



# **HCEO WORKING PAPER SERIES**

Working Paper



HUMAN CAPITAL AND  
ECONOMIC OPPORTUNITY  
GLOBAL WORKING GROUP

The University of Chicago  
1126 E. 59th Street Box 107  
Chicago IL 60637

[www.hceconomics.org](http://www.hceconomics.org)

**A Little Good is Good Enough:**  
**Ethical Consumption, Cheap Excuses, and Moral Self-Licensing**

**Jannis Engel <sup>a</sup>, Nora Szech <sup>b</sup>**

**March 2017**

We explore the role of cheap excuses in product choice. If a product improves upon one ethically relevant dimension, agents may care less about other, completely independent ethical facets of the product. This ‘static moral self-licensing’ would extend the logic of the well studied moral self-licensing over time. Our data document that static moral self-licensing exists. Furthermore, effects spill over to later, unrelated but ethically relevant contexts. Thus, static moral self-licensing and moral self-licensing over time amplify each other. Outsiders, though incentivized for correct estimates, are completely oblivious to effects of moral self-licensing, both, static and over time.

JEL-Codes: D03, D84

Keywords: moral self-licensing, moral spillovers, cheap excuses, outsider beliefs, moral personality, homo moralis, preference module

<sup>a</sup> Department of Economics, Karlsruhe Institute of Technology (KIT), jannis.engel@kit.edu

<sup>b</sup> Department of Economics, Karlsruhe Institute of Technology (KIT), Berlin Social Science Center (WZB), and CESifo, nora.szech@kit.edu

Corresponding authors: Jannis Engel, jannis.engel@kit.edu, Nora Szech, nora.szech@kit.edu

## 1. Introduction

Green, environmentally sustainable products are on the rise in a variety of areas, such as food, fashion, and electronics.<sup>1</sup> Many customers claim that they care about ethical dimensions of products, and that they are willing to pay extra in order to support ethical standards of the items they buy.<sup>2</sup> Some brands, such as Wholefoods, aim at a universal approach. Yet many companies seem to focus on one specific ethical facet of their products instead, and remain rather silent about others. For example, fashion-producer H&M advertises to replace conventional cotton in its “conscious collection,”<sup>3</sup> and to use organic or recycled cotton instead. Regarding the safety of workers in production, a comparably visible program does not exist. Likewise, Apple has publicly announced to audit all its suppliers with regard to the use of conflict minerals,<sup>4</sup> but so far has not come up with ensuring the safety of workers in the production processes as well. This is specifically remarkable as both companies have been involved in scandals regarding the safety of workers in production.<sup>5</sup>

Of course, for a company, it may be easier to focus on one ethically relevant dimension,<sup>6</sup> e.g., limiting environmental damage from growing cotton for garments, instead of improving upon all ethically relevant aspects of its products at the same time. It may also be the case that some facets such as environmental impact are easier to address than other facets such as labor standards of workers involved. This may partly explain why many companies stress specific ethical aspects in order to gain a more ethical reputation. Yet another reason for such practices could be that they serve customers’ moral interests very well by doing so. Maybe, for customers, fulfilling one ethical aspect is “enough” to ease their moral conscience when buying a product. Potentially, for many product decisions, “a little good is good enough.” If so, the need to improve other ethical facets may become quite irrelevant for companies as soon as they address one aspect.

---

<sup>1</sup> Compare overviews on different industries such as organic agriculture (FiBL & IFOAM - Organics International 2016, US Department of Agriculture 2016) and electronics (Consumer Technology Association 2015, UL 2014).

<sup>2</sup> This is the famous 30:3 syndrome. It often emerges when comparing stated intent in questionnaires and actual purchase behavior when it comes to ethical consumption; compare e.g., Cowe and Williams 2000.

<sup>3</sup> H&M’s announced goal is to obtain 100 percent of their cotton from sustainable sources. H&M defines organic cotton, recycled cotton, and Better Cotton (BCI) as sustainable. Better Cotton Initiative (BCI) is a non-profit organization that aims to reduce the environmental impact of cotton production. Compare e.g., the sustainability report by H&M (2015).

<sup>4</sup> The metals tantalum, tin, tungsten, and gold are essential for many electronic devices such as laptops, smartphones, and video cameras. They often serve as sources of income for armed groups in conflict zones such as the Democratic Republic of the Congo (DRC). The 2010 Dodd-Frank-Act requires U.S.-listed companies to investigate if their products contain at least one of the four conflict minerals that stems from the DRC or one of its neighboring countries (Apple 2016).

<sup>5</sup> Compare Burke (2013), Duhigg and Barboza (2012).

<sup>6</sup> We follow Bandura (2016) and define moral and ethical behavior in avoiding harm to other people and/or the environment. A definition independent of culture and time is notoriously difficult. Yet avoiding harm is central in many definitions of morality, compare also Gert (2012): “In this descriptive sense, although avoiding and preventing harm is common to all, ‘morality’ can refer to codes of conduct of different societies with widely differing content, and still be used unambiguously.”

This experiment explores whether subjects indeed care comparatively less about another ethical, unrelated dimension as soon as one ethical aspect of a product appears to be fulfilled. To this purpose, we elicit subjects' willingness to pay for improved manufacturing standards of textile products. We choose textile products, more specifically towels, for the following three reasons. First, the fashion sector is a \$1.75 – 3 Trillion (Joint Economic Committee 2015) market that employs over 57 million people worldwide (FashionUnited 2016). It is thus an important industrial sector. Second, the bulk of the production takes place in developing and newly industrialized countries.<sup>7</sup> The textile industry has often been in the focus when it comes to the need for improved production standards. Third, we specifically choose a product that most subjects may find useful, independent of personal characteristics such as gender or taste in fashion. Therefore, we employ towels in neutral colors as concrete products to decide upon.

Subjects know from the instructions that manufacturing (i.e., sewing the towel) is a production step unrelated to generating the raw material, in our case cotton, and that the two production steps often even take place in different countries. Thus subjects are informed that whether a towel is made from conventional versus organic cotton has nothing to do with the manufacturing standards in the sewing step. Nevertheless, our lab-in-the-field<sup>8</sup> data show that willingness to pay for secure working standards in the sewing process is highly significantly smaller if subjects know that they decide between towels that are made from organic cotton instead of conventional cotton. Ensuring better working standards becomes highly significantly more important to subjects when they decide about towels made from conventional cotton instead.

Our data demonstrate that organic cotton seems to serve as a kind of moral excuse for ignoring the working conditions in the sewing sector. The data thereby show that a static version of the classic moral self-licensing over time exists,<sup>9</sup> i.e., if one facet seems ethically fine, another unrelated ethical facet becomes significantly less relevant to subjects. The reason behind could be that subjects consider all these ethically relevant aspects as loading on one important factor, such as a positive self-image. This happens even though subjects know there is no correlation regarding the consequences of the different ethically relevant aspects. Obviously, workers in manufacturing do not work in more secure environments or earn more money just because the cotton was organic or recycled instead of conventional.

In a next step, we explore whether fulfilling one ethical dimension when buying a product has also spillovers on other, unrelated ethically relevant contexts later in time. Both effects, the well-known

---

<sup>7</sup> China, India, Vietnam, and Bangladesh are among the countries with the greatest number of employees in the textile and garment sector (International Labour Organization 2014, 2016).

<sup>8</sup> In line with Levitt and List (2007), we employ a real item in order to increase outside validity.

<sup>9</sup> For studies documenting moral self-licensing over time, see Monin and Miller (2011), Gneezy, Imas, and Madarász (2014) and other work as described in the literature overview.

moral self-licensing over time as well as the new moral self-licensing over different ethical facets at the same time, would then amplify each other. This is exactly what we find. In our study, about half an hour after the main part of the experiment (which focuses on the decision regarding towels), subjects are offered the opportunity to share money with refugees from a local refugee camp. Subjects do not know in advance that they will be confronted with this sharing decision. If fulfilling one ethical criterion when opting for a towel lowers interest in fulfilling another one, it may also be the case that willingness to share money with people in need becomes morally less relevant. This is exactly what we find. Thus, fulfilling one ethical criterion when deciding upon a product does not only lower interest in fulfilling another, orthogonal criterion of ethical production at the same point in time – it also reduces moral behavior in a later, completely unrelated context. Potentially, we find these spillovers because for subjects all these ethical behaviors load on one and the same personality factor, which would be “being an ethical person.”<sup>10</sup>

Knowing about the effects of moral self-licensing when it comes to product decisions could thus be an important first step in order to discuss social implications. Therefore, we incentivize other, unrelated subjects to predict the behavior of subjects in our product choice treatments. Our data show that these unrelated subjects are completely unaware of effects of static moral self-licensing! They also underestimate spillover effects on other morally relevant, later contexts, i.e., the well-known moral self-licensing over time. Thus, moral self-licensing does not appear to be intuitive to outsiders. Potentially, outsiders follow a different ethical guideline when not in the situation themselves and see the different product facets and contexts over time as they are: unrelated in their ethical consequences. Outsiders fail to anticipate that subjects, when weighing self-interest against something morally relevant, may prefer to find some kind of moral excuse for self-oriented behavior. Being moral at one point, regarding one relevant product facet, seems to be enough for that. Our findings thus indicate that exploring moral self-licensing in its different facets, but also spreading awareness of such effects, could be of major importance for social and political debate.

In addition to decision-making across contexts, we study the interactions between personality, ethical consumption, and moral self-licensing. To this purpose, we employ two well-established personality measures, the Big Five Inventory (BFI) (Rammstedt and John 2005) and the Mach-IV test for Machiavellianism (Christie and Geis 1970). In addition, we use the new Preference Survey Module (Falk et al. 2016),<sup>11</sup> which has been rather similarly implemented in the Gallup World Poll 2012. The latter is specifically interesting as it has been calibrated with incentivized laboratory experiments, and therefore, though not incentivized, may have some particular outside validity compared to other non-incentivized surveys. Our data support this view. In the well-established questionnaires as well as in the Preference Module, we find positive correlations between willingness to pay for secure

---

<sup>10</sup> See also Loewenstein (1999), Bénabou and Tirole (2002), Falk and Szech (2016).

<sup>11</sup> We use the 2013 version of the Preference Module.

manufacturing conditions and agreeableness as elicited in the Big Five, altruism as elicited via the Preference Module, and by tendency a low score in Machiavellianism. Thus, there seems to be a moral kind of personality.<sup>12</sup>

We thus find that context as well as personality matter. Studies of economic institutions in which agents take morally relevant decisions are a pressing and strongly growing field.<sup>13</sup> Our results indicate that conducting market analyses involving agents who apply moral self-licensing is an important next step for further research.

## 2. Related Literature

While we explore a *static* decision context, the classical literature on moral self-licensing has focused on effects *over time*. It has been documented that ethical behavior in the past serves as a justification to act less ethically later on. Moral self-licensing over time was first described by Monin and Miller (2001) in the contexts of racism and sexism in a two-stage experiment. Subjects had to make a hypothetical job decision in a neutrally framed environment. In one treatment group, the best applicant was African-American, while in the other group, the best applicant was white. After making their decision, all subjects were confronted with another hypothetical hiring scenario: A chief of police had to hire a new deputy in a racially charged job environment. All subjects were asked to rate whether the job was better suited for a white or a black person. Subjects with the opportunity to present themselves as non-prejudiced at stage one were significantly more likely to prefer a white person now. Thus, subjects who could demonstrate they were not racist in the first step, tended to prefer the white police applicant later on. A second study yielded similar findings in the field of sexism.

Sachdeva, Ilic, and Medin (2009) found that subjects use moral self-licensing over time to justify selfish behavior. Subjects who were asked to write positive short stories about their own good past deeds were later on more selfish than others when it came to making a real donation to charity. In a related vein, Mazar and Zhong (2010) demonstrated that subjects who acquire “green” products in a first step in time share significantly less money later on compared to subjects who could only acquire conventional products in a first step in time.<sup>14</sup> For more examples of moral self-licensing over time, see Merritt et al. (2010).

