

# Sorting and wage premiums in immoral work

Florian H. Schneider, Fanny Brun and Roberto A. Weber\*

April 6, 2022

We use surveys, laboratory experiments and administrative data to study how heterogeneity in the perceived immorality of work and in workers' aversion to acting immorally impact labor market outcomes. Immoral work is associated with higher wages, both in administrative data and in causal evidence from a laboratory experiment. Using distinct individual measures, we show that those who find immoral behavior least aversive are more likely to be employed in immoral work in the laboratory and exhibit a relative preference for work in sectors perceived as immoral. Such sorting may impact society through the resulting workforce composition across firms and sectors.

Keywords: Wage premium, immoral behavior, sorting, experiments

JEL Codes: C92, J31, D03

---

\* Department of Economics, University of Zurich, Blümlisalpstrasse 10, 8006 Zurich, Switzerland. Schneider: [florian.schneider2@econ.uzh.ch](mailto:florian.schneider2@econ.uzh.ch), Brun: [fanny.brun@econ.uzh.ch](mailto:fanny.brun@econ.uzh.ch), Weber: [roberto.weber@econ.uzh.ch](mailto:roberto.weber@econ.uzh.ch). We thank participants at several conferences and seminars for helpful comments and suggestions. We also gratefully acknowledge the financial support of the Swiss National Science Foundation (Grant No. 100018\_165943).

## 1. Introduction

Immoral workplace behavior has the potential to cause significant social harm. For example, financial industry scandals involving the intentional sale of toxic assets to unsuspecting clients (US Department of Justice, 2016) and the aiding of tax evasion (Hill, 2012) create significant burdens for public funds and erode trust in the financial sector. Tobacco companies have long been accused of engaging in misleading marketing regarding smoking's harmful effects (Heath, 2016) and intentionally targeting underage smokers (Bates and Rowell, 1998). Aggressive marketing of opioids by pharmaceutical firms contributed to a serious public health crisis (Okie, 2010). In cases like these, as well as in less extreme examples, corporate activities that many regard as “immoral” or “unethical”—but are nevertheless profitable—may have serious negative impacts on social welfare.

Rather than representing isolated incidents, there exists a widespread impression that some jobs—e.g., marketing tobacco, manufacturing weapons—likely involve inherently immoral acts. Conventional wisdom further posits that such jobs disproportionately attract those individuals who experience the least displeasure from acting immorally and that workers performing these kinds of jobs receive high wages, a form of compensating differential for unethical conduct.<sup>1</sup> Thus, “immoral” work shares features with other aspects of employment that people may find heterogeneously aversive, such as risk of physical harm (Rosen, 1986), and implies potentially higher labor costs for firms in industries tainted with the perception of immorality. Indeed, anecdotal evidence indicates that companies operating in industries perceived as immoral may face difficulties in recruitment.<sup>2</sup>

However, despite the intuitive appeal of such relationships, little empirical evidence links heterogeneity in the willingness to perform work generally perceived as immoral to resulting differential labor market outcomes; we review this and related work in the next section. Moreover, there are many reasons to believe that the impacts of moral concerns may be mitigated in competitive markets (Fehr, Kirchler, Weichbold, and Gächter, 1998; Levitt and List, 2008; Dufwenberg, et al., 2011; Lazear, Malmendier and Weber, 2013) and related

---

<sup>1</sup> This view dates back to Adam Smith (1776; Book I; Ch. X), who wrote, “The exorbitant rewards of players, opera-singers, opera-dancers, etc., are founded upon [...] the discredit of employing them in this manner. [...] Should the public opinion or prejudice ever alter with regard to such occupations, their pecuniary recompense would quickly diminish. More people would apply to them, and the competition would quickly reduce the price of their labour.” At the time such professions were widely seen as immoral; Smith equated them with “a sort of public prostitution.”

<sup>2</sup> For example, following the Cambridge Analytica scandal, Facebook struggled to attract top talent (CNBC, 2019). Tobacco companies deem difficulties in recruitment arising from their image as sufficiently important to warrant disclosure to regulators and shareholders (British American Tobacco, 2015, p. 37; Philip Morris International Inc., 2015, p. 14).

evidence that markets crowd out inherent moral preferences (e.g., Falk and Szech, 2013; Bartling, Weber and Yao, 2015; Kirchler, Huber, Stefan and Sutter, 2016; Ziegler, Romagnoli and Offerman, 2020). Such concerns might particularly apply to labor markets, where wages are set by the preferences of the marginal worker and where repeatedly forgoing profitable job opportunities may erode workers' aversion to performing immoral acts. On the other hand, moral concerns might be particularly relevant in labor markets, as job choices are typically visible to others and may significantly impact a worker's sense of identity and social image (Bénabou and Tirole, 2011; Oh, 2021).

In this paper, we provide novel evidence testing the above relationships between the perceived immorality of work, workers' heterogeneous concerns for acting morally and outcomes in labor markets. We employ a combination of surveys, laboratory experiments and administrative labor market data, in which we obtain measures of individuals' concerns for morality and relate these to variation in the perceived morality of work. The administrative data provides the clearest evidence of the economic relevance of these relationships, but in these data industries perceived as immoral might differ in other unobservable aspects from other industries, making it difficult to establish a causal relationship from the immorality of work to labor market outcomes. The control provided by laboratory experiments allows us to observe the kinds of outcomes that arise as the nature of work changes *only* in the extent to which it involves doing things generally perceived as immoral. We additionally use evidence from surveys to obtain insights into relationships between individuals' concerns for morality and their willingness to work in different real-world firms and industries. Thus, our paper highlights the value of employing complementary research methods to address complex economic phenomena.

