsible', 'Rather responsible' to 'Very responsible'. The answer options are normalized to the range from -1 ('Not responsible at all') to 1 ('Very responsible'). Whiskers represent 95% confidence intervals.

being significant (p=0.8521, 0.824 and 0.636 for the baseline vs. passive, baseline vs. active and passive vs. active treatment respectively using a t-test, and p=0.614, 0.778 and 0.783 for a Wilcoxon rank-sum test). This means, that irrespective of whether sellers share the profit or if they are first chosen by owners before making a decision, they are seen as equally responsible in absolute terms by all subjects. Therefore, the predictions are refuted and the result is surprising given the findings of punishment deflection through delegation (Bartling and Fischbacher, 2012) as well as findings from cognitive science, showing that subjects are less morally responsible if others influence them (e.g. Phillips and Shaw, 2015). Apparently for participants, the offering decision seems not to be influenced by delegation, which is a novel finding.

Looking at buyers, the assigned responsibility norm differs. Comparing the baseline to the passive treatment, the difference in the t-test is marginally significant (p=0.084), but not for the Wilcoxon rank-sum test (p=0.123). The difference between passive and active is marginally significant in both tests (p=0.069 and 0.081). Between the baseline and the active treatment there is again no statistical difference (p=0.980 and 0.860).

At the same time the assignment of responsibility for owners reveals that they are seen as rather not responsible in the passive treatment with -0.57 and more responsible in the active treatment with -0.24, which differs significantly between the treatments (p=0.000 for both Wilcoxon rank-sum and t-test). Interestingly, this value is still far from the assignment of responsibility for sellers and buyers. This result indicates that, for example, shareholders may avoid responsibility simply by having a CEO taking the final decisions for them.

Overall, this pattern seems to be consistent with the interpretation from the burden analysis in the market. For buyers, responsibility (weakly) drops from the baseline to the passive treatment and increases significantly in the active treatment relative to the passive treatment. This increase in responsibility complements the rise in responsibility of the owners. Also, relative to sellers, buyers' relative responsibility for the exchange between the two increases, which may induce them to be willing to prefer buying a fair product either out of the feeling of responsibility or because they indirectly punish sellers offering unfair products. Consequently, this complementary consumer responsibility potentially limits the ability of owners to delegate immoral behavior.

Turning to the third party, the findings are in line with predictions. As expected, third parties are assigned almost no responsibility at all, with ratings of -0.91, -0.95 and -0.91 in responsibility in the baseline, passive and active treatment respectively.[8] Therefore, there is overwhelming evidence that, if a player type has no active role and may be subject to a loss, he is not seen as responsible at all. Further, the difference between owners and the third party in the passive treatment is highly statistically significant (p=0.000 for both Wilcoxon rank-sum and t-test). This means that subjects regard individuals who

[8]P-values for the t-test are p=0.247, 0.930 and 0.251 and for Wilcoxon rank-sum test 0.325, 0.604 and 0.630, for baseline vs. passive, baseline vs. active and passive vs. active respectively.

receive spoils from someone else's action to be more responsible than individuals, who do not or who even suffer from someone else's actions. This corresponds to recent findings in cognitive psychology, showing that peoples' moral judgment is not only influenced by control or causal responsibility (Engl, 2018) of an individual over the outcome, but also what they think subjects desire (Cushman, 2008; Guglielmo and Malle, 2010; Phillips and Shaw, 2015). Similarly in Bartling and Fischbacher (2012), individuals with no active role, who profit from immoral behavior, are punished more than individuals, who profit from moral behavior.

In further analyses, I test whether responsibility assignment of subjects for their own player type is different than for other player types (Epley and Gilovich, 2016; Gino and Pierce, 2010). Support for this hypothesis would mean, that individuals would bias their answer in a self-serving manner, either consciously or unconsciously. However, in unreported tests, I do not find any evidence to support this claim when comparing the responsibility assignment of players for their own type with other players' ratings for that type for any of the player types. The result supports that truthful answers emerge from the norm elicitation game through incentivization, making the responses overall meaningful.

A potential limitation for the explanatory power of the elicited norm is the potential for belief updating in the market setting. Subjects may later update their beliefs on responsibility, when accumulating experience in the market and learning what other subjects do. Nevertheless, the difference in the market share of the fair product does not change over time between treatments, such that the relative responsibility assignments should still be meaningful.

Additionally, one might criticize comparing the responsibility assignments between the treatments and baseline, because one more player type is available in the passive and active treatment. This may lead subjects to implicitly reduce responsibility assignments for all player types, because of some potential fixed overall responsibility amount. However, the question specifically avoided such formulation as to assign relative responsibility. Additionally, if this were the case, then the overall increase in aggregate responsibility ratings in the active treatment would not be possible.