Another closely related concept from the field of behavioral economics is conscience accounting as proposed in Gneezy et al. (2014). In the first stage of their experiment, subjects could lie in order to

---

<sup>12</sup> This fits well to other recent studies linking real economic behavior to personality facets, such as Deckers et al. (2016) and Albrecht et al. (2016). See Rothenhäusler et al. (2016) for incorporating such moral heterogeneity into an institution design model.

<sup>13</sup> Recent contributions are e.g., Sobel (2010), Falk and Szech (2013), Kirchler et al. (2016), Pigors and Rockenbach (2016), Falk and Szech (2014), Rothenhäusler et al. (2015), Kerschbamer et al. (2016), Bartling et al. (2015), Friedrichsen and Engelmann (2014).

<sup>14</sup> Subjects had a 1 out of 25 chance to actually receive their selected products.

increase their own profit at the expense of their fellow players. In stage two, subjects had the opportunity to make a small donation to charity. The authors find that the share of donations is significantly higher among those who lied in stage one than among those who told the truth. This suggests that by donating, subjects atone for past moral norm violations. Furthermore, the share of donations among the liars drops significantly if the time delay between the two stages is increased. A possible explanation is that the remembrance of one's own unethical behavior fades out over time. Another finding is that subjects are more likely to lie if they know that they have the opportunity to donate at a later point in time. This leads to the conclusion that people set off their future ethical behavior against their current one, and goes well in line with the idea that all kind of moral behavior may load on one aspect of personality (Loewenstein 1999, Bénabou and Tirole 2002, Falk and Szech 2016). The authors also document that a comparatively cheap future possibility of "compensation" may be enough for subjects to keep a high moral self-image.

### 3. Design of the Study

We explore the effects of moral self-licensing across different ethical facets in purchase decisions and across different points in time. To this end, we implement a product choice regarding different towels, as well as a later donation situation. In addition to behavior, we study beliefs of outsiders about the behavior of decision-makers in order to analyze whether beliefs of outsiders are in line with real choice behavior, or not. We therefore conduct four treatments: **Product Baseline, Product Organic, Belief Baseline, and Belief Organic.**

**The Product Treatments:** In **Product Baseline** and in **Product Organic**, subjects take ethically relevant product decisions. Subjects choose between ethically different yet otherwise comparable towels. In a price list, they weigh money and receiving a more conventionally produced towel against receiving no money and a towel that fulfills more ethical production standards. We choose towels that could be useful for many of our subjects, independent of gender, body-size, and specific style in clothing. We let subjects choose between towels in neutral colors, in a standard, medium size,<sup>15</sup> and with a standard surface weight.<sup>16</sup>

**The Belief Treatments:** In **Belief Baseline** and in **Belief Organic**, we elicit in an incentivized way outsiders' beliefs about the behavior of subjects in Product Baseline and in Product Organic, respectively. The Belief treatments inform us about what uninvolved third parties, who do not take the decisions themselves, predict about those involved.

The instructions inform subjects from all treatments that there are orthogonal, ethically relevant dimensions when it comes to producing textiles such as towels. We focus on two important, separate

---

<sup>15</sup> The size of all towels is 100 cm x 50 cm.

<sup>16</sup> All towels weigh about 450 g/m<sup>2</sup>.

aspects. Firstly, the cotton used can be conventional or certified as “organic.”<sup>17</sup> All subjects know from the instructions that cotton certified as organic is produced without the use of agrochemicals, such as synthetic pesticides, herbicides, and fertilizers that are commonly used in conventional production and pollute the soil. Secondly, the manufacturing conditions under which workers have to work when sewing the towel can be certified by the well-established Fair Wear Foundation (FWF), or not.<sup>18</sup> Subjects are informed that under FWF conditions, safety, economic living standards, and political rights of workers are monitored. This is not the case in conventional production.<sup>19</sup>

Subjects further know from the instructions that the production chain in the garment industry is fragmented. The cotton is typically grown in different countries than where the manufacturing takes place. Subjects thus know that these two aspects, growing organic cotton and ensuring FWF controlled manufacturing standards, can be considered orthogonal to each other.

In order to analyze moral self-licensing within product decisions, we elicit the monetary value subjects assign to FWF manufacturing standards, in Product Baseline as well as in Product Organic. In the latter treatment, subjects know that the towel already fulfills another independent, ethically relevant criterion (organic cotton). Decisions are elicited in a price list. Subjects choose between a towel that was manufactured under FWF controlled standards and no additional money versus a towel without FWF controlled manufacturing standards and additional money. Monetary amounts vary from 0.25 to 12 euro in steps of 25 cent. Subjects know that one of their decisions will be randomly drawn and implemented.

We define the lowest monetary amount for which subjects still prefer no FWF certificate over the FWF certificate as the willingness to pay for FWF certification. The lower the switch-point, the less a subject cares about FWF certified working standards.

---

<sup>17</sup> We rely on the well-known Global Organic Textile Standard (GOTS) and the Organic Content Standard (OCS) which certify the organic origin of the cotton.

<sup>18</sup> The FWF is a well-established, independent non-profit organization with the aim of improving labor standards in factories based on the conventions of the International Labour Organisation (ILO).

<sup>19</sup> Indeed, the textile and garment industry is one of the ethically most frequently discussed industries, compare, e.g., Danzer and Grundke (2016).

Option A	Your Choice		Option B
organic + 12.00€	<input type="checkbox"/>	<input type="checkbox"/>	organic & fair wear + 0€
organic + 11.75€	<input type="checkbox"/>	<input type="checkbox"/>	organic & fair wear + 0€
organic + 11.50€	<input type="checkbox"/>	<input type="checkbox"/>	organic & fair wear + 0€
.	<input type="checkbox"/>	<input type="checkbox"/>	.
.	<input type="checkbox"/>	<input type="checkbox"/>	.
.	<input type="checkbox"/>	<input type="checkbox"/>	.
organic + 0.50€	<input type="checkbox"/>	<input type="checkbox"/>	organic & fair wear + 0€
organic + 0.25€	<input type="checkbox"/>	<input type="checkbox"/>	organic & fair wear + 0€

**Fig. 1: Price list as used in Product Organic.** In the Organic Product treatment, subjects choose in a price list between a towel without FWF Certificate and money versus a towel with FWF certificate and no additional money. The FWF certificate ensures controlled manufacturing standards for workers when sewing the towel. The towel is for sure made from organic cotton.

Figure 1 illustrates how the price list looked like in Product Organic. Product Baseline was similar, except that the cotton involved was conventional.<sup>20</sup> After this main decision, subjects take part in questionnaires about their political preferences, personal characteristics, and socio-economic background. We use standard tests on personality such as the Big Five (Rammstedt and John 2005) and on Machiavellianism (Christie and Geis 1970), as well as the rather new Preference Survey Module (Falk et al 2016). At the end of the study, about half an hour later, the instructions confront subjects with another ethically relevant decision. Subjects do not know this in advance.<sup>21</sup> Via the computer screen, subjects need to decide whether they want to share their show-up fee (which is 2 euro, i.e., a substantial amount of the total monetary amount subjects can earn in the study) with refugees from a local refugee camp or not.

In the **Belief treatments**, i.e., in Belief Baseline and in Belief Organic, subjects do not take any ethically relevant product or donation decision themselves. Instead, subjects indicate their expectations how subjects in the respective Product treatments behaved. We incentivize these estimates via quadratic scoring rules.<sup>22</sup> In Belief Baseline, subjects guess the average willingness to pay for controlled manufacturing standards in Product Baseline. Subjects in Belief Baseline know the instructions from the Product Baseline treatment for the towel decision. They thus have the exact same information as subjects in Product Baseline. After indicating their estimate, they also fill out the same questionnaires as subjects in the Product treatments. Then, the instructions confront them with a

<sup>20</sup> See the Appendix for details.

<sup>21</sup> Therefore, in contrast to the study of Gneezy et al. (2014), subjects cannot use the later sharing decisions as a justification for selfish behavior in the towel decision.

<sup>22</sup> The estimation for the mean willingness to pay is given in cent and has the following payout:  
 $\max \{500 - \frac{1}{4} (\text{true} - \text{guess})^2, 0\}$  cent.

second estimation task, requiring them to guess what percentage of subjects in Baseline opted to share their show-up fee with refugees. Again, we employ a quadratic scoring rule to incentivize this task.<sup>23</sup>

In Belief Organic, subjects accordingly guess the behavior from Product Organic. In total, subjects in Belief Baseline and in Belief Organic can earn up to 10 euro for good estimates, in addition to a show-up fee of 7 euro.

The study took place at the Karlsruhe Decision and Design Laboratory (KD<sup>2</sup> Lab) at the Karlsruhe Institute of Technology. We used ORSEE (Greiner 2004) in order to recruit 200 subjects in total (50 subjects per treatment). Subjects were placed in separate cubicles each. Instructions regarding the towels were presented on paper, while questionnaires were presented on computer screen, via SoSci Survey (Leiner 2014). See the Appendix for details.

#### 4. Hypotheses

The insights gained from the experiments on moral self-licensing (Monin and Miller 2001, Sachdeva et al. 2009, Mazar and Zhong 2010) and conscience accounting (Gneezy et al. 2014) show that people use acquired moral credentials in order to justify less ethical behavior at a later point in time. In a similar fashion, they morally redeem themselves from unethical behavior in the past by performing subsequent ethical deeds. Even if ethical deeds are quite small and cheap later on, they seem to serve as a good justification for rather selfish behavior at an earlier point in time (Gneezy et al. 2014).<sup>24</sup>

Moral cleansing may provoke a reduction of moral behavior in the first place specifically if subjects *know* about the possibility to morally cleanse themselves afterwards (compare Gneezy et al. 2014). Such a mechanism could also exist in the framework of a static decision across different ethically relevant facets of one product. The satisfaction of fulfilling one ethical dimension could lead to a decreased valuation of a second, though unrelated ethical dimension, at the very same point in time. This would imply that it is not necessarily the time dimension that renders moral cleansing relevant. Instead, it is the opportunity of having some kind of “moral excuse.” This logic would be specifically attractive if subjects did not care about consequences, but instead considered all these decision relevant for one specific aspect about themselves, i.e., a positive self-image (Loewenstein 1999, Bénabou and Tirole 2002, Falk and Szech 2016). Our first and main hypothesis is therefore that subjects in Product Organic have a lower willingness to pay (WTP) for controlled manufacturing conditions than subjects in Product Baseline, i.e., static moral self-licensing exists. We therefore expect that

---

<sup>23</sup> The payout for the estimated willingness to share is  $\max \{5 - \frac{1}{5} (\text{true} - \text{guess})^2, 0\}$  euro. Subjects need to type in integer percentage values.

<sup>24</sup> Gneezy et al. (2014) point out that a similar logic may have been behind the trade of pardons as practiced by the Catholic Church in medieval times. Yet there, redemption often appeared to be quite costly.

$$WTP_{\text{Product Organic}} < WTP_{\text{Product Baseline}}.$$

In addition to the static product decision, subjects later face another unrelated ethical decision by choosing whether to share their show-up fee with local refugees or not. In line with the findings on moral self-licensing over time, the certainty of having acquired a product that fulfills (at least) one ethical facet in the first stage should “free” the subjects to behave less ethical afterwards. This leads to our second hypothesis: Subjects in Product Organic are less likely to share their show-up fee than subjects in Product Baseline, i.e., static moral self-licensing and moral self-licensing over time can amplify each other, i.e.,

$$\text{Donation}_{\text{Product Organic}} < \text{Donation}_{\text{Product Baseline}}.$$

Exploratively, we moreover investigate the incentivized beliefs of outsiders about the behavior of subjects in the respective Product treatments. To this purpose, we conduct two Belief treatments, Belief Baseline and Belief Organic. In these treatments, subjects do not face trade-offs between more ethical decisions versus money. Instead, they are incentivized to predict the decisions of other subjects from the respective Product treatment.