Our work is focused on two specific hypotheses that arise from a simple theoretical analysis of how individuals' heterogeneous aversion to performing work that is generally perceived as immoral may interact with jobs that vary in the immorality of the work they require. The two hypotheses reflect the perceptions that we note above: first, that work perceived as more immoral yields higher wages—as long as workers care enough about avoiding work activities that are immoral—and, second, that immoral work attracts those workers least concerned with immorally conduct.

Our results provide support for these two hypotheses, both in and out of the laboratory. Table 1 provides an overview of our main findings.

First, we show that work perceived as immoral commands a wage premium over comparable work that is not perceived as immoral. We show this in Section 4, using

administrative labor market data, where we find that industries generally perceived as immoral yield higher wages, controlling for observable worker and industry characteristics. Moreover, in our laboratory labor markets—which vary by treatment *only* whether being employed requires doing something generally perceived as immoral—we observe a causal relationship indicating that wages are persistently higher for such immoral work (Section 5). This wage premium is large and does not decrease with market experience, reflecting a stable aversion to immoral work.

**Table 1. Overview of our evidence on wage premiums and sorting**

	<i>Laboratory labor market</i>	<i>Labor markets outside the laboratory</i>
<i>Immorality premium</i>	Causal evidence for a wage premium for immoral work (Section 5; <a href="#">Figure 5</a> )	Correlation between perceived industry immorality and wages in the Swiss labor market (Section 4; <a href="#">Figure 1</a> , <a href="#">Table 2</a> )
<i>Sorting</i>	Immoral types are more likely to be hired, but only for immoral work (Section 5; <a href="#">Figure 6</a> , <a href="#">Table 5</a> )	Immoral types state a greater willingness to work in firms and industries perceived by others as immoral (Section 6; <a href="#">Figure 9</a> , <a href="#">Table 7</a> )

Second, we provide evidence of sorting by “immoral” types into “immoral” work, both in the laboratory and in the field. We obtain two measures of individuals’ aversion to acting immorally, from a behavioral task and from a survey conducted several days apart, and find that the measures are strongly correlated. These two measures of moral concern predict individual labor market outcomes. In our laboratory labor markets (Section 5), immoral types are employed more frequently, but only when work involves doing something immoral. In our survey data (Section 6), immoral types report a significantly greater willingness to work in firms and industries that are generally perceived by others as immoral.<sup>3</sup>

Our findings have potentially important policy implications, which we discuss further in the Conclusion. For example, our evidence of sorting in laboratory and non-laboratory labor markets suggests that if industries with greater potential for social harm—e.g., weapons manufacturing—are perceived as more immoral, they may consequently attract those individuals least concerned with minimizing negative social impacts. Moreover, this sorting process may also mean that perceptions that a firm or industry involves immoral behavior

---

<sup>3</sup> While these job choices are hypothetical, Wiswall and Zafar (2018) and Gill, Heinz, Schumacher and Sutter (2020) provide evidence that stated labor market preferences predict subsequent employment.

may be self-reinforcing, if they lead to a workforce mainly comprising those individuals least concerned with acting morally. Thus, even in industries that may not be inherently immoral, like financial services, the perception that they involve immoral activities may be sufficient to tilt the labor force toward individuals more likely to commit immoral acts.

The rest of our paper proceeds as follows. In the next section, we discuss how this paper relates and contributes to previous literature. Section 3 describes a simple theoretical framework that we use to develop hypotheses. In Section 4, we use Swiss labor market data to investigate the relationship between the perceived immorality of work and wages. Section 5 presents the design and results of our laboratory experiment, while Section 6 describes our survey study. Section 7 concludes by discussing implications of our findings for policy.

## **2. Related Literature**

Our research directly contributes to better understanding the impacts of moral concerns on market behavior and outcomes (Fehr, Kirchler, Weichbold, and Gächter, 1998; Falk and Szech, 2013; Bartling, Weber and Yao, 2015; Kirchler, Huber, Stefan and Sutter, 2016; Ziegler, Romagnoli and Offerman, 2020). Levitt and List (2007, 2008) question whether social preferences matter in markets, due to factors including high stakes, market competition and experience. Our observation of substantial wage premiums for immoral work—both inside and outside of the laboratory—suggest that many workers are willing to forgo financial gains in order not to work in immoral industries. Our survey data indicates that individuals’ aversion to acting immorally predicts sorting into industries. Finally, in the laboratory labor market, we find that neither the wage premium nor sorting decrease with market experience. These findings provide novel evidence on how moral preferences impact both market wages and individual labor market outcomes.

There is related evidence on sorting by pro-social “mission-oriented” types into the public sector (Carpenter and Myers, 2010; Dur and Zoutenbier, 2014; Fisman et al, 2015; Hanna and Wang, 2017; Ashraf, et al., 2019; Barfort, et al., 2019; Friebel, Kosfeld and Thielmann, 2019).<sup>4</sup> However, none of this work tests how differential sorting by heterogeneous types impacts outcomes such as wages. Moreover, our focus on sorting into immoral work can have distinct policy implications. For example, sorting by immoral types

---

<sup>4</sup> There is also some correlational evidence that people in specific industries have different values (Ashraf, Bandiera and Delfino, 2020) and lower trustworthiness (Gill, Heinz, Schumacher and Sutter, 2020). Moreover, Carter and Irons (1991) find that economics students exhibit lower concerns for morality, however, this study does not document that these moral concerns drive differential selection into different kinds of work rather than the relationship being perhaps the other way around (Frank, Gilovich and Regan, 1993; Cohn, Fehr and Maréchal, 2014; Ashraf and Bandiera, 2017).

into jobs where they can cause social harm may often exacerbate the negative social impacts of such work. That is, from society's perspective we may want those people most inclined to act morally working in industries where they can have the greatest impacts on mitigating social harm, but our results suggest the opposite may occur. Conversely, sorting by pro-social types into the public sector is likely to increase the positive social impacts of their work.<sup>5</sup>

Our work also relates to the literature on compensating differentials. A substantial number of studies investigate whether nonprofit employees earn less than for-profit employees (e.g., Leete, 2001; Mocan and Tekin, 2003; Jones, 2015). These studies yield mixed correlational evidence, likely due to methodological challenges in estimating compensating wage differentials using observational data (see the discussion in Mas and Pallais, 2017). Recent papers on compensating differentials instead rely on experimental methods and/or stated preferences (Carpenter, Matthews and Robbett, 2017; Mas and Pallais, 2017; Wiswall and Zafar, 2018). We differ from this work in that we explore variation in the perceived immorality of work and the aversion to acting immorally as the driving sources of heterogeneity and we present the first causal evidence on compensating wage differentials based on the interaction of these job characteristics and workers' preferences.