4.3 Individual Type Behavior

The responsibility norm analysis reveals significant differences in the assigned responsibility only for buyers, but not for sellers, which also seems consistent with evidence from the burden sharing analysis. In this section, I provide further evidence in line with the proposed mechanism of complementary consumer responsibility.

Usually, data from prevailing product sales are difficult to interpret, because they are endogenously determined by supply and demand. Consequently, I cannot make a statement concerning whether buyers are the sole reason for the market outcome. However, in the experiment I perfectly observe the product menu available to buyers and can therefore estimate the choice of buyers conditionally on the available options between treatments and controlling for price differences. The order in which buyers saw the products was randomly drawn in each round, which leads to observations for each individual with a considerable range of choice sets. After controlling for prices, the treatment dummy should measure the treatment effect on willingness to buy a fair product, excluding the price and availability channel. Importantly, because the price differential does not significantly differ between the treatments, as shown above, this analysis should also be close to the overall effect the treatments have on the buyer's choice.

Table 4 reports the treatment effect on buyer behavior, again comparing the treatments pairwise. In column (1), (3) and (5) I measure the treatment effects and in the other columns I control for the lowest price a customer has faced from each product type and period effects. Only observations for which buyers saw at least one product of each type are considered. The number of observations does not vary significantly between the treatments, showing that the results are unlikely to be driven by a selection effect. When comparing the passive treatment to the baseline, buyers are significantly less likely to buy a fair product by marginally -14.6%, which may be driven by the interaction with sellers though. When controlling for the minimum prices buyers faced for both types of product, the effect of the higher minimum price of the fair product is negative, whereas the effect of a higher minimum price of an unfair product is positive. A price increase of the fair product of 1 point decreases the marginal buyer willingness to buy this product

Table 4: Buyer Treatment Effect

	Baseline vs. Passive		Baseline vs. Active		Active vs. Passive	
	(1)	(2)	(3)	(4)	(5)	(6)
TreatPassive	-0.146**	-0.173*			-0.215***	-0.225**
	(0.064)	(0.094)			(0.081)	(0.102)
TreatActive			0.072	0.031		
			(0.083)	(0.135)		
Period		-0.003		-0.004		-0.004
		(0.003)		(0.003)		(0.004)
MinPriceFairProd		-0.032***		-0.032***		-0.040***
		(0.008)		(0.010)		(0.008)
MinPriceUnfairProd		0.031***		0.029***		0.033***
		(0.007)		(0.007)		(0.005)
Random Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	723	723	705	705	706	706
Unique N	48	48	48	48	48	48

Note: Dependent variable is a dummy equal to one if the product type bought does not cause an externality and zero otherwise. In the first 4 columns, baseline is the omitted category, in the column (5) and (6) active treatment is the omitted category. Cases in which no product was bought are omitted. TreatPassive is a dummy equal to one, if the observation is from the passive treatment, TreatActive if it is from the active treatment. Period is an integer taking on the value between 1 and 21. MinPriceFairProd is the lowest price a buyer faces for a fair product, MinPriceUnfairProd for an unfair product. The table reports raw probit coefficients. All standard errors are clustered at the market level.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

by 3.2%. On the other hand, if the price of the unfair product increases by 1 point, the likelihood for a buyer to buy the fair product increases by 3.1%. The magnitude of the coefficients is not significantly different (Wald test, p=0.909). Both signs are consistent with standard market logic, where the two product types are at least to some degree substitutes. When controlling for the minimum prices, the treatment effect is numerically even larger at -17.3% and marginally significant. When comparing the baseline with the active treatment in column (3) and (4), the effect is positive but not significant, while the treatment effect of the passive treatment, using the active treatment as the reference category in column (5) and (6), is again significantly negative, with an effect of -22.5% in the specification controlling for prices.

Therefore, the overall pattern in buyer behavior is consistent with the idea of complementary consumer responsibility, showing that buyers drive the treatment effect to a

certain degree in line with their assigned responsibility. Buyers are significantly more likely to buy a fair product in the active treatment relative to the passive treatment and treatment effect sizes for buyers are comparable to the overall treatment effect.

For robustness, I check if price sensitivity to fair and unfair products differs between the treatments and do not find any significant effect. Moreover, buyers may see more or less fair or unfair product types in the different treatments, which may induce a framing effect representing a certain norm of appropriateness. When controlling for the number of fair products or the share of fair to unfair products seen additionally to prices, the results do not change.