It may be that subjects in the Belief treatments predict that moral self-licensing will take place in the Product treatments. This is not so clear, however, for at least three reasons. First, subjects in the Belief treatments do not face a trade-off between more ethical decisions and money. In the literature, it has been documented that hypothetical questions and real behavior often correlate when it comes to altruistic versus selfish behavior. Yet typically, in real behavior, subjects care much more about their self-interest than when hypothetically asked. Accordingly, the literature on ethical consumption speaks of the famous 30:3 ratio (Cowe and Williams 2000) comparing stated intent from questionnaires to real decisions in the market. Second, it has been documented that there can be differences in “hot” versus “cold” morally relevant decision-making (see Teper et al. 2011, FeldmanHall et al. 2012).<sup>25</sup> Maybe subjects do not figure in these effects when predicting the product and donation decisions of others. For example, it could be that subjects underestimate willingness to pay for ethically improved production decisions in the first place as they are outside of the situation, but then they do not count in effects of moral self-licensing in subsequent donation decisions either. If so, effects of moral self-licensing could become underestimated. Third, it has been shown that context can matter drastically for morally relevant judgments (see e.g., O’Fallon and Butterfield 2005, Loe et al. 2000, Bandura 1999, 2016, Falk and Szech 2013, 2016, Bartling et al. 2015, Kirchler et al. 2016). We therefore expected that

$$\text{Belief}_{WTP\_Product Organic} \leq \text{Belief}_{WTP\_Product Baseline},$$

---

<sup>25</sup> See also Van Boven and Loewenstein (2005) on empathy gaps.

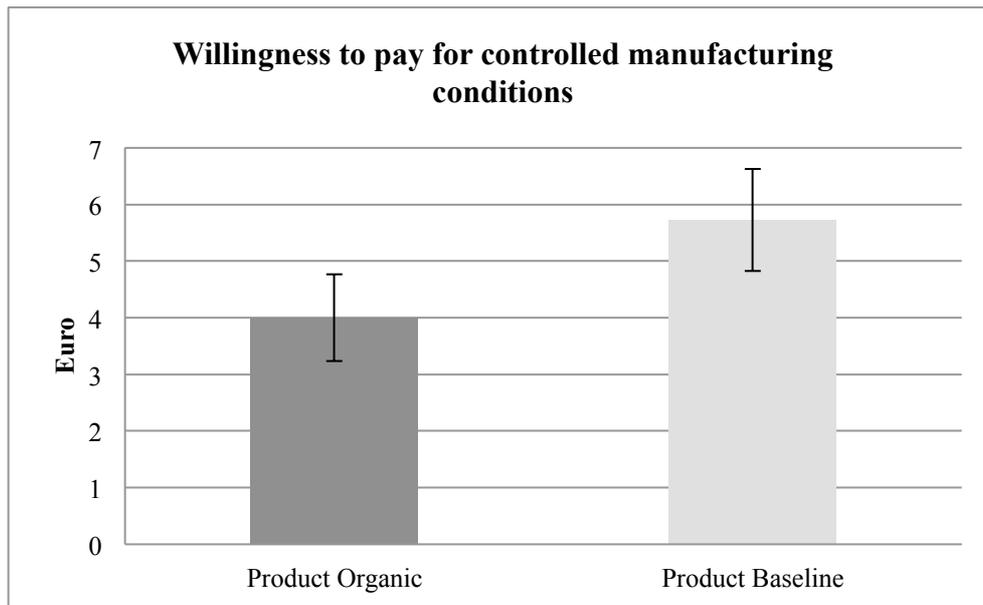
and accordingly

$$\text{Belief}_{\text{Donation\_Product Organic}} \leq \text{Belief}_{\text{Donation\_Product Baseline}}.$$

If subjects do not predict moral self-licensing to take place, i.e., if both formulas hold with equality, it may be of specific relevance to emphasize the importance of moral self-licensing in social and public debate.

## 5. Results

**Static Moral Self-Licensing.** We hypothesized that static moral self-licensing takes place, i.e., subjects treat orthogonal, ethically relevant aspects as if they were substitutes in product choice. Therefore, we expected that subjects in Product Organic would pay less money to ensure FWF-controlled manufacturing standards than in Product Baseline, as reflected in a lower switch-point in the price list. This is exactly what we find. In Product Organic, the average switch-point and therefore willingness to pay is 4.00 euro. In Product Baseline, the average switch-point is 5.73 euro.<sup>26</sup> This is an increase of 43 percent, a highly significant increase ( $p=0.002$ , one-sided).



**Fig. 2:** In Product Organic, subjects pay highly significantly less money to ensure controlled manufacturing standards compared to Product Baseline. Static moral self-licensing thus takes place.

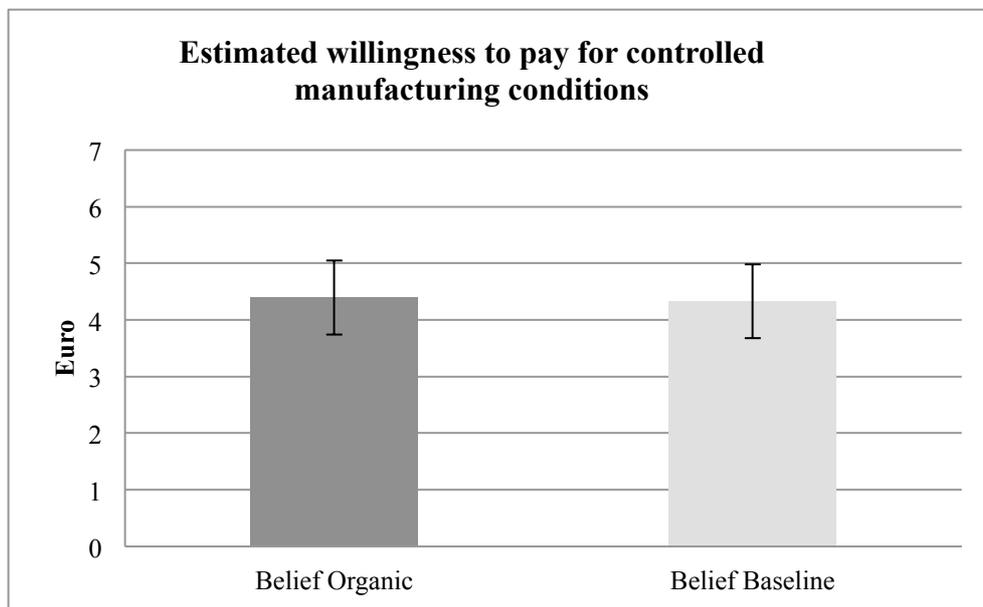
Subjects know from the instructions that sewing the towel is a very different production step than growing the cotton, with totally different people involved and different ecological consequences. Nevertheless, they attribute less value to guaranteeing minimum working standards when they know

---

<sup>26</sup> In case of multiple switching, we use the mean switch point for our analysis. We have seven multiple switchers in the sample. Our findings remain robust if we exclude them or use the first switch point, last switch-point or median switch-point in order to approximate their willingness to pay. See the robustness checks in Section 6 for details.

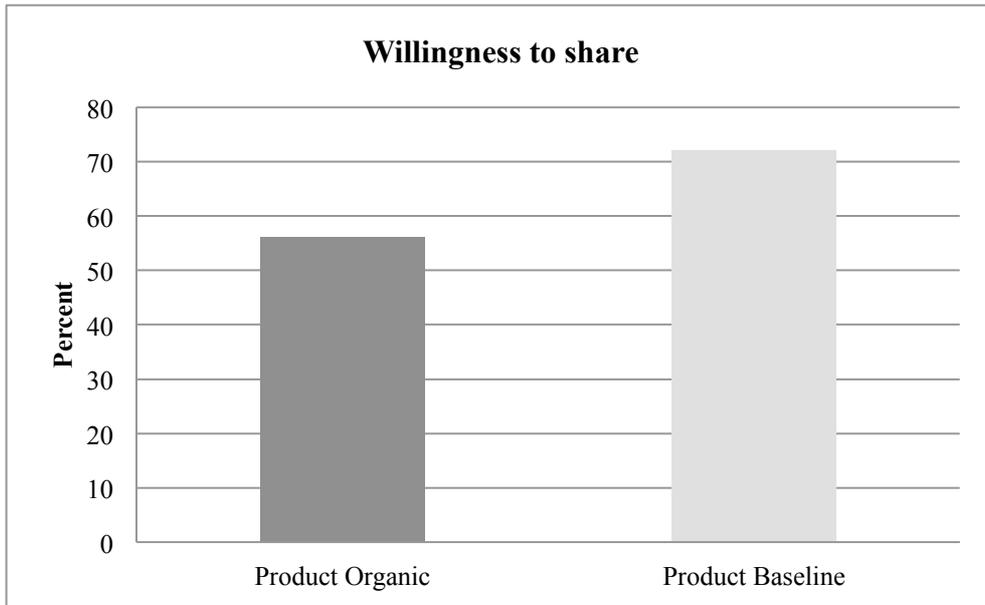
that their towel is made from organic cotton. Potentially, subjects already have a quite good moral conscience in the latter case, and accordingly care less about the workers in manufacturing. We thus find that static moral self-licensing exists.

**Beliefs about Static Moral Self-Licensing.** Do uninvolved people expect static moral self-licensing to take place? In order to discuss effects of static moral self-licensing, people must be aware of it. We hypothesized that for at least three reasons, this may not be the case. Our data from Belief Organic demonstrate that indeed, subjects are not aware of effects of static moral self-licensing. Estimates for switch-points are at any conventional level statistically comparable to those from Belief Baseline (4.39 vs. 4.33,  $p=0.89$ , two-sided). If anything, there is a (non-significant) tendency that subjects expect decision-makers to pay a little bit more in Product Organic than in Product Baseline, which would be even contradictory to effects of moral self-licensing.



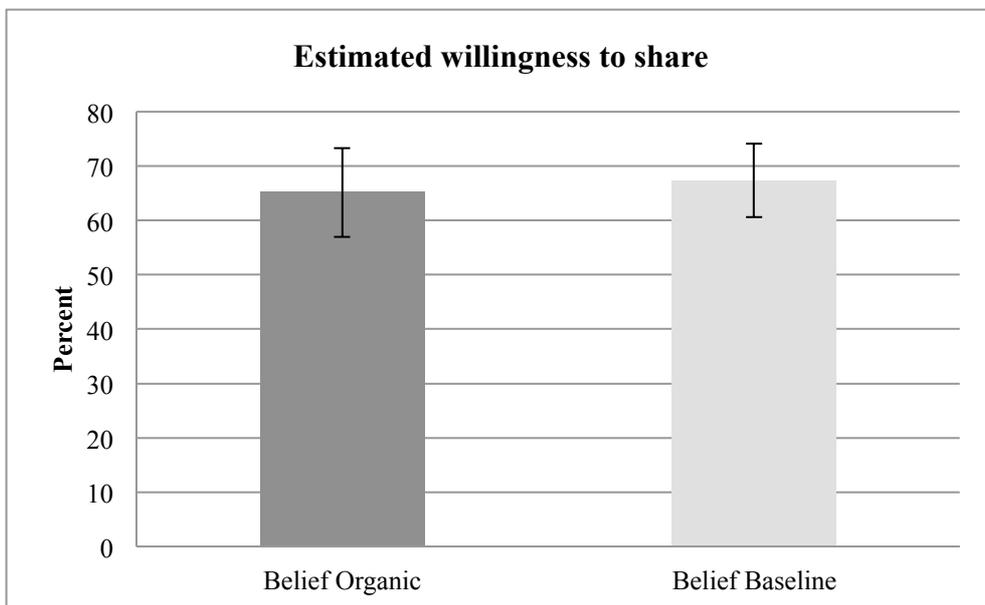
**Fig. 3:** In the Belief treatments, subjects guess the average willingness to pay for controlled manufacturing in Product Organic, resp. Product Baseline. The data display no significant difference in estimates at any conventional level. Thus, subjects in Belief Organic are not aware that static moral self-licensing may take place.