The closest prior evidence on the compensating differentials we investigate comes from studies documenting wage premiums for presumably immoral behavior, like prostitution (Moffatt and Peters, 2004; Arunachalam and Shah, 2008), and for work in professions and firms with low levels of perceived social responsibility (Frank, 1996). This work shares features with the correlational evidence we provide in Section 4, but focuses on more limited samples and does not study the role of heterogeneous individual preferences. This prior evidence is consistent with the notion that an aversion to performing immoral work yields differential wages across occupations and industries. However, there remain important gaps in documenting that the observed relationships are really the result of sorting, heterogeneous moral preferences and compensating differentials. Most obviously, correlational evidence like this might result from other unobserved worker and job characteristics. For instance, the wage premium for prostitution is typically attributed to a compensating differential; but it is unclear whether the compensation is for the perceived immorality or for other aversive aspects of the

---

<sup>5</sup> Our study also relates to research on effort and sorting by mission-oriented types (Besley and Ghatak, 2005; Delfgaauw and Dur, 2008; Ariely, Bracha and Meier, 2009; Fehrler and Kosfeld, 2014; Tonin and Vlassopoulos, 2015; Carpenter and Gong, 2016; Cassar and Meier, 2018; Dur and van Lent, 2019), though this line of research typically focuses on worker motivation and effort within firms and not on worker morality and labor demand in competitive labor markets.

job (Edlund and Korn, 2002; Gertler, Shah and Bertozzi, 2005).<sup>6</sup> Moreover, such studies fail to measure a critical element—the identification of workers’ heterogeneous concerns for morality, which is a fundamental part of our study—as a key driver of the relationship.

### 3. Theoretical framework

Our study is guided by hypotheses obtained from a simple model of labor markets with varying degrees of perceived job immorality and workers’ heterogeneous concern for avoiding immoral behavior.<sup>7</sup> We present the detailed model and analysis in Appendix C; in this section, we focus only on the model’s key predictions.

We examine a labor market for a job,  $j \in J$ , which might involve doing work perceived as immoral. The perceived immorality of  $j$  is measured by a function  $I: J \rightarrow [0, \infty)$ , where  $I(j') > I(j)$  means that job  $j'$  is *more immoral* than job  $j$ . Firms decide whether to hire a worker to do  $j$  at the market wage,  $w$ , and workers decide whether to accept work  $j$  for the market wage. Workers differ in degree to which they are concerned with acting morally,  $\theta_i \geq 0$ . A worker of type  $\theta_i$  accepts job  $j$  if the utility from doing so is higher than that of an outside option, or

$$u_i^{accept}(j, w) = w - c - \theta_i * I(j) \geq \underline{u},$$

where  $c \geq 0$  is the worker’s cost of effort and  $\underline{u} \geq 0$  is the workers’ reservation utility.<sup>8</sup>

Our first two results derive the primary hypotheses that we test across our analysis. The first result (Proposition 1) shows that there is an immorality premium for immoral jobs: an increase in the immorality of a job,  $I(j)$ , decreases labor supply and therefore increases the equilibrium wage. However, this wage premium will be insignificant if workers care little about morality (Corollary). Our second main result (Proposition 2) states that those individuals least concerned with avoiding immoral work—i.e., those with low  $\theta_i$ —sort into accepting immoral jobs, while those more concerned with morality refuse to do the job for the

---

<sup>6</sup> Related work in finance (Hong and Kacperczyk, 2009; Colonnello, Curatola and Gioffré, 2019) demonstrates that investing in firms that engage in immoral activities (“sin stocks”) yields higher returns. However, other industry and firm characteristics, such as litigation risk, may also differ between these types of investments (Blitz and Fabozzi, 2017).

<sup>7</sup> Our framework is a simplification of the theoretical literature on compensating wage differentials (see, e.g., Rosen, 1986). We do not seek to expand this literature, but rather to apply it to a context in which the relevant job dimension is immorality.

<sup>8</sup> Models about “mission-oriented” employees commonly assume very similar additive utility functions (e.g. Cassar and Meier, 2018). We assume that workers are not concerned about the consequential harm produced by immoral work, but rather have an aversion to personally implementing immoral work. For example,  $\theta_i$  can be understood as a reduced-form representation of social image concerns in a signaling model (Bénabou and Tirole, 2006; DellaVigna, List, Malmendier and Rao, 2016) or as impure altruism (Andreoni, 1989).

equilibrium wage. That is, wage premiums arise precisely because those who find immoral work most distasteful opt out of such jobs, which also has implications for the resulting composition of the workforce in industries and firms performing immoral work. This important relationship has, to our knowledge, not been previously empirically tested.

Our next two results are less central, but nevertheless provide useful testable insights into outcomes in labor markets for immoral work. First, we show that the most immoral types profit from an increase in the immorality of work (Proposition 3). Second, more widespread concern for moral behavior—i.e., more people caring about avoiding immoral work—can have positive externalities for the most immoral types (Proposition 4). That is, if the distribution of types shifts toward greater concern for morality (higher  $\theta_i$ ), then the supply at any given wage decreases, increasing the equilibrium wage and the utility of those least concerned with morality.