Table 5: Seller Treatment Effect

	Baseline vs. Passive		Baseline vs. Active		Active vs. Passive	
	(1)	(2)	(3)	(4)	(5)	(6)
TreatPassive	-0.160**	-0.164***			-0.177**	-0.146**
	(0.065)	(0.060)			(0.078)	(0.068)
TreatActive			0.027	-0.005		
			(0.076)	(0.069)		
Period		-0.008**		-0.004		-0.006**
		(0.003)		(0.003)		(0.003)
ExpProfitPremium		0.008**		0.006***		0.012
		(0.003)		(0.002)		(0.009)
Random Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	1260	1025	1260	935	1260	920
Unique N	60	60	66	66	66	66

Note: Dependent variable is a dummy equal to one if the product type offered does not cause an externality and zero otherwise. In the first 4 columns, baseline is the omitted category, in the column (5) and (6) active treatment is the omitted category. TreatPassive is a dummy equal to one, if the observation is from the passive treatment, TreatActive if it is from the active treatment. Period is an integer taking on the value between 1 and 21. LastPeriodProfitPrem is the mean profit of fair products sold in the previous round minus the mean of unfair products sold in the previous round. The table reports raw probit coefficients. All standard errors are clustered at the market level.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

From these results on buyers alone, I cannot conclude that other player types are not affected by the treatments. Testing for influence from a seller perspective is more complicated than from a buyer perspective, as in contrast to picking from a menu of offers, sellers need to form expectations on which product will likely be sold for which price. However, the market outcome in all treatments does not correspond to standard predictions even

when taking into consideration inequality aversion. Under the assumptions for social preferences from above, seller profits should always be zero for each type of product, which is clearly not the case. Therefore, in order to find a clean treatment effect, conditional on endogenously determined profit premium and selling probability for each product type that correspond to the rationale of sellers is very difficult. One way to calculate a type of expected profit premium is to use the last period as an approximation and calculate the difference between the average profit for fair products and the average profit for unfair products excluding zero profits (for not sold products). In table 5, the product choice of sellers is regressed on the treatment dummies as well as the last period profit premium from offering the fair product. In column (1) and (2), the passive treatment is compared to the baseline, controlling for period and the last round profit premium in column (2). Consistent with the predictions, sellers are less willing to offer a fair product in the passive treatment, which is significant even after controlling for the different returns. Comparing the active treatment with the baseline, the effect is insignificant, similar to buyer behavior. When comparing the passive to the active treatment, the sellers are significantly less willing to offer the fair product in the passive treatment. An alternative way to test for seller treatment effects given the problematic adjustment for expected profits, is to look at the product decision after not having sold a product. Sellers, who have not sold a product in the previous period, should feel the most pressure to adapt their strategies by either reducing the price and/or changing the product type. If sellers are less willing to compete for the fair product in the passive treatment, then we should see sellers offer more unfair products after being left out than in the baseline. If, on the other hand, they have the same strategy of adaptation in all market settings, for example always change the product type when not sold, then this is an indication that they just try to find a product type and price combination to sell to the buyers irrespective of the treatment. Table 6 shows the results corresponding to the seller adaption strategy. The dependent variable is again a dummy equal to one, if a seller offers a fair product. Observations are limited to those participants, which have not sold the product in the previous round. There is no significant influence of the treatment dummies. The estimates are also numerically

smaller and standard errors comparable than in the previous specification, showing that the insignificance is not solely driven by the loss of observations. The results are robust to using expected seller profits including zero profit and using the maximum difference in payoffs or controlling for the probability of selling a fair product in the previous period.

Table 6: Seller Product Decision After Not Having Sold

	Baseline vs. Passive		Baseline vs. Active		Active vs. Passive	
	(1)	(2)	(3)	(4)	(5)	(6)
TreatPassive	-0.073	-0.054			-0.073	-0.093
	(0.073)	(0.068)			(0.073)	(0.064)
TreatActive			0.083	0.069		
			(0.107)	(0.092)		
Period		-0.008		-0.004		-0.003
		(0.006)		(0.006)		(0.006)
LastPeriodProfitPrem		0.001		0.001*		0.003***
		(0.001)		(0.001)		(0.001)
Random Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	273	273	219	219	273	218
Unique N						

Note: Dependent variable is a dummy equal to one if the product type offered does not cause an externality and zero otherwise. In the first 4 columns, baseline is the omitted category, in the column (5) and (6) active treatment is the omitted category. TreatPassive is a dummy equal to one, if the observation is from the passive treatment, TreatActive if it is from the active treatment. Period is an integer taking on the value between 1 and 21. LastPeriodProfitPrem is the mean profit of fair products sold in the previous round minus the mean of unfair products sold in the previous round. The table reports raw probit coefficients. The table only includes observations of subjects who have not sold a product in the previous round. All standard errors are clustered at the market level.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Overall, findings for sellers are mixed concerning treatment effects. Positive findings in the probability of offering an unfair product may be due to incorrect controlling for the expected profit premium and therefore driven by buyers. A null finding on the other hand would be in line with the findings from the responsibility norm evaluation, which only finds differences between treatments for buyers.