**Moral Self-licensing over time.** At the end of the study, about half an hour later, subjects in the Product treatments were asked whether they wanted to share their show-up fee of two euro with refugees from a local refugee camp. Subjects did not know beforehand that they would be confronted with this decision. If static moral self-licensing has spillovers in time, subjects in Product Organic should less likely share their show-up fee with refugees than subjects in Product Baseline. This is what we find. 56 percent of subjects are willing to share in Product Organic, while 72 percent of subjects are willing to share in Product Baseline ( $p=0.048$ , one-sided test of proportions).



**Fig. 4:** Subjects in Product Organic are significantly less likely to share their show-up fee of two euro with local refugees than subjects in Product Baseline.

Looking into the estimates from the two Belief treatments, there is again no significant difference in estimates of willingness to share with refugees for the two Product treatments (65 percent in Belief Organic vs. 67 percent in Belief Baseline,  $p=0.68$ , two-sided). Outsiders thus do neither expect static moral self-licensing nor moral self-licensing over time.



**Fig. 5:** Estimates from subjects in the Belief treatments for average willingness to share in the respective Product treatment.

All in all, while effects of moral self-licensing with regard to specific ethical dimensions are pronounced within decisions and over time, uninvolved parties fail to anticipate its effects. This happens even though uninvolved parties were monetarily incentivized, and could earn much more money from making accurate guesses than from the flat payment they received anyways for

participating in the study. Potentially, subjects in the Belief treatments just indicate what they consider morally appropriate, and forget that effects of moral-licensing provide excuses to subjects in the Product treatments that justify more selfish donation decisions in the Product Organic treatment than in Product Baseline. Data from the product choices indicate that also a hot versus cold empathy gap may play a role. In the product choice data, we see that beliefs of outsiders tend to be closer to the real product decisions from Product Organic than to the more generous decisions from Product Baseline.

## 6. Robustness

In the following, we study the robustness of our effects. First, we look in detail into the behavior of multi-switchers. Second, we control for income effects. Neither of both aspects affects the results.

When eliciting the willingness to pay with the help of a price list, there is the possibility of subjects giving inconsistent answers, i.e., switching multiple times between the two product options. In order to ensure that those multi-switchers do not affect the results, we test for robustness using different measures of willingness to pay for multi-switchers. We therefore analyze the following specifications in the data for handling multi-switchers as well: excluding multi-switchers from the analysis, using the first switch-point, using the mean switch-point, using the median switch-point, and using the last switch-point.

<b>Switch-point without multi-switchers</b>			
Product Baseline	Product Organic	p-value (one-sided)	n
5.64	3.78	0.0015**	93
<b>First switch-point (including multi-switchers)</b>			
Product Baseline	Product Organic	p-value (one-sided)	n
5.61	3.70	0.0008***	100
<b>Mean switch-point (including multi-switchers)</b>			
Product Baseline	Product Organic	p-value (one-sided)	n
5.73	4.00	0.0021***	100
<b>Median switch-point (including multi-switchers)</b>			
Product Baseline	Product Organic	p-value (one-sided)	n
5.71	3.85	0.0011***	100
<b>Last switch-point (including multi-switchers)</b>			
Product Baseline	Product Organic	p-value (one-sided)	n
5.89	4.29	0.0075***	100

**Table 1:** Switch-points in the Product treatments, depending on different specifications for multi-switchers; p-values presented are from the t-tests between the treatments. Significance levels marked as follows: \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

The differences in willingness to pay between the treatments remain significant for all types of measurement. Thus, the effect is robust.

Another important aspect regarding the robustness of the findings is the potential role of disposable income. Willingness to pay for controlled manufacturing conditions could be dependent on the financial background of subjects. As our sample is made up of students, maximizing the payoff from the experiment in order to get by, could specifically matter for students with low disposable income. Reassuringly, we do not see any income effects. Subjects above median willingness to pay for controlled manufacturing neither have significantly more money at their disposal ( $p=0.23$ , two-sided),<sup>27</sup> nor are they significantly less concerned about their own financial situation ( $p=0.34$ , two-sided) than subjects with a below median willingness to pay for controlled manufacturing. We thus find that our results are robust, both to multi-switching and to financial background of subjects.

## 7. Personality facets and preferences

While the data display highly significant differences in willingness to pay across Product treatments, we also observe significant differences regarding *what kind of personalities* are willing to pay for controlled manufacturing standards for workers.<sup>28</sup> In this section, we therefore analyze the personality differences between those who pay above versus those who pay below median in their respective treatment.

In a nutshell, as expected, we find that subjects above median willingness to pay score higher on agreeableness than subjects below. Furthermore, they show no significant differences among the other traits of the Big Five<sup>29</sup>, which indicates discriminatory validity. In the Preference Module, we obtain a similar pattern in the data. Where expected, we see significant or at least marginally significant results. Items for which we did not expect a correlation do not correlate indeed, with the one exception of present-biasedness (which correlates positively with a high willingness to pay for controlled manufacturing). In the Appendix, we provide a detailed analysis of the results. Here, we stress the main findings regarding personality facets, economic preferences, donation behavior, and political orientation.

Among the Big Five personality facets which we elicited according to the Big Five Inventory (BFI) (Rammstedt and John 2005), agreeableness is the one facet one would expect to correlate with willingness to pay for improved manufacturing. Agreeableness encompasses positive interpersonal attitudes and behaviors, such as generosity, trustworthiness and compassion (Costa and McCrae 1992,

---

<sup>27</sup> If anything, they have less monthly money at their disposal (378 euro vs. 431 euro after rent).

<sup>28</sup> In the Appendix, we present additional analysis for the respective Product treatments separately. By and large, personality facets have comparable impacts in both Product treatments which indicates robustness. See the Appendix for details.

<sup>29</sup> The other traits of the Big Five are conscientiousness, openness to experience, neuroticism and extraversion.

Evans and Revelle 2008). Agreeable individuals have been found to show a high interest in maintaining positive interpersonal relationships (Jensen-Campbell and Graziano 2001). Furthermore, they frequently engage in helping others and in altruistic behavior. Agreeableness has been positively linked to fairness (Kalshoven et al. 2011) and cooperation (Koole et al. 2001), and negatively linked to opportunism (Sakalaki and Fousiani 2012). Therefore, we expected a positive correlation between agreeableness and willingness to pay for controlled manufacturing standards. Exactly as expected, our data document that subjects with a willingness to pay above median for controlled manufacturing are highly significantly more agreeable ( $p=0.0003$ , one-sided) than subjects below median. The findings are further backed up by the data from the long version of the Preference Survey Module (Falk et al. 2016).<sup>30</sup> Altruism, as captured in hypothetical donations and the stated willingness to donate, correlates significantly with the willingness to pay for improved manufacturing standards. Furthermore, subjects above median willingness to pay score higher in positive reciprocity, negative reciprocity, and trust. While overall, results are thus as expected, we observe one pattern in the module data that surprises us: Several times, self-assessment and stated intent do not display the same pattern. This indicates that subjects who tend to score high are nevertheless rather critical with themselves. It may be interesting to explore this in more detail in future research.

Looking for altruism in our incentivized data, we obtain the same result. Subjects above median willingness to pay are also more likely to share their show-up fee with refugees. Aggregating the Product treatments, subjects with a willingness to pay for controlled manufacturing conditions above median are more likely to donate in the later decision context compared to subjects below median (73 percent vs. 55 percent,  $p=0.058$ , two-sided test of proportions). This difference is not that pronounced and non-significant if we focus on Product Baseline alone, possibly because donation rates are high in this treatment for all subjects (76 percent for above median subjects, 68 percent for below median subjects). In Product Organic, however, the contrast is drastic. While 70 percent of subjects above median donate, only 42 percent do so among subjects below median ( $p=0.042$ , two-sided test of proportions).

We furthermore find that subjects with a high willingness to pay for improved manufacturing conditions also tend to be more present-biased than other subjects.<sup>31</sup> We did not expect this correlation. A possible explanation could be that impulsive subjects are more inclined to immediately forfeit money in order to behave ethically if directly presented with the opportunity. Such subjects may be specifically reactive to a “hot” situation, i.e., a context in which they can directly act. In line with Albrecht et al. (2016), subjects willing to pay above median for improved manufacturing conditions furthermore state a higher hypothetical willingness to pay for better husbandry conditions

---

<sup>30</sup> See the Appendix for details.

<sup>31</sup> See the Appendix for details.

of livestock ( $p=0.001$ , two-sided). They also state to be more engaged in voluntary work ( $p=0.07$ , two-sided).

The regression table 2 shows that willingness to pay for controlled manufacturing is driven by two factors: moral self-licensing in the organic treatment, and a general ethical personality that manifests itself in agreeableness and willingness to donate. In contrast, gender, disposable money, as well as financial concerns do not seem to significantly influence willingness to pay for improved manufacturing conditions.

We furthermore investigate if the subjects' political and social backgrounds affect the willingness to pay. We find that the willingness to pay for improved manufacturing varies across different political orientations. Subjects with a willingness to pay above median are more likely to vote for liberal parties ( $p=0.01$ , two-sided test of proportions) than those with a below median willingness to pay, who favor more conservative parties.<sup>32</sup> These results are also in line with findings from Albrecht et al. (2016) as well as Deckers et al. (2016). As in the latter, we find that context and personality facets both have significant impact on morally relevant behavior. Thus personality differences play an important role in addition to decision context.

treatment organic	-1.73*** (0.59)	-1.74*** (0.59)	-1.76*** (0.59)	-1.78*** (0.60)	-1.57*** (0.57)	-1.48** (0.56)
female		0.44 (0.67)	0.40 (0.68)	0.48 (0.68)		
disposable income			-0.001 (0.001)			
financial concerns				0.17 (0.36)		
willingness to donate					0.33*** (0.11)	0.28** (0.11)
agreeableness						0.24** (0.10)
Adj. R-squared	0.07	0.07	0.06	0.06	0.14	0.17

**Table 2:** OLS regressions of mean changes in the willingness to pay for controlled manufacturing conditions in euro. Standard errors in parentheses,  $n=100$ , significance levels marked as follows: \*  $p<0.1$ , \*\*  $p<0.05$ , \*\*\*  $p<0.01$ .

<sup>32</sup> We asked for the preferred party in case of a federal election in Germany. We clustered the Social Democrats (SPD), the Greens (Bündnis 90/Die Grünen) which is a party with environmentally orientated policies, and The Left (Die Linke) as a more politically liberal, socially oriented party. We cluster the Christian Democrats (CDU/CSU) and the Free Democrats (FDP) as more politically conservative parties.

## **8. Conclusion**

Our data document that people behave as if there existed strong substitutive relations between ethically relevant, non-correlated product dimensions. We thus find that static moral self-licensing exists. Fulfilling one ethical dimension in a product choice further eases moral conscience in a later, unrelated yet morally relevant context. Thus, different aspects of moral self-licensing, static (across morally relevant dimensions at one point in time), and over different points in time, can act together and thereby amplify overall effects of moral self-licensing.

Outsiders do not infer that moral self-licensing occurs. Even though monetarily incentivized to make accurate estimates, subjects are completely oblivious to effects of moral self-licensing, both, static and over time. They are unaware of its importance, though effects are drastic. Therefore, social and political debate needs to be informed about the importance of moral self-licensing, within decisions regarding different ethical dimensions as well as over time.

In future research, market analyses involving customers who apply facets of moral self-licensing should be carried out. It may be that for firms, offering products that fulfill exactly one ethical criterion, e.g., ecological cotton in textile products, is a smart strategy for selling products to customers with easy-to-ease ethical concerns. If such practices turn out profitable, political and societal debate should be aware of these effects.