#### **4. Evidence of an immorality premium in the Swiss labor market**

Our analysis first explores whether individuals working in industries generally perceived as immoral obtain wage premiums. We use data from the Swiss Labor Force Survey (SLSF), a representative sample of Swiss workers compiled by the Swiss Federal Statistical Office, to study the relationship between the perceived immorality of work and the portion of wages that cannot be explained by observable worker and industry characteristics.

We initially identified industries that we jointly perceived as involving work activities likely to be widely seen as immoral; we did so before looking at any data from these industries, including wages.<sup>9</sup> This yielded six “immoral” industries: gambling and betting activities, monetary intermediations, credit granting, manufacture of tobacco products, wholesale of tobacco products and manufacture of weapons and ammunition. These include the industries regularly classified as “sin industries” in financial research (e.g., Hong and Kacperczyk, 2009; Blitz and Fabozzi, 2017; Colonnello, Curatola and Gioffré, 2019). We then chose comparison industries from within the same industrial branch with similar distributions of education levels, as well as nine additional industries representing large shares of employment in Switzerland.<sup>10</sup> While we did not look at wages when selecting any of these

---

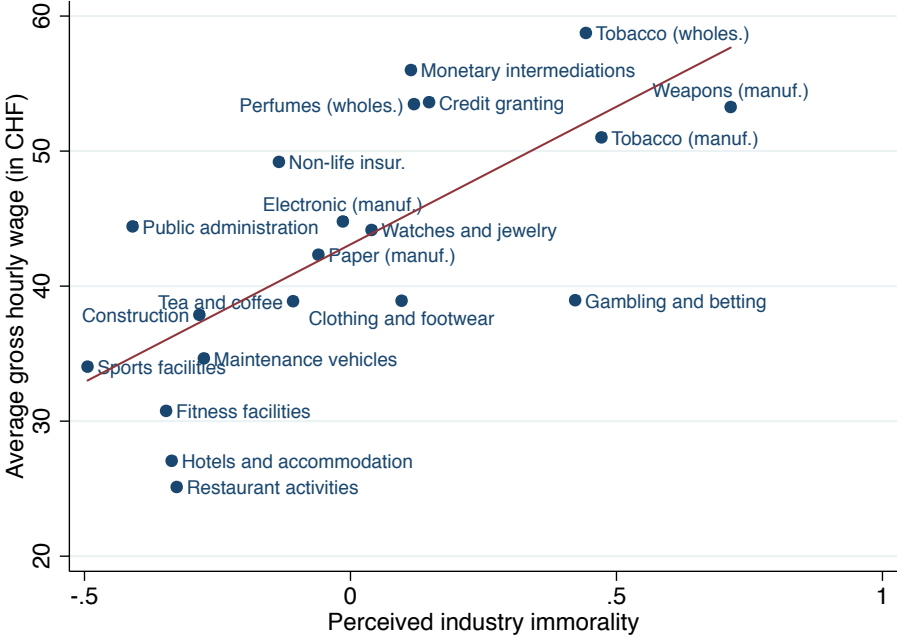
<sup>9</sup> Specifically, we started with the complete list of industries listed in the SLSF. Each of the authors went through the list and indicated any industries that he or she believed was widely perceived to have a significant immoral component. We selected those industries for which all three authors agreed. We proceed this way, rather than using the entire set of Swiss industries, to keep the number of questions we ask in subsequent surveys manageable.

<sup>10</sup> We chose five comparison industries: non-life insurance (for monetary intermediations; credit granting), organization and operation of sport facilities (for gambling and betting activities), processing of tea and coffee



industries, a natural concern is that choosing the sample of industries ourselves possibly (unconsciously) biases the sample toward those likely to confirm our hypothesis. In Appendix B we provide evidence for an immorality premium that does not rely on our sample of industries. Specifically, we first show that “immoral industries” (based on survey respondent’s ratings) are also estimated to pay a wage premium when we add all other industries as control industries. Second, we replicate the entire analysis with another set of industries selected by research assistants who were not familiar with the research question.<sup>11</sup>

**Figure 1: Correlation between wages and perceived industry immorality**



Source: Weighted data from the SLFS, years 2010-2016 (wage) and our own survey (perceived industry immorality). Notes: Perceived immorality is in [-1, 1] where -1 means very moral, 0 means neutral and 1 means very immoral. Real gross hourly wage in 2010 CHF (1 CHF ≈ 1 US Dollar). N = 32,638.

We next obtained independent ratings of the perceived (im)morality of the selected industries. We asked a sample of 177 students on the campus of the University of Zurich and the Federal Institute of Technology (ETH) to rate each industry on a 5-point Likert scale ranging from “very moral” to “very immoral,” re-scaled the responses to lie on the -1 to 1 interval and averaged the responses. (These survey data were collected as part of our survey studies, which we describe in more detail in Section 5.) We interpret this variable as a

---

(for manufacture of tobacco products), manufacture of electronic components (for manufacture of weapons and ammunitions), wholesale of perfume and cosmetics (for wholesale of tobacco products). These twenty industries make up a substantial share of the Swiss labor market: they employ 20.6% of the Swiss labor force (STATENT, 2016).

<sup>11</sup> We thank Uri Gneezy for suggesting this approach.

measure of the perceived immorality of working in industry  $j$ , or  $I(j)$ , a key component of our theoretical analysis. The mean ratings for each industry are shown on the horizontal axis of Figure 1. They confirm that our initial judgments with respect to the perceived immorality of certain industries are shared by the survey respondents.