Given the findings that buyer responsibility increases with that of owners in the active setting, it is worth to ask how owners influence the market, i.e. what aspects they care for when selecting sellers. They choose sellers from a fixed menu of sellers who retain their ID over the course of the experiment, such that sellers can build up reputation

over time. After every round, owners see a summary of the sellers' behavior in the last round, informing them which product type was offered, if it was sold and for what price. This information is the only information for owners to condition their selection on and they may infer from previous profit and product choice, what the sellers will do in the next round. Indeed, the Spearman correlation coefficient between product type offered by sellers and the lagged product type is 0.46. This means, that if a seller offered a fair product in period t, the probability of him offering a fair product in t+1 is 76%, which makes the information useful as a selection criterion.

The seller choice by owners is displayed in table 7. In the first four columns, I estimate a conditional logit model, where the dependent variable is a dummy equal to one, if a seller is chosen, and zero otherwise for all periods from 2 to 21. The independent variables are FairProduct, which is equal to one, if the seller chose a fair product in the previous round and zero otherwise, Profit, which is the earnings an owner (and the seller) received in the last period and NotChosen, which is a dummy equal to one if a player was not chosen in the previous round. Consistent with prior findings in Hamman et al. (2010), owners choose sellers only based on profit, which is always significant, while offering a fair product is never positively significant. In column (5), I estimate the driver behind early choices. In contrast to analyzing all owners' choices, the first owners making a decision may focus on different things than later ones, because they might want to make a statement. In column (5), the dependent variable is a dummy equal to one, if the seller was among the two first sellers chosen in this round. Offering a fair product has no significant influence on being chosen early. Additionally, owners can rely on information on more rounds than just the last round through the constant seller ID. Note that this does not work for buyers, because the order of the products displayed was always randomly shuffled. Using the average share of the product decision of the last five rounds as well as the average profit, regressing this on the seller choice after period 4, I still find no positive significant effect of offering the fair product.

The results from the owner analysis are important. Without buyers stepping up and exerting their influence on the market by buying more fair products, the influence of

Table 7: Owner Choice

	Overall Choice				Early	Period > 4	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
FairProduct	0.123		-0.161	0.170	0.181	0.160	
	(0.177)		(0.203)	(0.225)	(0.307)	(0.248)	
Profit		0.007***	0.008***	0.087***	0.113***	0.091***	
		(0.002)	(0.002)	(0.018)	(0.021)	(0.020)	
NotChosen				8.791***	10.585***	9.223***	
				(1.962)	(2.235)	(2.237)	
MeanFairLast5							-0.438
							(0.316)
MeanProfitLast5							0.171***
							(0.028)
N	2400	2400	2400	2400	2400	1920	1920

Note: The model is estimated using a conditional logit. FairProduct is a dummy equal to one if a seller has offered a fair product in the previous round. Profit equals the profit a seller made in the previous round. NotChosen is a dummy equal to one if a manager was not chosen in the previous round. MeanFairLast5 is the mean share of fair products offered in the last 5 rounds from all products being offered and MeanProfitLast5 the mean profit from the last 5 rounds (with zero profit if a product was not sold or not offered). The dependent variable in all columns except column (5) is a dummy equal to one for the seller which is selected by a specific owner in a round and zero for all others. In column (5) the dependent variable is a dummy equal to one, if a seller was selected as first or second overall choice. The first three columns contain only observations for sellers selected in the previous rounds for all periods between 2 and 21, in column(4) they include the managers who have not been previously chosen. Therefore, the variables FairProduct and Profit are interacted with a dummy equal to one if the manager was chosen in the previous round and zero otherwise. Columns (6) and (7) are restricted to observations after the fourth period. The table reports raw probit coefficients. All standard errors are clustered at the market level.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

owners through the selection only based on profit would go in the opposite direction, because the profits from selling an unfair product are higher. The signal to sellers, who also observe the selection process while owners are making their choices, is clear. Profit seems to be the only thing that matters and there is no bonus from offering a fair product for selection. Delegation in this way would facilitate immoral behavior. Only because buyers are more willing to buy the fair product, does the market share of the fair product increase, which makes the implications of this finding more important.

4.4 Robustness

As robustness checks, I primarily test whether participants understood the experiment or are enforcing other norms. One potential criticism could be that the active treatment is overall more complicated than either the baseline or the passive treatment, due to the additional stage of seller selection by owners. If subjects do have less understanding of the interactions, this could potentially lead to behavioral changes unrelated to the treatment.

I check the understanding of the game using several control questions at the end of the instructions. When subjects were in the experimental room, they were verbally reminded that if they could not solve or have difficulties answering these control questions, they could ask the experimenter, questions were then answered individually at the place of the subject. Additionally, the answers to the control questions were gathered after the experiment.Overall, 78% of all subjects answered all control questions correctly. The difference between the treatments is not significant.[9]

Further, I looked at choices, which seem suboptimal, where subjects chose a product for which a cheaper product of the same type was available. This behavior could indicate an alternative norm buyers would try to enforce by rewarding a certain type of offer. Overall, this happened in 67 cases out of 1,465 trades (4.6% of the times). If this behavior would drive the result, it should vary substantially between treatments. Using a random effect regression as before, I do not find significant differences.