## 9. References

- Albrecht K, Krämer F, Szech N (2016) Ethics and animal welfare: A personality study. Working Paper, Karlsruhe Institute of Technology, Germany.
- Apple (2016) Conflict Minerals Report 2016.
- Bandura A (1999) Moral disengagement in the perpetration of inhumanities. *Personality and social psychology review* 3(3): 193-209.
- Bandura A (2016) *Moral Disengagement How People Do Harm and Live with Themselves* (Worth Publishers, Macmillan Learning, New York).
- Bartling B, Weber RA, Yao L (2015) Do markets erode social responsibility? *Quart. J. Econom.* 130(1): 219–66.
- Bénabou R, Tirole J (2002) Self-confidence and personal motivation. *Quart. J. Econom.* 117(3): 871-915.
- Burke J (2013) Bangladesh factory fires: fashion industry's latest crisis. *The Guardian* (December 8), <http://www.theguardian.com/world/2013/dec/08/bangladesh-factory-fires-fashion-latest-crisis>
- Christie R, Geis F (1970) *Studies in Machiavellianism* (Academic Press, New York).
- Costa PT, McCrae RR (1992) Four ways five factors are basic. *Personality and Individual Differences* 13 (6): 653–65.
- Consumer Technology Association (2015) 2015 Sustainability Report.
- Cowe R, Williams S (2000) *Who are the ethical consumers?* (The Co-operative Bank, London).
- Danzer AM, Grundke R (2016) Coerced labor in the cotton sector: How global commodity prices (don't) transmit to the poor. Working paper.
- Deckers T, Falk A, Kosse F, Szech N (2016) Homo moralis: Personal characteristics, institutions, and moral decision-making. Working Paper, Karlsruhe Institute of Technology, Karlsruhe, Germany.

Duhigg C, Barboza D (2012) In China, human costs are built into an iPad. *New York Times* (January 25), [http://www.nytimes.com/2012/01/26/business/ieconomy-apples-ipad-and-the-human-costs-for-workers-in-china.html?\\_r=2&](http://www.nytimes.com/2012/01/26/business/ieconomy-apples-ipad-and-the-human-costs-for-workers-in-china.html?_r=2&)

Evans AM, Revelle W (2008) Survey and behavioral measurements of interpersonal trust. *Journal of Research in Personality* 42: 1585–1593.

Falk A, Becker A, Dohmen TJ, Huffman D, Sunde U (2016) The preference survey module: A validated instrument for measuring risk, time, and social preferences. *SSRN Electronic Journal*.

Falk A, Szech N (2013) Morals and markets. *Science* 340 (6133): 707–11.

Falk A, Szech N (2014) Organizations, diffused pivotality and immoral outcomes. Discussion Paper No. 9522, Center for Economic Policy Research, London, UK.

Falk A, Szech N (2016) Pleasures of skill and moral conduct. Working paper, Karlsruhe Institute of Technology, Karlsruhe, Germany.

FashionUnited (2016) Global fashion industry statistics - International apparel. Accessed on February 10, 2017.

<https://fashionunited.com/global-fashion-industry-statistics>

FeldmanHall O, Mobbs D, Evans D, Hiscox L, Navrady L, Dalgleish T (2012) What we say and what we do: the relationship between real and hypothetical moral choices. *Cognition* 123(3):434-41.

FiBL & IFOAM (2016) The World of Organic Agriculture – Statistics and Emerging Trends 2016.

Friedrichsen J, Engelmann D (2014). Who cares for social image? Interactions between intrinsic motivation and social image concerns. Working paper, CESifo Working Paper Series No. 4514, Munich, Germany.

Gert B (2012) The definition of morality. Zalta EN, ed. *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/spr2016/entries/morality-definition>

Gneezy U, Imas A, Madarász K (2014) Conscience accounting: Emotion dynamics and social behavior. *Management Sci.* 60(11): 2645–58.

Greiner B (2004) An online recruitment system for economic experiments. Kremer K, Macho V, eds. *Forschung und wissenschaftliches Rechnen 2003*, GWDG Bericht 63 (Ges. für Wiss. Datenverarbeitung, Göttingen), 79–93.

H&M (2015) Conscious Actions – Sustainability Report 2015.

International Labour Organization (2014) Wages and working hours in the textiles, clothing, leather and footwear industries.

International Labour Organization (2016) Improving working conditions in the ready made garment industry: Progress and achievements. Accessed on November 6, 2016.

[http://www.ilo.org/dhaka/Whatwedo/Projects/WCMS\\_240343/lang--en/index.htm](http://www.ilo.org/dhaka/Whatwedo/Projects/WCMS_240343/lang--en/index.htm)

Joint Economic Committee (2015) The economic impact of the fashion industry. Washington, DC: United States Congress.

Jensen-Campbell LA, Graziano WG (2001) Agreeableness as a moderator of interpersonal conflict. *Journal of Personality* 69(2): 323–62.

Kalshoven K, Den Hartog D, De Hoogh A (2013) Ethical leadership and follower helping and courtesy: Moral awareness and empathic concern as moderators. *Applied Psychology* 62: 211–235.

Kerschbamer R, Neururer D, Sutter M (2016) Insurance coverage of customers induces dishonesty of sellers in markets for credence goods. *Proc. Natl. Acad. Sci. USA* 113(27): 7454–58.

Kirchler M, Huber J, Stefan M, Sutter M (2016) Market design and moral behavior. *Management Sci.* 62(9): 2615-2625.

Koole SL, Jager W, van den Berg AE, Vlek CAJ, Hofstee WKB (2001) On the social nature of personality: Effects of extraversion, agreeableness, and feedback about collective resource use on cooperation in a resource dilemma. *Personality and Social Psychology Bulletin* 27(3): 289–301.

Leiner DJ (2014) SoSci Survey (Version 2.5.00-i) [Computer software]

Levitt S, List J (2007) What do laboratory experiments measuring social preferences reveal about the real world? *The Journal of Economic Perspectives* 21(2): 153-174.

Loe TW, Ferrell L, Mansfield P (2000) *J. Bus. Ethics* 25: 185.

Loewenstein G (1999) Because it is there: The challenge of mountaineering...for utility theory *Kyklos* 52(3): 315–43.

Mazar N, Zhong CB (2010) Do green products make us better people? *Psych. Sci.* 21(4): 494–98.

Merritt AC, Effron DA, Monin B (2010) Moral self-licensing: When being good frees us to be bad: Moral self-licensing. *Social and Personality Psychology Compass* 4(5): 344–57.

Monin B, Miller DT (2001) Moral credentials and the expression of prejudice. *J. Personality Soc. Psych.* 81(1): 33.

O'Fallon M, Butterfield K (2005) A review of the empirical ethical decision-making literature: 1996-2003. *J. Bus. Ethics* 59(4): 375-413.

Pigors M, Rockenbach B (2016) Consumer social responsibility. *Management Sci.* 62(11): 3123-3137.

Rammstedt B, John OP (2005) Kurzversion des Big Five inventory (BFI-K) *Diagnostica* 51(4): 195–206.

Rothenhäusler D, Schweizer N, Szech N (2016) Guilt in voting and public good games. Working paper.

Rauthmann JF, Kolar GP (2012) How ‘dark’ are the dark triad traits? Examining the perceived darkness of narcissism, Machiavellianism, and psychopathy. *Personality and Individual Differences* 53(7): 884–89.

Sachdeva S, Iliev R, Medin DL (2009) Sinning saints and saintly sinners the paradox of moral self-regulation. *Psych. Sci.* 20(4): 523–28.

Sakalaki M , Fousiani K (2012) About some personality misfortunes of opportunists: The negative correlation of economic defection with autonomy, agreeableness, and well-being. *Journal of Applied Social Psychology* 42(2): 471–87.

Sakalaki M, Richardson C, Thépaut Y (2007) Machiavellianism and economic opportunism. *Journal of Applied Social Psychology* 37(6): 1181–90.

Sobel J (2010) Markets and other-regarding preferences. Working paper.

Teper R, Inzlicht M, Page-Gould E (2011) Behavior and moral forecasting. *Psych. Sci.* 22(4): 553-55.

Textile Exchange (2014) Organic cotton market report 2014.

UL (2014) Verifying environmental sustainability in the electronics marketplace. Underwriter Laboratories.

US Department of Agriculture ERS (2016) *Organic market overview*. Accessed November 8, 2016. <https://www.ers.usda.gov/topics/natural-resources-environment/organic-agriculture/organic-market-overview/>

Van Boven L, Loewenstein G (2005) Empathy gaps in emotional perspective taking. Malle BF, Hodges SD, eds. *Other Minds: How Humans Bridge the Divide between Self and Others*, (Guilford New York), 284-297.

## Appendix

### Index

I.	Instructions .....	25
a.	Instructions of the Product treatments .....	25
b.	Instructions of the Belief treatments .....	29
II.	Personality Traits .....	33
a.	Preference Survey Module .....	33
i.	Altruism .....	33
ii.	Positive Reciprocity .....	35
iii.	Negative Reciprocity .....	36
iv.	Trust .....	38
v.	Time Discounting .....	39
vi.	Risk Taking .....	41
b.	The Big Five and Machiavellianism .....	42
i.	Agreeableness .....	42
ii.	Conscientiousness .....	43
iii.	Neuroticism .....	44
iv.	Openness to experience .....	44
v.	Extraversion .....	45
vi.	Machiavellianism .....	46
III.	References (Appendix) .....	47

## **I. Instructions**

### **a. Instructions of the Product treatments**

*In the following, we present the instructions from Product Organic. The instructions from Product Baseline were identical, except that subjects in Product Baseline had the choice between two towels always of conventional instead of organic cotton. The instructions were presented on paper sheets, including a cover sheet and a prize list on the back of the cover sheet.*

### **Welcome to this Study!**

Thank you for participating in an economic study at the Karlsruhe Institute of Technology (KIT).

As in all economic studies at KIT, all circumstances described in the following are true. Your decisions will be implemented exactly as described.

We would ask you to keep quiet during the study. If you have questions, please indicate by slightly opening your door. Your questions will then be answered at your cubicle.

**Here, please fill in your subject ID.**

**Subject-ID:** \_\_\_\_\_

**Please start reading the two pages inside the cover sheet before attending to the list on the back page.**

## **The production line of the garment industry**

In the course of this study, you will choose between two towels, which differ in their way of production.

En route from the cotton plant to the final product ready to sell, a towel passes several production steps. Two essential production steps are the cultivation of cotton and the sewing of the final product. These production steps are often done in regions far away from each other. This is why the conditions of the different production steps often vary widely.

Relating to the production standards in both production steps – cultivation and sewing - there exist established labels.

### **What does “certified organic cotton” mean?**

This label is relevant to the cultivation of cotton. In contrast to conventionally grown cotton, certified organic cotton (organic cotton) is cultivated without chemically produced fertilizers, defoliants, or pesticides. Also the use of genetically modified organisms (GMOs) is prohibited. These conditions must be met in order to become certified, and additionally the soil must have been free from synthetic agrochemicals for three years. In the context of the certification process, the plantations are audited at least once a year from independent institutions.

### **What does “Fair Wear Foundation” mean?**

This label refers to the sewing of the final product. The Fair Wear Foundation (FWF) is an international verification initiative. The affiliated producers in the garment industry commit to the step-by-step implementation of social minimum standards in their manufacturing factories. This is audited by independent parties. Among the labor standards of the FWF are: employment is freely chosen, no discrimination in employment, no exploitation of child labor, freedom of association and the right to collective bargaining, payment of a living wage, no excessive working hours, safe and healthy working conditions, and legally binding employment relationships.

## **Your task**

In this study, you will make a choice for a towel. You will receive in any case a towel made out of organic cotton of the size of 100cm x 50cm that will be handed out to you when the study has ended. All the towels in this study have the same grammage of 450g/m<sup>2</sup> and are held in neutral colors.

## **Price list and remuneration**

On the backside of the cover sheet is a list. In each row you will have the choice between Option A and Option B.

**Option A:** You receive a towel made out of organic cotton without Fair Wear certification. Additionally, you receive a monetary amount.

**Option B:** You receive a towel made out of organic cotton with Fair Wear certification. You do not receive an additional monetary amount.

The monetary amount in Option A varies over the rows of the table. It starts at 12.00 euro and decreases in every row by 25 cent.