The vertical axis of Figure 1 plots the mean real gross hourly wage (in 2010 Swiss Francs) in each industry. These data are the reported hourly wages of employees surveyed in the SLFS. We use data from the 2010 to 2016 waves. The strong positive relationship supports the hypothesis that work in industries perceived as less moral yields a wage premium. This result is not due to our choice of the industry sample—Figure B1 in the Appendix shows similar correlations in a set of industries independently selected by research assistants blind to the research question.

Of course, the relationship in Figure 1 ignores the potential role of individual worker characteristics, which may vary across industries, and other characteristics of the industries that may account for the wage gaps. To partially address this concern, Table 2 reports regressions of the hourly wage reported by individuals in different industries on the perceived immorality of each industry, along with several additional control variables. Model 1 displays the results of a simple regression of the natural logarithm of real gross hourly wages on the perceived industry's immorality, supporting the positive relationship in Figure 1. Model 2 adds observable worker and industry characteristics, while Model 3 additionally includes indicator variables for each year and indicator variables for the region where the employer is located.<sup>12</sup> The immorality premium remains large and statistically significant: according to Model 3, individuals working in an industry as immoral as manufacture of tobacco products (i.e., Perceived immorality = 0.47) have (geometric) mean hourly earnings approximately 35 percent higher than people working in a neutral industry (i.e., Perceived immorality = 0).<sup>13</sup> Appendix Table B1 shows a statistically significant wage premium using alternative approaches for classifying immoral industries, including firms selected independently by research assistants.

---

<sup>12</sup> Since the number of clusters is relatively small, we also computed significance levels using the wild bootstrap (Cameron, Gelbach and Miller, 2008) with 400 replications; p-values for industry immorality are 0.070 for Model 1, 0.000 for Model 2, and 0.005 for Model 3. Note also that numbers of observations differ substantially between industries. If we weight observations by industry size (instead of using survey weights), estimates for perceived immorality are smaller (Model 1: 0.489, Model 2: 0.349, Model 3: 0.344), but still significant ( $t=3.88$ , 3.10, and 4.03, respectively). Appendix Table A1 provides summary statistics of all variables.

<sup>13</sup> We obtain this number by doing the following calculation:  $e^{0.638*0.47} - 1 \approx 0.350$ .

**Table 2: Relationship between wages and perceived industry immorality**

Dependent variable: ln of real gross hourly wage (in 2010 CHF)			
	(1)	(2)	(3)
Perceived industry immorality	0.929*** (2.95)	0.736*** (4.15)	0.638*** (3.85)
Age		0.005*** (3.32)	0.005*** (3.64)
Male		0.204*** (6.60)	0.203*** (6.80)
Married		0.033* (1.88)	0.037** (2.13)
Education high		0.469*** (8.38)	0.442*** (8.64)
Education middle		0.177*** (4.53)	0.170*** (4.77)
Swiss		0.028 (1.15)	0.036* (1.97)
Experience		0.005** (2.81)	0.005*** (3.17)
Full-time equivalent		-0.037 (-0.51)	-0.038 (-0.54)
Managerial duties		0.059 (0.71)	0.065 (0.82)
Industry sales		0.027 (0.94)	0.034 (1.25)
Industry size (employees)		0.001*** (3.77)	0.001*** (3.69)
Constant	3.759*** (70.71)	2.930*** (26.43)	
N	32,638	32,638	32,638
Adjusted R <sup>2</sup>	0.140	0.379	0.397
Year FE	No	No	Yes
Region FE	No	No	Yes

Source: Weighed data from the SLFS, years 2010-2016 (wage and demographics), STATENT, years 2011-2016 (industry size, industry sales), Value Added Tax Statistics, years 2010-2016 (industry sales) and our own survey (perceived industry immorality).

Notes: Perceived immorality is in  $[-1, 1]$  where  $-1$  means very moral,  $0$  means neutral and  $1$  means very immoral. Control variables: Male in  $\{0, 1\}$ , Married in  $\{0, 1\}$ , Education high: higher vocational education and training or university/college, Education middle: apprenticeship, full-time vocational school, matura or pedagogical training, Education low (reference category): compulsory schooling or pre-vocational education, Swiss in  $\{0, 1\}$ , Experience = number of years in the firm, Full-time equivalent = (working hours /42), set to 1 for working hours  $\geq 42$ , managerial duties in  $\{0, 1\}$ , Industry size = number employees in this industry / 1000 (2010 data is not available, we substitute it with 2011 data), Industry sales = Industry sales/number employees in this industry. Model (3) controls for company region fixed effects (26 Swiss cantons) and year fixed effects (2010-2016). Standard errors clustered at the industry level, t-statistics in parentheses; \*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$ .

The above analysis is consistent with workers' concern with the perceived immorality of work requiring compensation. However, the correlational aspect of the relationship leaves open the possibility that additional unobserved industry characteristics may explain the relationship in Figure 1. Moreover, the analysis does not tell us whether the workers employed in these industries differ in their concerns for morality. In the following two sections, we explore the model's predictions more carefully—controlling for unobservable aspects of work and making a clearer connection to subjects' heterogeneous concerns for morality.

## **5. Sorting and wage premiums in a laboratory labor market**

Our experimental study investigates the preferences and behavior of subjects in the role of workers in a labor market. These subjects complete an online questionnaire, and then participate in a laboratory experiment approximately one week later. The questionnaire measures subjects' concern for morality using survey instruments, as well as their preferences regarding future employment possibilities. We discuss the online questionnaire in detail in Section 6, where we also report the analysis of the resulting data. In this section, we focus on the design and results of the laboratory experiment.