Another potential limitation of my findings may be that markets do not have converged completely. However, with 21 periods and relatively slow movement and constant differences between the treatments, it seems unlikely that they would converge completely, given more time.

Finally, I checked for spiteful buying decisions. Buying an unfair product for a higher price than a fair product would involve spiteful behavior towards a third party, while at the same time rewarding a seller (because he gains much more than the seller who offers the fair product for a similar price). There was only one case in all trades in the experiment, which can be characterized as irrational or spiteful trade and it happened in the control

[9]P-values are p=0.818 for baseline vs. passive, p=0.660 for baseline vs. active and p=0.442 for active vs. passive.

treatment. Overall, it seems unlikely that the results are driven by misunderstanding of the experiment or different norm enforcement rather than by the intended treatment itself.

5 CONCLUSION

In this study, I test whether moral eroding effects of delegation persist even in market settings where consumers may respond to this behavior. The results from the market reveal that passive delegation, where sellers share their profit with an owner, leads to a significantly lower share of the fair product being traded consistent with the findings on responsibility diffusion in the delegation literature. In contrast the market share of the fair product in the active delegation, where owners first selected sellers, is significantly higher than in the passive delegation treatment and insignificantly different from the baseline.

Answers from a responsibility norm elicitation reveal that buyers' responsibility differs between the treatments and that, when owners have an active choice over sellers, they are assigned a kind of complementary consumer responsibility. The results from analyzing the different market participants' behavior are in line with the findings from the responsibility norm analysis and reveal strong support for the moral eroding effect of passive delegation. Buyers are less willing to buy the fair product, when the sellers share the profit with a passive owner. Findings for sellers are not robust to different specifications. In contrast, buyers are less willing to buy the unfair product in the active delegation treatment, effectively limiting the scope for delegating immoral decisions found in other research. A promising direction for future research would be to investigate in more detail why the owners' selection process makes buyers feel more responsible, inducing them to buy more responsibly.

Experimental tests on socially responsible behavior in markets are an important source of advice for real world interaction. Therefore, my findings suggest that consumers may potentially limit the power of delegation on the proliferation of immoral actions in markets through delegation, but that the effect depends on the structure of delegation. Apparently, consumers need to be provided with a strong enough focus on owners' choice, for

complementary consumer responsibility to emerge. Therefore, initiatives enhancing transparency on the producer side potential to delegate to different entities may be beneficial for limiting immoral firm behavior.

References

Appelbaum, E. and Batt, R. (2014). *Private Equity at work: When wall street manages main street*. Russel Sage Foundation.

Bartling, B. and Fischbacher, U. (2012). Shifting the blame: On delegation and responsibility. *Review of Economic Studies*, 79(1):67–87.

Bartling, B., Grieder, M., and Zehnder, C. (2017a). Competitive pricing reduces wasteful counterproductive behaviors. *Journal of Public Economics*, 156:34–47.

Bartling, B. and Özdemir, Y. (2017). The Limits to Moral Erosion in Markets: Social Norms and the Replacement Excuse. *ECON - Working Papers Universität Zürich*, No. 263.

Bartling, B., Valero, V., and Weber, R. (2017b). On the Scope of Externalities in Experimental Markets. *Experimental Economics*.

Bartling, B., Weber, R. A., and Yao, L. (2015). Do Markets erode social responsibility? *Quarterly Journal of Economics*, 130(1):219–266.

Bauer, T., Epstein, G. S., and Gang, I. N. (2005). Enclaves, language, and the location choice of migrants. *Journal of Population Economics*, 18(4):649–662.

Behnk, S., Hao, L., and Reuben, E. (2017). Partners in Crime: Diffusion of Responsibility in Antisocial Behaviors. *IZA Discussion Paper Series*, No. 11031.

Bénabou, R. and Tirole, J. (2006). Incentives and Prosocial Behavior - Slides. *American Economic Review*, 96(5):1652–1678.

Benabou, R. and Tirole, J. (2010). Individual and corporate social responsibility. *Economica*, 77(305):1–19.

Berenson, A. (2006). A cancer drug's big price rise is cause for concern. *New York Times*, pages 25, March 12.

Berle, A. A. and Means, G. C. (1932). *The Modern Corporation and Private Property*. Transaction Publishers.

Bolton, G. E. and Ockenfels, A. (2000). A Theory of Equity, Reciprocity, and Competition. *American Economic Review*, 90(1):166–193.

Brodzinky, S. (2013). Central American Farmers Stay One Step Ahead of Profit-hungry 'Coyotes'. *The Guardian, available online: https://www.theguardian.com/global-development/2013/jun/25/central-american-farmers-coyotes*.