Please make a decision for every row by marking exactly one of the two options in every row.

Afterwards, one of the rows will be randomly selected. The probability is 1 out of 48 for every row. Your decision in the selected row will be exactly implemented as described. That means that you will receive your chosen towel and – depending on the decision – additional money.

Please note: If a row is chosen in which you did not mark exactly one option, you will not receive anything.

For your participation in the study, you will additionally receive 2 euro.

There is a colored piece of paper with your participation number attached to the cover sheet. This piece of paper serves as your coupon with which you can pick up your payout (towel and money) at the KD2 Lab as soon as you have received the respective email. Your remuneration will be in a package so that the person handing it out will not know the content.

Please keep the coupon until your remuneration. Otherwise, you cannot be remunerated.

Take your time to make yourself aware of the differences between the two options.

*After the subjects made their choice and filled out the price list, they proceeded to fill out the personality and preference tests on a computer. These included the Big Five Inventory (BFI) (Rammstedt and John 2005), the Mach-IV test for Machiavellianism (Christie and Geis 1970), as well as the Preference Survey Module (Falk et al. 2016), and further questions on social and political preferences.*

*After completing the questionnaires, subjects from Product Baseline and Product Organic could decide whether they wanted to share their show-up fee with local refugees. They did not know in advance to be confronted with this decision.*

**Now, you have the opportunity to share your show-up fee of 2 euro in a 1:1 ratio. You keep one half and the other half benefits local refugees.**

- Share the show-up fee (keep 1 euro, 1 euro for refugees)
- Don't share the show-up fee (keep 2 euro, 0 euro for refugees)

## **b. Instructions of the Belief treatments**

*In each Belief treatment, subjects were asked to guess the behavior of subjects from a Product treatment, respectively. Therefore, in addition to the instructions of the respective Product treatment (see Appendix I.a.), subjects in a Belief treatment also received the following, additional instructions. The additional instructions were identical for Belief Baseline and Belief Organic. The footnotes beneath the payout table were part of the instructions.*

### **Welcome to this Study!**

Thank you for participating in an economic study at the Karlsruhe Institute of Technology (KIT).

As in all economic studies at KIT, all circumstances described in the following are true. Your decisions will be implemented exactly as described.

We would ask you to keep quiet during the study. If you have questions, please set the wheel on your door to red.

In this study, you will be shown the instructions of subjects of a previous study. These instructions are on the colored sheets.

Please read these instructions carefully. Later, you will be asked to estimate how the subjects of the previous study behaved. This determines a part of your remuneration.

In any case you will receive a show-up fee of 7€ for participating in this study.

**Please read the colored sheets first and take a look at the list. Then continue with the sheets on white paper that contain your instructions.**

### Your task

You had the opportunity to read the instructions for previous subjects. In that study, subjects were asked for their willingness to pay<sup>I</sup> for Fair manufacturing conditions with the aid of a price list. This willingness to pay could be located within the whole range of the price list.

**Please estimate the average willingness to pay among the previous subjects.<sup>II</sup>**

Depending on the accuracy of your estimation, you can earn additional money. The closer you get to the average willingness to pay of the previous subjects, the more additional money you can earn.<sup>III</sup> This can be seen in detail in the following table.

**Payout table**

<b>Deviation in cent</b>	<b>Additional money in euro</b>
0	5.00
10	4.75
15	4.44
20	4.00
25	3.44
30	2.75
35	1.94
40	1.00
44	0.16
45 or more	0

---

<sup>I</sup> The willingness to pay is the amount of money for which the subjects of the previous study switched from the right column (Fair Wear without additional money) to the left column (no Fair Wear, additional money).

<sup>II</sup> The arithmetic mean of the willingness to pay of every subject is calculated. [If a subject has switched multiple times between the columns, we use the arithmetic mean of every willingness to pay. If a subject always preferred a towel without a Fair Wear certification, we define his willingness to pay as 0.25€. If a subject always preferred the Fair Wear towel, we define his willingness to pay as 12.25€.]

<sup>III</sup> Your payout for this estimation amounts to  $\max \{500 - \frac{1}{4} (r - x)^2, 0\}$  cent, whereas x is your estimation and r the mean willingness to pay.

**Please indicate how high was the average willingness to pay of previous subjects (only values between 0.25 € and 12.25 € are technically permissible).**

\_\_\_\_\_ Euro

**Your Subject-ID:** \_\_\_\_\_

*After the subjects in the Belief treatments completed the personality and preference questionnaires (these were identical to the questionnaires in the Product treatments), they were asked to estimate how many percent of the previous subjects in the respective Product treatment were willing to donate to local refugees. Subjects did not know they would be confronted with this second estimation task beforehand.*

In the following, you can see a screenshot of another decision option of previous subjects.

-- Begin Screenshot --

**Now, you have the opportunity to share your show-up fee of 2 Euro in a 1:1 ratio. You keep one half and the other half benefits local refugees.**

- Share the show-up fee (keep 1 Euro, 1 Euro for refugees)
- Don't share the show-up fee (keep 2 Euro, 0 Euro for refugees)

-- End Screenshot --

**Please estimate what percentage of the subjects was willing to share their show-up fee of 2 euro in a 1:1 ratio, thus picking the upper option.**

Depending on the accuracy of your estimation, you can earn additional money.\*

<b>Payout table</b>	
<b>Deviation in percentage points</b>	<b>Additional money in euro</b>
0	5
1	4.8
2	4.2
3	3.2
4	1.8
5	0

**Please indicate what percentage of the previous subjects, in your opinion, was willing to share their show-up fee of 2 Euro in a 1:1 ratio. Please enter the percentage in whole numbers. Values between 0 and 100 percent are technically permissible.**

\_\_\_\_\_ Percent

\* Your payout for this estimation amounts to  $\max \left\{ 5 - \frac{1}{5} (r - x)^2, 0 \right\}$  euro, whereas x is your estimation and r is the percentage of previous subjects that were willing to share their show-up fee.

## II. Personality Traits

In the following, we explore in more depth the relation between personality facets and preferences and willingness to pay above versus below median for controlled manufacturing.<sup>36</sup>

### a. Preference Survey Module

We employ the 2013 long version of the Preference Survey Module by Falk et al. to elicit the correlations between preferences and willingness to pay for improved manufacturing conditions. The Preference module was developed by Falk et al. (2016), who calibrated it with students from Bonn: Hypothetical questions that correlated well with incentivized behavior were selected to become part of the Preference Module. Our data thus also serves as a validation of the selection of questions used in the Preference Module. All in all, we find that real behavior goes well together with preferences as elicited via the module.

The module contains six preference measures with three items each, focusing on risk taking, time discounting, altruism, trust, positive reciprocity, and negative reciprocity. We elicit the preferences as recommended by the authors, i.e., we elicit all of them. The results indicate convergent as well as discriminatory validity (i.e., preferences that one would not expect to be of relevance for ethical product choices indeed do not correlate). In the following, we state the results, and restate the items as administered to the subjects.

#### i. Altruism

Altruism has been defined as the willingness to do things that favor others, even if they come at a cost for oneself (Fehr and Fischbacher 2003). The Preference Module measures altruism via the items *hypothetical donation*, *willingness to donate to charity*, and *altruistic inclination*. Reducing one's own payoff for controlled manufacturing conditions can be seen as an act of altruism. We therefore had the strong hypothesis that subjects above median willingness to pay for improved manufacturing score higher on the items of altruism than subjects below median willingness to pay. As expected, we find a positive correlation between altruism and willingness to pay for improved manufacturing.

	WTP above median	WTP below median	p-value (one-sided)
Altruism 1	118.35	58.88	0.015**
Altruism 2	7.29	5.98	0.005***
Altruism 3	3.98	3.78	0.356

**Table 3:** Subjects above median willingness to pay give away higher hypothetical donations and are more willing to donate to charity than subjects below the median. One-sided t-test, n=98, \* p<0.1, \*\* p<0.05, \*\*\* p<0.01.

---

<sup>36</sup> We leave subjects exactly at the median out of the median split. As there are only two subjects exactly at the median, the findings remain robust to counting subjects at the median to the subgroups.

Exactly as expected, subjects with a willingness to pay for improved manufacturing above median give significantly larger hypothetical donations and are highly significantly more willing to donate to charity as elicited in the Preference Module. Yet, both groups indicate the same understanding for altruistically driven people.

	Altruism 1	Altruism 2	Altruism 3
Altruism 1	1.000		
Altruism 2	0.396 (0.000)***	1.000	
Altruism 3	0.028 (0.780)	0.017 (0.868)	1.000

**Table 4:** Spearman rank-order correlation, n=100, significance levels in parentheses. While we see a highly significant correlation between the two questions on donation behavior, neither is correlated to the self-assessed apprehension of altruistic behavior. Significance levels marked as follows: \* p<0.1, \*\* p<0.05, \*\*\* p<0.01.

The spearman correlations of the items draw a similar picture: Hypothetical donation and the willingness to donate to charity are highly significantly correlated (p=0.000), while neither of them is significantly correlated to altruistic *inclination*. This may be due to the fact that the first two items are both linked to charitable donations whereas the latter is a more abstract question on the understanding of altruistic motives.

The items from the Preference Module were the following.

**Altruism 1. Hypothetical donation** *Imagine the following situation: Today you unexpectedly received 1000 Euro. How much of this amount would you donate to charity? (Values between 0 and 1000 are allowed)*

**Altruism 2. Self-assessment: Willingness to donate to charity** *How would you assess your willingness to share with others without expecting anything in return, for example your willingness to give to charity? Please use a scale from 0 to 10, where a 0 means you are “completely unwilling to give up something” and a 10 means you are “very willing to give up something.” You can also use the values in-between to indicate where you fall on the scale.*

**Altruism 3. Self-assessment: Altruistic inclination** *How well does the following statement describe you as a person? I do not understand why people spend their lifetime fighting for a cause that is not directly beneficial for them. Please use a scale from 0 to 10, where 0 means “does not describe me at all” and a 10 means “describes me perfectly”. You can also use the values in between to indicate where you fall on the scale.*

## ii. Positive Reciprocity

Positive reciprocity has been understood as willingness to return kindness to the benefactor (see e.g., Caliendo et al. 2012). In line with Burger et al. (2009), we argue that engaging in positive reciprocity can be seen as an indication of acting in accordance to social norms. Therefore, we expected that subjects above median willingness to pay may tend to score higher on positive reciprocity.

	WTP above median	WTP below median	p-value (one-sided)
Positive Reciprocity 1	12.79	11.80	0.161
Positive Reciprocity 2	5.20	4.51	0.163
Positive Reciprocity 3	8.88	8.27	0.047**

**Table 5:** Subjects above the median report a significantly higher effort to return a favor and send back more money as a second mover in a hypothetical trust game. One-sided t-test, n=98, \* p<0.1, \*\* p<0.05, \*\*\* p<0.01.

In line with our expectation, more subjects above median willingness to pay put more effort into returning a favor (p=0.047, one-sided). By a (non-significant) tendency, they also score higher on the second mover behavior in a hypothetical trust game and are more willing to return a favor.

The effort to return a favor correlates significantly with the willingness to return a favor (p=0.002), indicating that subjects' self-assessment on the reaction to favors remains consistent. The second mover behavior is furthermore significantly correlated to the effort to return a favor (p=0.099), but not to the willingness to return a favor. This may be caused by the posing of the question that is more specific and hence more similar to the second mover behavior than the more general item 2.

	Positive Reciprocity 1	Positive Reciprocity 2	Positive Reciprocity 3
Positive Reciprocity 1	1.000		
Positive Reciprocity 2	0.096 (0.343)	1.000	
Positive Reciprocity 3	0.166 (0.099)*	-0.307 (0.002)***	1.000

**Table 6:** Spearman rank-order correlation, n=100. The willingness to return a favor and the effort to return a favor significantly correlate to a highly significant extent. The negative correlation is due to the inverse scale. Significance levels in parentheses. Significance levels marked as follows: \* p<0.1, \*\* p<0.05, \*\*\* p<0.01.