In the laboratory sessions, subjects perform two choice tasks. First, we elicit a measure of concern for morality ( $\theta$  in the theoretical framework in Section 3) using an incentivized behavioral task that creates a tradeoff between personal monetary gain and moral conduct. We use this measure, which closely parallels the moral nature of work in the subsequent market activity, to investigate whether participants with low concern for morality sort into immoral work. We also validate this measure using the more general measure of moral concern from our survey instruments (see Section 6).

Second, subjects participate in a laboratory labor market for 15 periods. In each period, they submit reservation wages for performing a task and are potentially hired based on their and other workers' wage requests and the demand for work. Labor demand is simulated according to a fixed demand schedule; that is, wage offers are made by computerized employers.

The key feature of our laboratory experiment is that we exogenously vary *only* the degree to which work involves doing something generally perceived as immoral—holding constant all other job characteristics, including the specific actions subjects take when employed—to investigate the causal impact on labor market outcomes. We attempt to design an “immoral” act that is unambiguously harmful and for which there is likely widespread

agreement regarding its immorality. We opt for an act akin to giving bad financial advice to a non-profit organization, like UNICEF, thereby harming the non-profit's financial standing and thus harming the organization's employees and its ability to help aid recipients. Providing harmful information and misleading customers is a realistic feature of many existing jobs perceived as immoral.

We operationalize this kind of scenario in our experiment by informing subjects that each session is endowed with an initial donation to a UNICEF fund that provides malaria treatments for children (the aid recipients).<sup>14</sup> However, the actual final donation for a session is influenced by the behavior of participants in the session. Specifically, subjects in our experiment are hired to provide written advice to a "client" (a subject who participates in a subsequent survey and serves a role analogous to an employee of the non-profit making financial decisions on its behalf). We vary, by treatment, whether workers are assigned to a market with *neutral* jobs that involve honest advice that has little impact on the client and the UNICEF fund or to a market with *immoral* jobs that involve dishonest advice that hurts the client and UNICEF. A worker's choice of whether to accept work is visible to other workers in the labor market.<sup>15</sup>

After the laboratory sessions, we recruited a separate sample of individuals at public locations. These participants serve two functions. First, they serve in the role of "clients" who receive written recommendations from laboratory subjects and act on this advice. From these choices, the clients accumulate money and determine the size of the UNICEF donation. Second, these participants complete a survey in which they evaluate the extent to which various industries and firms are "moral" or "immoral." These are the ratings that we already used in Figure 1 and Table 2.

## 5.1 Design

Laboratory sessions consisted of 24 participants. Before entering the lab, we took a portrait photograph of each subject to make labor market outcomes public. Participants were asked to make a neutral face while the picture was taken. As subjects could influence the amount of a donation to a UNICEF fund providing malaria treatment to children, participants read an

---

<sup>14</sup> To strengthen the moral context and link the money to UNICEF to "donors," these initial donations are linked to a donation generated by a third party. Prior to the laboratory session, we approached individuals who had just donated blood as part of a donation campaign and asked whether they would agree that the University of Zurich potentially make a donation to UNICEF as a complement to their blood donation. Most donors we approached agreed. We did not otherwise obtain any information or choices from these donors.

<sup>15</sup> Our behavioral measures of aversion to immoral acts confounds both internally-driven concerns for acting immorally and concerns for being perceived as willing to act immorally. Since real-world labor markets typically also confound both motives, we do not draw a distinction in our study.

information sheet at the beginning of the experiment about the consequences of malaria and the need for treatments—we adopted wording from public UNICEF materials and referred to each donation unit as helping to “save a child” from malaria by providing a treatment.

In the following, we describe each of the choices subjects completed in the laboratory session, in detail. We also provide details on the recruitment and role of the clients.

### 5.1.1 Behavioral measure of concern for morality ( $\theta^{Exp}$ )

Participants first played an incentivized game that measures their willingness to lie for personal gain while causing harm to others in a non-market environment. The task builds on a game by Gneezy, Rockenbach and Serra-Garcia (2013), and modifies it such that it mimics the consequences of a lie in the immoral work treatment condition in our experimental labor market.

In the game, Participant A privately observes a computerized die roll and sends a message reporting the observed number to Participant B. Participant A may claim that the observed number  $r$  is either “1”, “2,” “3,” “4,” “5,” or “6,” regardless of the actual number. Participant A receives  $5+r$  CHF (1 CHF  $\approx$  1 US Dollar),<sup>16</sup> which means that she has an incentive to lie if  $r$  is less than 6. Participant B then decides whether “to follow” or “not to follow” the message sent by Participant A. If Participant B does not follow the message, he receives 1.5 CHF and the donations to UNICEF are unaffected. If he follows the message and Participant A truthfully reported the observed number, Participant B earns 5 CHF and the initial donation to UNICEF is increased by an amount corresponding to one additional anti-malarial treatment.<sup>17</sup> However, if Participant B follows the message and Participant A lied, Participant B does not earn any money and the donation *decreases* by one treatment.

Every participant initially plays the role of Participant A. We use the strategy method to elicit Participant A’s message for every possible die roll as this allows us to classify all subjects by their strategies. At the end of the experimental session, 5 of the 24 participants in the session have their role changed from Participant A to Participant B. These Participants B are then matched with five of the remaining Participants A and decide whether or not to follow the corresponding message. All participants whose role is not switched—who remain

---

<sup>16</sup> During the experiment, subjects accumulated earnings in “points,” which we converted to money at the rate of 20 points = 1 CHF. We present the design and results in terms of ultimate payments in Swiss francs (CHF) to provide a clearer indication of the economic significance.

<sup>17</sup> The actual cost of providing 30 malaria treatments for children, according to UNICEF, was CHF 29. In order to create small units with a strong tangible moral component, our instructions always referred to the amount of money corresponding to treating “one child” and did not specify the exact monetary amounts.

as Participant A—are paid based on their choice as Participant A, independently of whether or not they are matched with a Participant B.<sup>18</sup>

Participants were informed that, at the end of the session and after all choices had been made, their decisions as Participant A would be publicly displayed to other participants in the session, along with their portrait photograph. This was all explained clearly and publicly at the beginning of the experiment.