Cappelen, A. W., Nielsen, U. H., Sørensen, E., Tungodden, B., and Tyran, J. R. (2013). Give and take in dictator games. *Economics Letters*, 118(2):280–283.

Carpenter, J. and Dolifka, D. (2017). Exploitation aversion: When financial incentives fail to motivate agents. *Journal of Economic Psychology*, 61:213–224.

Charness, G. and Rabin, M. (2002). Understanding Social Preferences with Simple Tests Author. *The Quarterly Journal of Economics*, 117(3):817–869.

Chen, D. L., Schonger, M., and Wickens, C. (2016). oTree-An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97.

Clean Clothing Campaign (2017). Labour Without Liberty. Female Migrant Workers in Bangalore's Garment Industry. Technical report.

Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, 108(2):353–380.

Dana, J., Weber, R. A., and Kuang, J. X. (2007). Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness. *Economic Theory*, 33(1):67–80.

Darley, J. M. and Latane, B. (1968). Bystander Intervention in Emergencies: Diffusion of Responsibility. *Journal of Personality and Social Psychology*, 8(4.1):377–383.

Duch, R., Przepiorka, W., and Stevenson, R. (2015). Responsibility Attribution for Collective Decision Makers. *American Journal of Political Science*, 59(2):372–389.

Engelmann, D., Friedrichsen, J., and Kübler, D. (2018). Fairness in Markets and Market Experiments. *WZB Discussion Paper SP*, (No. II 2018-203).

Engl, F. (2018). A Theory of Causal Responsibility Attribution. *SSRN*, No.2932769.

Epley, N. and Gilovich, T. (2016). The Mechanics of Motivated Reasoning. *Journal of Economic Perspectives*, 30(3):133–140.

Falk, A. and Szech, N. (2014). Diffusion of Being Pivotal and Immoral Outcomes. *DIW Berlin*.

Fehr, E. and Gächter, S. (2000). Fairness and Retaliation: The Economics of Reciprocity. *Journal of Economic Perspectives*, 14(3):159–181.

Fehr, E. and Schmidt, K. (1999). A Theory of Fairness, Competition and cooperation. *The Quarterly Journal of Economics*, 114(3):817–868.

Fershtman, C. and Gneezy, U. (2001). Strategic Delegation : An Experiment. *The RAND Journal of Economics*, 32(2):352–368.

Finlay, W. and Coverdill, J. E. (2007). *Headhunters: Matchmaking in the labor market*. Cornell University Press.

Fiske, S. T. (2013). *Social beings: Core motives in social psychology*. John Wiley & Sons.

Friedrichsen, J. (2017). Is Socially Responsible Production a Normal Good? *DIW Berlin Discussion Paper Series*, No.1644.

Gino, F., Norton, M. I., and Weber, R. A. (2016). Motivated Bayesians: Feeling Moral While Acting Egoistically. *Journal of Economic Perspectives*, 30(3):189–212.

Gino, F. and Pierce, L. (2010). Lying to Level the Playing Field: Why People May Dishonestly Help or Hurt Others to Create Equity. *Journal of Business Ethics*, 95(SUPPL. 1):89–103.

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association*, 1(1):114–125.

Guglielmo, S. and Malle, B. F. (2010). Enough skill to kill: Intentionality judgments and the moral valence of action. *Cognition*, 117(2):139–150.

Hamman, J. R., Loewenstein, G., and Weber, R. A. (2010). Self-Interest through Delegation: An Additional Rationale for the Principal-Agent Relationship. *American Economic Review*, 100(4):1826–1846.

Harrison, B. A. and Scorse, J. (2010). Multinationals and Anti-Sweatshop Activism. *American Economic Review*, 100(1):247–273.

Herkenhoff, P. and Krautheim, S. (2018). The International Organization of Production in the Regulatory Void. *CESifo Working Paper Series*, (No. 6922).

Hill, A. (2015). Does delegation undermine accountability? Experimental evidence on the relationship between blame shifting and control. *Journal of Empirical Legal Studies*, 12(2):311–339.

Jakob, M., Kübler, D., Steckel, J. C., and van Veldhuizen, R. (2017). Clean up your own mess: An experimental study of moral responsibility and efficiency. *Journal of Public Economics*, 155(September):138–146.

Jensen, M. C. and Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4):305–360.

Kirchler, M., Huber, J., Stefan, M., and Sutter, M. (2015). Market Design and Moral Behavior. *Management Science*, 62(9):2615–2625.

Krakovsky, M. (2015). *The Middleman Economy - How Brokers, Agents, Dealers, and Everyday Matchmakers Create Value and Profit*. Palgrave MacMillan, New York.

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

List, J. A. (2007). On the Interpretation of Giving in Dictator Games. *Journal of Political Economy*, 115(3):482–493.