This is how the Preference Module picks up positive reciprocity:

**Positive Reciprocity 1. Hypothetical trust game: Second mover behavior** *Please consider the following situation: You and another person, whom you do not know, both participate in a study where you can decide on how to assign a certain amount of money and thereby determine the outcome. The rules are as follows. Both participants get an account with 20 Euros. At the beginning, both*

participants thus own 20 Euros. The other person decides first. She can transfer money to your account. She can transfer any amount: 0, 1, 2 Euro, etc. up to 20 Euro. Each Euro that she transfers to you is tripled by the conductors of the study and booked to your account. After this first stage the other person therefore has 20 Euro minus the amount she transferred to you in her account. You have 20 Euro plus the tripled amount of the transfer of the other person on your account. Now you get to decide: you have the opportunity to transfer money back to the other person. You can transfer any amount up to 80 Euro, depending on how much you have in your account. This will be the end of the study and the account balances will be final. The other person has in her account 20 Euros minus the amount she transferred to you plus the amount you transferred back. You have 20 Euro plus the tripled amount of what the other person transferred to you minus the amount you transferred back to her. We would like to know how much you would choose to transfer back to the other person, for a given transfer of her to you.

Suppose the other person transfers 5/10/15/20 Euro to your account. After the first stage you then own  $20+3*5/10/15/20=35/50/65/80$  Euro the other person owns  $20-5/10/15/20=15/10/5/0$  Euro. What amount do you choose to transfer back?

**Positive Reciprocity 2. Self-assessment: Willingness to return a favor** How would you assess your willingness to return a favor to a stranger? Please use a scale from 0 to 10, where 0 means you are “not willing to return a favor” and a 10 means you are “very willing to return a favor”. You can also use the values in-between to indicate where you fall on the scale.

**Positive Reciprocity 3. Self-assessment: Effort to return a favor** How well does the following statement describe you as a person? I put a lot of effort in it to help someone who helped me in the past. Please use a scale from 0 to 10, where 0 means “does not describe me at all” and a 10 means “describes me perfectly”. You can also use the values in-between to indicate where you fall on the scale.

### iii. Negative Reciprocity

Negative reciprocity occurs if someone who was subjected to unfavorable behavior answers with a similar, non-friendly reaction (see, e.g., Perugini et al. 2002). Negative reciprocity has also been interpreted as an indicator for inequality aversion (Fehr and Schmidt 1999). Therefore, we expected a positive link between the willingness to pay for controlled manufacturing conditions and the score on negative reciprocity. Indeed, in the data, we find some support that this is the case.

	WTP above median	WTP below median	p-value (one-sided)
Negative Reciprocity 1	38.90	32.63	0.017 **
Negative Reciprocity 2	6.86	6.41	0.170
Negative Reciprocity 3	5.57	5.98	0.771

**Table 7:** Subjects above the median have higher minimum acceptable offers in an Ultimatum Game. One-sided t-test, n=98, \* p<0.1, \*\* p<0.05, \*\*\* p<0.01.

Subjects above the median report significantly higher minimum acceptable offers in a hypothetical trust game (p=0.017). This rejection of low offers is an indication for the stronger inequality aversion of subjects above median willingness to pay. The other two items of negative reciprocity, however, are not significantly linked with the willingness to pay. The data display a highly significant correlation between the minimum acceptable offer in an ultimatum game and the willingness to punish unfair behavior (p=0.000). This indicates that subjects use the rejection of low offers as a means to punish unfair behavior.

	Negative Reciprocity 1	Negative Reciprocity 2	Negative Reciprocity 3
Negative Reciprocity 1	1.000		
Negative Reciprocity 2	0.399 (0.000)***	1.000	
Negative Reciprocity 3	-0.013 (0.897)	0.152 (0.130)	1.000

**Table 8:** Spearman rank-order correlation, n=100, significance levels in parentheses. There is a highly significant correlation between the minimum acceptable offer and the willingness to punish unfair behavior. Significance levels marked as follows: \* p<0.1, \*\* p<0.05, \*\*\* p<0.01.

The exact wording of the items from the Preference Module is as follows.

**Negative Reciprocity 1. Minimum acceptable offer in an ultimatum game** *Consider the following: Together with a person you do not know you have won 100 Euro in a contest. The rules are as follows. One of the two of you has to make a proposal on how to divide the money between the two of you. The other one is informed about the proposal and has two options. Either he can accept the proposal or he can reject it. If he accepts the proposal, the proposed division of the money is implemented. If he rejects the proposal both get nothing. Suppose the other person makes the proposal about the division of the money. You then have to decide whether to accept or to reject the proposal. Which minimum amount does the other person have to offer you for you to accept the proposal?*

**Negative Reciprocity 2. Self-assessment: Willingness to punish unfair behavior** *Are you a person who is generally willing to punish unfair behavior even if this is costly? Please use a scale from 0 to*

10, where 0 means you are “not willing at all to incur costs to punish unfair behavior” and a 10 means you are “very willing to incur costs to punish unfair behavior”. You can also use the values in-between to indicate where you fall on the scale.

**Negative Reciprocity 3. Self-assessment: Willingness to take revenge** *How well does the following statement describe you as a person? If somebody hurts me on purpose, I will try to hurt them back. Please use a scale from 0 to 10, where 0 means “does not describe me at all” and a 10 means “describes me perfectly”. You can also use the values in-between to indicate where you fall on the scale.*

**iv. Trust**

Trust is essential for social interactions and the basis for any economic activity due to the omnipresence of incomplete contracts (Arrow 1972). We find that subjects above median willingness to pay for controlled manufacturing conditions have more favorable expectations of other people’s intentions. It is plausible that subjects who can trust others also put more value in the adherence to social norms in general. Going a step further, subjects who see unfamiliar people in a more positive light may also want to ensure good working conditions to them.

	WTP above median	WTP below median	p-value (one-sided)
Trust 1	11.86	12.16	0.599
Trust 2	6.89	5.74	0.007***
Trust 3	6.12	5.69	0.112

**Table 9:** Subjects above median willingness to pay have highly significant better expectations of other’s intentions than subjects below. Yet, they assess themselves not as significantly more trusting than subjects below the median. One-sided t-test, n=98, \* p<0.1, \*\* p<0.05, \*\*\* p<0.01.

The data display that subjects above median willingness to pay are significantly more willing to trust into others’ intentions. Yet in self-assessment, scores are not different from subjects with a low willingness to pay for controlled manufacturing. Again, we see that self-assessment and stated intent do not display the same patterns.

	Trust 1	Trust 2	Trust 3
Trust 1	1.000		
Trust 2	0.121 (0.231)	1.000	
Trust 3	0.139 (0.167)	0.513 (0.000)***	1.000

**Table 10:** Spearman rank-order correlation, n=100, significance levels in parentheses. The willingness to trust others and the belief about other’s intentions are highly significantly correlated, while the first mover behavior shows no significant correlation to any of the items. Significance levels marked as follows: \* p<0.1, \*\* p<0.05, \*\*\* p<0.01.

The exact items of the Preference Module were as follows.

**Trust 1. Hypothetical trust game: First mover (after reading the instructions for the Trust Game, see paragraph on Positive Reciprocity)** *Suppose you were assigned the role of the other person. Which amount would you choose to transfer?*

**Trust 2. Self-assessment: Belief about other’s intentions** *How well does the following statement describe you as a person? As long as I am not convinced otherwise I always assume that people have only the best intentions. Please use a scale from 0 to 10, where 0 means “does not describe me at all” and a 10 means “describes me perfectly”. You can also use the values in-between to indicate where you fall on the scale.*

**Trust 3. Self-assessment: Trust** *How would you assess your willingness to trust others, for example your willingness to trust strangers? Please use a scale from 0 to 10, where 0 means you are “completely unwilling to trust” and a 10 means you are “very willing to trust”. You can also use the values in-between to indicate where you fall on the scale.*

#### v. Time Discounting

Subjects ranking high in time discounting as elicited in the Preference Module strongly prefer a sooner consumption of a good compared to a later consumption.

	WTP above median	WTP below median	p-value (two-sided)
Time Discounting 1	140.39	126.39	0.049**
Time Discounting 2	7.33	8.59	0.006***
Time Discounting 3	7.08	6.51	0.285

**Table 11:** Subjects with a WTP above median for controlled manufacturing show a significantly higher time discounting than subjects below. Two-sided t-test, n=98, \* p<0.1, \*\* p<0.05, \*\*\* p<0.01.

We find that subjects with a willingness to pay for controlled manufacturing above median tend to be more present-biased than subjects below median. We did not expect this correlation. Potentially, more present-biased subjects are more prone to immediately seize the opportunity of behaving ethically, if confronted with such an option. The staircase measure is significantly correlated to the self-assessed willingness to wait and the self-assessed tendency to postpone things. Yet, the latter two items do not significantly correlate. Thus again, the data display a discrepancy in stated intent and self-assessment across the different sub-groups. Thus, the difference between behaviors and intents on the one side, and self-assessment on the other, may be an interesting topic of investigation for future research.

	Time Discounting 1	Time Discounting 2	Time Discounting 3
Time Discounting 1	1.000		
Time Discounting 2	-0.319 (0.001)***	1.000	
Time Discounting 3	0.2393 (0.017)**	-0.121 (0.230)	1.000

**Table 12:** Spearman rank-order correlation, n=100, significance levels in parentheses. The staircase measure correlates highly significantly with the other two items, the willingness to wait and the tendency to postpone things. The latter two items however do not significantly correlate. Significance levels marked as follows: \* p<0.1, \*\* p<0.05, \*\*\* p<0.01.

Here we restate the items exactly as they were used in the Preference Module, and we used them as well.

**Time Discounting 1: Staircase measure** *Suppose you were given the choice between the following: receiving a payment today or a payment in 12 months. We will now present to you 5 situations. The payment today is the same in each of these situations. The payment in 12 months is different in every situation. For each of these situations we would like to know which you would choose. Please consider the following: would you rather receive 100 Euro today or \_\_\_ Euro in 12 Months?*

**Time Discounting 2: Self-assessment: Willingness to wait** *How well does the following statement describe you as a person? I abstain from things today so that I will be able to afford more tomorrow. Please use a scale from 0 to 10, where 0 means “does not describe me at all” and a 10 means “describes me perfectly”. You can also use the values in-between to indicate where you fall on the scale.*

**Time Discounting 3: Self-assessment: Tendency to postpone things** *How well does the following statement describe you as a person? I tend to postpone things even though it would be better to get them done right away. Please use a scale from 0 to 10, where 0 means “does not describe me at all” and a 10 means “describes me perfectly”. You can also use the values in-between to indicate where you fall on the scale.*

## vi. Risk Taking

Risk preference measures the tendency to opt for more risky or less risky options when exposed to uncertainty (see e.g., Păunescu and Cantaragiu 2012). The three main types of risk preference are risk-seeking, risk neutral and risk-averse preferences. We did not expect a correlation between the risk preferences and the willingness to pay for controlled manufacturing conditions.

	WTP above median	WTP below median	p-value (two-sided)
Risk 1	105.31	107.55	0.826
Risk 2	6.16	6.20	0.926
Risk 3	434.08	396.92	0.227

**Table 13:** There is no significant link between the willingness to pay and risk attitudes as subjects above and below the median do not significantly differ in any of the items. Two-sided t-test, n=98, \* p<0.1, \*\* p<0.05, \*\*\* p<0.01.

Indeed, we do not find any significant difference in risk preferences between subjects above and subjects below the median. The staircase measure and the estimate of certainty equivalent are highly significantly correlated. The estimated certainty equivalent is also correlated to the self-assessed willingness to take risks.

	Risk 1	Risk 2	Risk 3
Risk 1	1.000		
Risk 2	0.110 (0.270)	1.000	
Risk 3	0.704 (0.000)***	0.176 (0.08)*	1.000

**Table 14:** The staircase measure correlates to a highly significant extent with the estimate of the certainty equivalent. The latter also correlates significantly with the willingness to take risks. Spearman rank-order correlation, n=100, significance levels in parentheses. Significance levels marked as follows: \* p<0.1, \*\* p<0.05, \*\*\* p<0.01.