### *5.1.2 Labor market experiment*

In the labor market, subjects play the role of workers competing to be hired by automated firms. Before interacting in the market, subjects receive general instructions about the labor market and hear a recording of the instructions. Then, participants answer comprehension questions about the market, including how prices and quantities are determined. Only after the above instructions about market procedures, subjects receive information about the nature of the job, which varies by treatment. This ensures that subjects in both conditions interpret the market instructions similarly. Their understanding of these new instructions about the job is again tested through comprehension questions.

**The job.** In both treatment conditions, workers have the opportunity to be hired as an “advisor” whose job is to give advice to another participant outside the laboratory, the “client.” Specifically, the advisor has to write a recommendation to a client to select one among ten choice options (labeled by the letters “A” through “J”). The treatment varies which option the advisor must recommend to the client. The client receives this recommendation, and then selects one of the ten options. The client only knows that the option he selects determines his financial reward for completing a survey and influences a donation to UNICEF, but does not know the consequences of any specific option. However, the client knows that the advisor had complete payoff information at the time of writing the recommendation. The client is free to choose the recommended option or any other option.<sup>19</sup>

The payoffs associated with each of the ten options are indicated in Table 3. Nine options increase the client’s reward by 1 CHF ( $\approx$  1 US Dollar) and increase the donation to UNICEF by an amount corresponding to the treatment of one child. However, one of the 10 options—in this case, *D*—gives 0 CHF to the client, and *reduces* the UNICEF donation.

---

<sup>18</sup> This implies that Participant As (whose role was not switched) received their own payment with certainty. Their decision, however, only had consequences for Participants B (and UNICEF) with a probability of 26.3 percent. This corresponds, roughly, to the stochastic impacts in our experimental labor market.

<sup>19</sup> Clients followed the recommendation in 84% of all cases.

Our two conditions vary only in the recommendation that the advisor is hired to make. In *neutral work*, the advisor’s job is to recommend a specific option that is beneficial to the client and to UNICEF (e.g., option *G* in Table 3). Note that a client is very likely to make such a choice independently of any advice, thereby making the impact of such advice largely neutral. In *immoral work*, the job is to recommend the single option with negative consequences (option *D*). In both cases, the advisor makes a recommendation by completing a form stating that the recommended option “will save the highest number of children” and “will give you the highest financial reward.” By recommending option *D*, however, the advisor increases the chance that the client selects the single option that will not increase his earnings and that will reduce the donation to UNICEF. We vary the letter of the recommended bad (neutral) option across immoral (neutral) laboratory sessions. Note that conditions only differ in the moral nature of the job; everything else, including effort costs, is kept constant.

**Table 3: Options available to the “client”**

	A	B	C	D	E	F	G	H	I	J
Additional number of children receiving the anti-malarial treatment	1	1	1	-1	1	1	1	1	1	1
Financial reward for client (CHF)	1	1	1	0	1	1	1	1	1	1

**The market.** Participants are randomly allocated to markets consisting of 6 workers who compete to be hired by 6 automated firms. Each worker can provide up to two units of labor—one at a low cost (CHF 2.50) and one at a high cost (CHF 5.50). All workers face the same induced costs, which the instructions clearly explain.

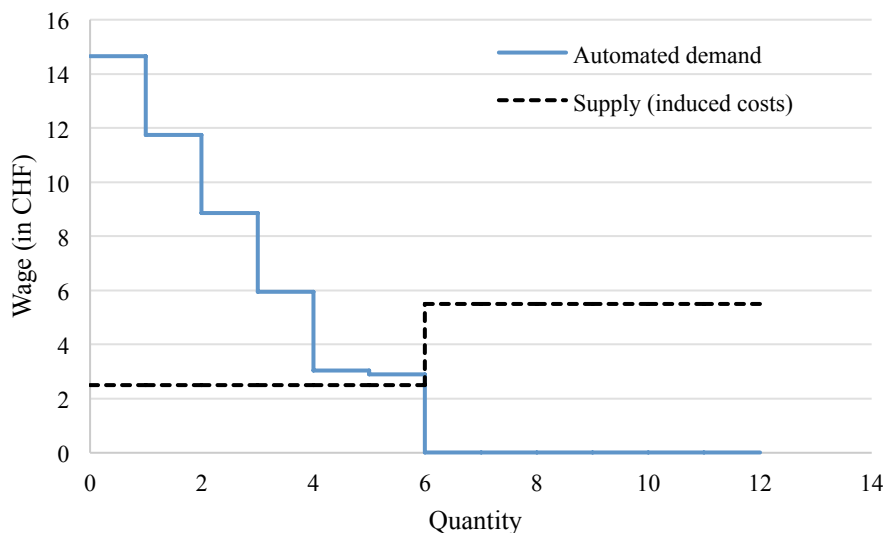
At the beginning of every market period, each worker decides whether or not to participate in the labor market. In she decides to participate, she then (privately) provides two wage requests, one for each of the possible units of labor she can provide. Workers may only submit wage requests that are at least as high as the corresponding cost of providing that job.

Firms are simulated by the computer. Each firm can hire up to one unit of labor per period. Firms are identical except for the wage that they offer to the workers. Figure 2 displays the automated demand for labor as well as the induced costs of labor supply. In



equilibrium, all workers provide one unit of labor and the market wage is between CHF 2.5 and 2.9.<sup>20</sup> The workers have no information about the shape of the automated demand.