Moskin, J. (2004). Helping the Third World One Banana at a Time. *New York Times*, pages 1, May 5.

Oexl, R. and Grossman, Z. J. (2013). Shifting the blame to a powerless intermediary. *Experimental Economics*, 16(3):306–312.

Paharia, N., Coffman, L. C., and Bazerman, M. H. (2012). Intermediation and Diffusion of Responsibility in Negotiation: A Case of Bounded Rationality. In Bolton, G. E. and Croson, R. T., editors, *Oxford Handbook of Economic Conflict Resolution*, pages 37–46. Oxford University Press, New York.

Paharia, N., Kassam, K. S., Greene, J. D., and Bazerman, M. H. (2009). Dirty work, clean hands: The moral psychology of indirect agency. *Organizational Behavior and Human Decision Processes*, 109(2):134–141.

Parrino, R., Sias, R. W., and Starks, L. T. (2003). Voting with their feet: Institutional ownership changes around forced CEO turnover. *Journal of Financial Economics*, 68(1):3–46.

Phillips, J. and Shaw, A. (2015). Manipulating Morality: Third-Party Intentions Alter Moral Judgments by Changing Causal Reasoning. *Cognitive Science*, 39(6):1320–1347.

Pigors, M. and Rockenbach, B. (2016a). Consumer Social Responsibility. *Management Science*, 62(11):3123–3137.

Pigors, M. and Rockenbach, B. (2016b). The competitive advantage of honesty. *European Economic Review*, 89:407–424.

Spulber, D. F. (1999). *Market microstructure: intermediaries and the theory of the firm*. Cambridge University Press.

Stigler, G. J. and Friedland, C. (1983). The Literature of Economics: The Case of Berle and Means. *The Journal of Law and Economics*, 26(2):237–268.

Taub, A. (2016). Meet VelvetJobs: Transforming The $5 Billion Career Transition Industry. *Forbes (Available online: https://www.forbes.com/sites/alextaub/2016/09/13/meet-velvetjobs-transforming-the-5-billion-career-transition-industry/#54da962db42e Accessed August 15, 2016).*

Tirole, J. (2002). Rational irrationality: Some economics of self-management. *European Economic Review*, 46(4):633–655.

Wiltermuth, S. S. (2011). Cheating more when the spoils are split. *Organizational Behavior and Human Decision Processes*, 115(2):157–168.

6 APPENDIX A - THEORETICAL APPENDIX

Theoretical Equilibrium

In the baseline, nothing changes compared to the original in Bartling et al. (2015), which is based on the inequity aversion model of Fehr and Schmidt (1999). The baseline assumptions are that since sellers and consumers are both free to choose what to trade and at which price, the social responsibility concept only applies to the third parties payoff and not to other market participants. The assumptions seem reasonable, as the models fairly well describes the outcome in their paper.

Following their analysis, I derive the price premium for the fair product i) the buyer is willing to pay ii) the producer needs to be willing to offer the fair product and finally show iii) the equilibrium predictions.

(i) **Buyer willingness**

The corresponding utilities for buyers to buy a fair or unfair product are:

$$U^f = v - p^f - \beta \cdot max(v - p^f; 0)$$
$$U^u = v - p^u - \beta \cdot max(e + v - p^u; 0)$$

U^f is the utility of a buyer for a fair product, U^u for an unfair product, v is the valuation for the product and $p^{f,u}$ the respective price. The parameter β is the aheadness aversion parameter and e is the value of the externality to the third party. Buyers will buy a product type j only if $v \geq p^j$. This leads to a corresponding price premium, below which the buyer is willing to buy a fair product:

$$\Delta p = p^f - p^u \leq \frac{\beta}{(1-\beta)} e$$

For completely selfish buyers, the fair product will only be bought if the price is equal or below the unfair product price. If a buyer exhibits other-regarding preferences in

the form of inequality aversion with $\beta > 0$, then buyers may be willing to buy a fair product for a price premium. The result remains unchanged in the two treatments.

(ii) **Seller price premium**

In the Baseline:

$$\Pi^f = p^f - c - \beta \cdot max(p^f - c; 0)$$
$$\Pi^u = p^u - \beta \cdot max(e + p^u; 0)$$

Π is the profit of a seller, and c is the cost for the fair product. Sellers sell if $\Pi^f \geq \Pi^u$. Given that a seller should never offer a product with $p^f \leq c$ or $p^u \leq 0$, the solution to the problem is

$$\Delta p = p^f - p^u \geq c - \frac{\beta}{(1-\beta)} e$$

The trade-off is intuitively clear, the higher the cost c, the larger the price differential between fair and unfair price has to be in order for the seller to offer the fair product. If the seller is completely selfish ($\beta = 0$), then the price differential needs to be exactly the cost. Contrarily, if the seller cares about the third party, the externality e reduces the necessary price differential. If $\beta/(1-\beta)e \geq c$, then is it is optimal to sell the fair product, even if both products are sold for the same price. For the parameters of $c = 20$ and $e = 60$, this is the case for $\beta \geq 1/4$.