These were the items as used in the Preference Module.

**Risk 1: Staircase measure** *Please imagine the following situation: You can choose between a sure payment and a lottery. The lottery gives you a 50 percent chance of receiving 300 Euro. With an equally high chance you receive nothing. Now imagine you had to choose between the lottery and a sure payment. We will present to you 5 different situations. The lottery is the same in all situations. The sure payment is different in every situation. What would you prefer: a 50 percent chance of winning 300 Euro, when at the same time there is a 50 percent chance of receiving nothing, or would you rather have the amount of \_\_\_ Euro as a sure payment?*

**Risk 2: Self-assessment: Willingness to take risks** *How do you see yourself: are you a person who is generally willing to take risks, or do you try to avoid taking risks? Please use a scale from 0 to 10, where a 0 means you are “completely unwilling to take risks” and a 10 means you are “very willing to take risks”. You can also use the values in-between to indicate where you fall on the scale.*

**Risk 3: Estimate of certainty equivalent** *Please imagine the following situation: you have won a prize in a contest. Now you can choose between two different payment methods, either a lottery or a sure payment. If you choose the lottery there is a 50 percent chance that you receive 1000 Euro and an equally high chance that you receive nothing. Please consider: what would the sure payment need to be in order for you to prefer the sure payment over playing the lottery?*

**b. The Big Five and Machiavellianism**

In the context of ethical decision making and personality, Machiavellianism and the Big Five are two important measures that add additional insights on the relationship between personality facets and the willingness to pay for safety of workers in production. We elicit them via the Big Five Inventory (BFI) (Rammstedt and John 2005) and the Mach-IV test for Machiavellianism (Christie and Geis 1970).

**i. Agreeableness**

Within the five factor model of personality, agreeableness is associated with cooperation and altruism. People scoring high on agreeableness typically strive for social harmony and tend to be more empathetic, and trustworthy (Costa and McCrae 1992, Evans and Revelle 2008). In the context of our experiment, we expected agreeableness to be positively linked to the willingness to pay for controlled manufacturing conditions. This is exactly what we find. Subjects above median willingness to pay score highly significantly higher on agreeableness than subjects below the median. This finding is robust across treatments.

	WTP above median	WTP below median	p- value (one-sided)
Agreeableness	13.59	11.69	0.0003 ***

**Table 15:** Subjects above median willingness to pay are highly significantly more agreeable than subjects below. One-sided t-test, n=98, \* p<0.1, \* p<0.1, \*\* p<0.05, \*\*\* p<0.01.

	<b>Product Baseline</b>		
	WTP above median	WTP below median	p-value (one-sided)
Agreeableness	13.6	12.2	0.032**
	<b>Product Organic</b>		
	WTP above median	WTP below median	p-value (one-sided)
Agreeableness	13.58	11.17	0.0016***

**Table 16:** In both treatments, subjects above median willingness to pay are highly significantly more agreeable than subjects below. Two-sided t-test, n=98, \* p<0.1, \*\* p<0.05, \*\*\* p<0.01.

The five factor model contains four additional personality traits besides agreeableness: conscientiousness, neuroticism, openness, and extraversion.

We did not have a hypothesis for a connection between any of those traits and the willingness to pay for controlled manufacturing conditions. We present the data in an explorative way to deliver a more complete picture of our subjects.

## ii. Conscientiousness

Conscientiousness is a measure for the discipline, self-control, and precision. Conscientious persons tend to be more organized, act more dutifully, and more achievement orientated, while people scoring low on conscientiousness are often more unreliable, more spontaneous and flexible (see Costa and McCrae 1992). As mentioned above, we did not expect correlation between conscientiousness and willingness to pay for controlled manufacturing standards.

	WTP above median	WTP below median	p- value (two-sided)
Conscientiousness	14.84	14.16	0.174

**Table 17:** There are no significant differences in conscientiousness above and below median willingness to pay. Two-sided t-test, n=98, \* p<0.1, \*\* p<0.05, \*\*\* p<0.01.

As expected, there are no significant differences in conscientiousness between subjects above and subjects below median willingness to pay. This is also the case for Product Baseline and Product Organic considered separately.

	<b>Product Baseline</b>		
	WTP above median	WTP below median	p- value (two-sided)
Conscientiousness	14.64	14.28	0.556
	<b>Product Organic</b>		
	WTP above median	WTP below median	p- value (two-sided)
Conscientiousness	15.042	14.04	0.212

**Table 18:** In both treatments, there are no significant differences in conscientiousness between subjects above and below median willingness to pay. Two-sided t-test, n=98, \* p<0.1, \*\* p<0.05, \*\*\* p<0.01.

### iii. Neuroticism

Neuroticism is a measure for the emotional instability of a person. Neurotics are more frequently subjected to fear, anxiety, sensitivity to stress and moodiness, while low scores of neuroticism are associated with a calm and even temper (Rammstedt et al. 2013). We had no specific hypothesis regarding the interplay of neuroticism and the willingness to pay. The data confirm this.

	WTP above median	WTP below median	p- value (two-sided)
Neuroticism	11.08	11.20	0.851

**Table 19:** We find no significant differences in neuroticism between subjects above and below median willingness to pay. Two-sided t-test, n=98, \* p<0.1, \*\* p<0.05, \*\*\* p<0.01.

Subjects above and below median willingness to pay indeed show no significant difference in neuroticism, neither across, nor within treatments.

	<b>Product Baseline</b>		
	WTP above median	WTP below median	p- value (two-sided)
Neuroticism	11.4	11.52	0.886
	<b>Product Organic</b>		
	WTP above median	WTP below median	p- value (two-sided)
Neuroticism	10.75	11.88	0.903

**Table 20:** In both treatments, there are no significant differences in conscientiousness between subjects above and below the median. Two-sided t-test, n=98, \* p<0.1, \*\* p<0.05, \*\*\* p<0.01.

### iv. Openness to experience

People scoring high on openness to experience are rather creative types, curious, flexible, have many interests, and open to emotion. On the contrary, people scoring low on openness tend to like predictable structures, routines, and have a narrower field of interests (Costa and McCrae 1992, Rammstedt et al. 2013).

	WTP above median	WTP below median	p- value (two-sided)
Openness	17.92	17.90	0.978

**Table 21:** We find no significant differences in openness to experience between subjects above and below the median. Two-sided t-test, n=98, \* p<0.1, \*\* p<0.05, \*\*\* p<0.01.

As expected, the data show no significant difference in openness to experience between subjects with a willingness to pay above and below the median. The difference remains insignificant in Product Baseline and Product Organic.

	<b>Product Baseline</b>		
	WTP above median	WTP below median	p- value (two-sided)
Openness	17.56	17.64	0.936
	<b>Product Organic</b>		
	WTP above median	WTP below median	p- value (two-sided)
Openness	18.29	18.17	0.911

**Table 22:** In both treatments, there are no significant differences between subjects with a willingness to pay above and below the median in openness to experience. Two-sided t-test, n=98, \* p<0.1, \*\* p<0.05, \*\*\* p<0.01.

#### v. Extraversion

Extraversion is a trait that marks the behavior in social situations. Extraverted people are often outgoing, talkative, and are perceived as energetic. On the contrary, introverts tend to prefer solitude and behave more passively in groups (compare e.g., Rammstedt et al. 2013).

	WTP above median	WTP below median	p- value (two-sided)
Extraversion	14.14	14.12	0.977

**Table 23:** There are no significant differences in extraversion between subjects above and below median willingness to pay. Two-sided t-test, n=98, \* p<0.1, \*\* p<0.05, \*\*\* p<0.01.

As hypothesized, we find no significant differences in extraversion between subjects above and below median willingness to pay. This finding is robust across treatments.

	<b>Product Baseline</b>		
	WTP above median	WTP below median	p- value (two-sided)
Extraversion	13.68	13.60	0.933
	<b>Product Organic</b>		
	WTP above median	WTP below median	p- value (two-sided)
Extraversion	14.71	14.583	0.906

**Table 24:** In both treatments, there are no significant differences in extraversion between subjects above and below median willingness to pay. Two-sided t-test, n=98, \* p<0.1, \*\* p<0.05, \*\*\* p<0.01.

#### vi. Machiavellianism

Machiavellianism encompasses a variety of world views and actions that are characterized by a general disregard for other people in order to increase personal gain. This behavior is named after the renaissance politician and author Niccolò Machiavelli who proposed that a competent ruler has to be willing to engage in ruthless behavior, such as lying, cheating, deception and treachery in order to succeed (Machiavelli 1513/1992). Sacrificing money in order to improve the working conditions of unknown people without personal benefit is certainly not in line with Machiavellian behavior. This is why we expected a negative correlation between Machiavellianism and the willingness to pay for controlled manufacturing conditions. The results go a bit into this direction, yet not as clearly as hypothesized.

	<b>Product Baseline</b>		
	WTP above median	WTP below median	p- value (one-sided)
Machiavellianism	58.4	55.36	0.0999*
	<b>Product Organic</b>		
	WTP above median	WTP below median	p- value (one-sided)
Machiavellianism	57.42	57.88	0.415

**Table 25:** The difference in Machiavellianism between subjects above and below median willingness to pay is only significant in Product Baseline. One-sided t-test, n=98, \* p<0.1, \*\* p<0.05, \*\*\* p<0.01.

In Product Baseline we obtain a tendency as expected, yet not so in Product Organic. In the aggregate, the correlation is not significant (p=0.131).

	WTP above median	WTP below median	p- value (one-sided)
Machiavellianism	56.37	58.14	0.131

**Table 26:** In line with our expectations, subjects above median willingness to pay score significantly lower on Machiavellianism than their counterparts. One-sided t-test, n=98, \* p<0.1, \*\* p<0.05, \*\*\* p<0.01.

### III. References (Appendix)

Arrow K (1972) Gifts and exchanges. *Philosophy and Public Affairs* 1 (4): 343-362.

Burger JM, Sanchez J, Imberi JE, Grande LR (2009) The norm of reciprocity as an internalized social norm: Returning favors even when no one finds out. *Social Influence* 4 (1).

Caliend M, Fossen F, Kritikos A (2012) Trust, positive reciprocity, and negative reciprocity: Do these traits impact entrepreneurial dynamics? *Journal of Economic Psychology* 33 (2): 394–409.

Christie R, Geis F (1970) *Studies in Machiavellianism* (Academic Press, New York).

Costa PT, McCrae RR (1992) Four ways five factors are basic. *Personality and Individual Differences* 13 (6): 653–65.

Evans AM, Revelle W (2008) Survey and behavioral measurements of interpersonal trust. *Journal of Research in Personality* 42 (6): 1585–93.

Doyle JR (2013) Survey of time preference, delay discounting models. *Judgment and Decision Making* 8(2): 116–135.

Fehr E, Fischbacher U (2003) The nature of human altruism. *Nature* 425: 785-791.

Fehr E, Schmidt KM (1999) A theory of fairness, competition, and cooperation. *Q J Econ* 114 (3): 817-868.

Machiavelli N, Thompson NH (1992) *The Prince. (Dover Thrift Editions)*. (Dover Publications, New York).

Păunescu C, Cantaragiu R (2012) How does the risk preference impact the need for achievement in business contexts. *Management & Marketing Challenges for the Knowledge Society* 7 (1): 89-106.

Perugini M, Gallucci M, Presaghi F, Ercolani AP (2002) The personal norm of reciprocity. *Eur. J. Pers.* 17: 251–283.

Rammstedt B, John OP (2005) Kurzversion des Big Five Inventory (BFI-K). *Diagnostica* 51 (4): 195–206.

Rammstedt B, Kemper CJ, Klein MC, Beierlein C, Kovaleva A (2013) A short scale for assessing the big five dimensions of personality (BFI-10). *Methoden, Daten, Analysen* 7(2): 233-249.