In the active and passive treatment:

In both treatments the incentives are the same, assuming that in the active treatment, the sellers and buyers do not consider the payoff of the seller left out of the trade, which seems quite reasonable.

$$\Pi^f = \frac{1}{2}(p^f - c) - \beta \cdot max\left(\frac{1}{2}(p^f - c); 0\right)$$
$$\Pi^u = \frac{1}{2}p^u - \beta \cdot max\left(e + \frac{1}{2}p^u; 0\right)$$

this leads to:

$$\Delta p = p^f - p^u \geq c - \frac{\beta}{(1-\beta)} 2e \tag{11}$$

Again, intuitively this results makes sense. The inequality aversion induced negative utility from the externality is fully incorporated, while at the same time the return is only halved. For a selfish seller, the incentive change does not influence the decision for the product type. For non-selfish sellers, the necessary β for the willingness to offer a fair product however is reduced.

(iii) **Market Equilibrium**

Conditions for an equilibrium

Following the arguments from Bartling et al. (2015), the baseline is unchanged, allowing me to calculate the potential equilibrium conditions in the same way. Assuming homogeneous agents, if $\beta \geq c/(2e+c)$, then buyers are willing to pay at least as much for the fair product as the seller demands as compensation for offering it. It follows that there exists an equilibrium with a positive fair product price premium where the fair product is indeed sold. Given the parameters chosen in this study, this is the case if $\beta \geq 1/7$.

Theoretical equilibrium:

As in the original paper in all treatments the number of sellers is larger than the number of buyers. This means that through price competition, profits for sellers should be zero and prices should correspond to the marginal costs, in this case $p^f = c$ and $p^u = 0$. In this equilibrium, sellers do not care which product they offer as their profit is zero both if they sell and if they do not sell. Buyers will at the prevailing prices only buy the unfair product, which gives them the highest utility. This leads to a market share of the unfair product of exactly 100% in all treatments.

Heterogenous Agent equilibrium:

The equilibrium in the market depends on the assumption of the distribution over β. Consistent with the original paper which uses the distribution found in the paper

of Fehr and Schmidt (1999), 30% of the subjects (and by randomization buyers and sellers) are characterized by $\beta = 0$, 30% have a $\beta = 0.25$ and 40% have a $\beta = 0.60$. Using this preference distribution as well as the assumption of perfect divisibility in order to ease computation, I derive the equilibrium predictions for the three different treatments.

Baseline

The Equilibrium is characterized by a market share of $[0.4, 0.7]$ of the fair product and $p^f = 20$ and $p^u = 0$. The selfish sellers offer the unfair product at a price of zero, the selfish consumers buy the unfair product at zero. The buyers with $\beta = 0.25$ are indifferent between buying a fair or an unfair product, therefore these 30% can buy each product. The buyers with $\beta = 0.6$ strictly prefer the fair product. The sellers with $\beta = 0.6$ will offer the fair product, the other sellers are indifferent between the product types. For each preference type there are more sellers than buyers, all buyers can buy the product they prefer and some sellers can't sell their product.

To see that this is an equilibrium, consider deviations from it. Consider first consumers. The 30% buyers have no incentive to switch to a fair product, because it would mean a decrease in utility for them. The buyers with $\beta = 0.6$ would suffer a utility loss from switching to the unfair product and for the buyer with $\beta = 0.25$ any product is okay. Overall deviations from the equilibrium range on the buyer side are not utility increasing.

For the sellers, the argument is similar. Sellers with $\beta = 0.6$ would incur a utility loss if they sold the unfair product at price zero, such that they do not have an incentive to switch with the product decision. Decreasing the price would lead to a personal loss and an increased price would not sell the product. For this group of sellers, a deviation is not maximizing. Sellers with $\beta = 0.25$ are indifferent between offering the different product types and switching would again not increase profits,

nor would changing the price. The same holds for sellers with $\beta = 0$.

Passive and Active Treatment

Given the zero profit conditions, the equilibrium in the treatments is characterized by a market share of 70% of the fair product and again $p^f = 20$ and $p^u = 0$. The only difference is for sellers with $\beta = 0.25$. They are not willing to offer the unfair product at $p^u = 0$ anymore. Therefore, they strictly offer the fair product at marginal cost. If they were to offer the fair product at a higher price, buyers would not buy it anymore. Comparing the equilibria in the two treatments with the equilibrium in the baseline, the market share of the fair product, θ^i should have the following relationship:

$$\theta^{Baseline} \leq \theta^{Passive} = \theta^{Active}.$$