**Figure 2: The automated demand and the induced costs of labor supply**



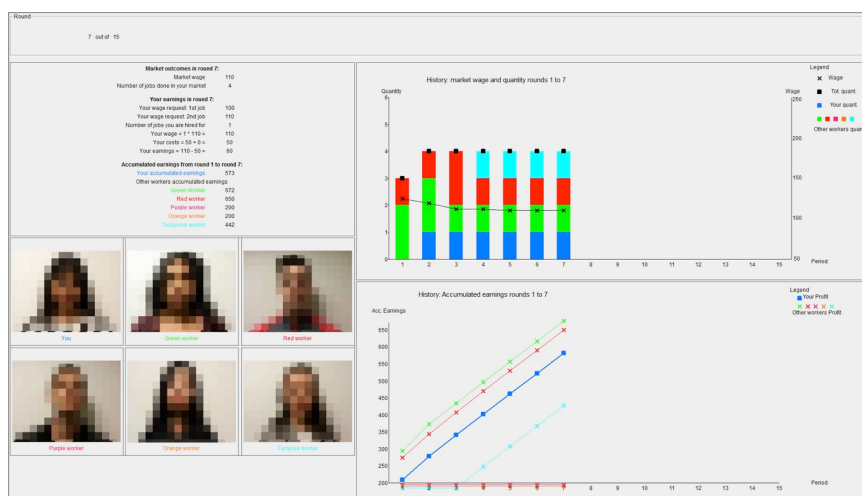
We use a uniform-price sealed-offer auction as the market mechanism, as this provides desirable features. First, Smith et al. (1982) show that this type of market typically converges to the equilibrium prediction. Second, and more importantly, this mechanism allows us to automate labor demand (see also Sausgruber and Tyran, 2011) and therefore keep demand constant between conditions. Once all six workers have submitted their wage requests, the computer ranks them from lowest to highest and compares the workers' wage requests to the firms' wage offers, ranked from highest to lowest. The *market wage* is then the lower of two potential candidates: (i) the last wage offer that is higher than the wage request with the same rank, and (ii) the first wage request that is higher than the wage offer with the same rank. This mechanism clears the market in that, for the market wage, labor supply equals labor demand and all workers with wage requests below the market wage are hired. A worker's earnings in a round equal the market wage times the units of labor provided by that worker (0, 1 or 2), minus the corresponding costs; workers who do not participate in the market do not make any money.

The market repeats for a total of 15 periods. The composition as well as the type (immoral or neutral) of each market is fixed across periods. At the end of each period, the computer reports the market wage, displays the picture of every worker in the market and summarizes information regarding each workers' outcomes across all periods (see Figure 3).

<sup>20</sup> We selected this specific labor demand function to facilitate equilibrium convergence. As long as the wage is higher than CHF 3.05, at least two workers will be unemployed, putting downward pressure on wages.

Specifically, subjects observe employment outcomes, wages and cumulative earnings for all workers in their market across periods, and can connect these to the other workers' identities through the photographs. After observing outcomes, those participants who are hired in a period complete the paper forms with the recommendations—they write their own initials and the appropriate letter (e.g., “G” or “D” in the earlier example).<sup>21</sup> If a firm does not succeed in hiring a worker in a period, the firm's client will not receive any recommendation. This implies that in the immoral condition, if a participant is not willing to do the job for the market wage, the number of clients who receive bad recommendations (weakly) decreases.

**Figure 3: Example of feedback provided after every period**



### 5.1.3 Procedural details

All sessions took place at the Decision Sciences Laboratory (DeSciL) at the Federal Institute of Technology in Zurich (ETH) in February through May of 2017. Participants were recruited using hroot (Bock, Baetge and Nicklisch, 2014) from the joint subject pool of the University of Zurich and the ETH. Every session consisted of 24 participants, who were only accepted at the session if they had previously completed the online survey described in Section 6.<sup>22</sup> To start a session, subjects had to enter an identifier that allows us to link, anonymously, their

<sup>21</sup> Subjects are informed that each firm has a probability of 25 percent of having a client in every period, independently of whether or not the firm hires a worker. If the firm does not have a client, then the worker's recommendation will be unused, although the worker still completes the recommendation and receives the market wage. However, subjects do not know at the time of submitting wage requests or completing the forms at the end of a period whether or not there will be a client for that period. This represents, for instance, a case in which a worker is hired to prepare promotional materials for a harmful product, which may or may not ultimately be seen by potential customers. At the end of the experiment, subjects learn which of their written recommendations will be distributed. We did this to lower the number of clients we needed to recruit in the follow-up survey. This procedure implies that writing a recommendation only has consequences with a probability of 25 percent and, therefore, works against our hypothesized treatment effect.

<sup>22</sup> We made an exception if less than 24 subjects who completed the survey showed up to the experiment. In total, three subjects were allowed to participate despite not completing the online survey.

answers in the online survey with their behavior in the lab. We conducted ten sessions, resulting in a total of 240 participants, allocated to 28 markets for immoral work and 12 markets for neutral work. The laboratory experiment was implemented with zTree (Fischbacher, 2007). Our study obtained ethical approval from the Human Subjects Committee of the Faculty of Economics, Business Administration and Information Technology at University of Zurich.

All instructions were delivered both on paper and with pre-recorded audio files. Instructions and materials are available at <https://osf.io/ym657/>.<sup>23</sup>

#### 5.1.4 Survey study with “clients”

We subsequently recruited a different sample of students on the campus of the University of Zurich and the Swiss Federal Institute of Technology (ETH) (N=177). We invited student passersby to participate in a brief choice experiment in which they could earn money and generate a donation for UNICEF aimed at providing treatments for children infected by malaria. They were told that they would earn CHF 2 plus possibly some additional money for a 5-minute study. These subjects performed two functions.

First, they served the role of “clients” for the recommendations from the laboratory labor market. Each participant made up to six decisions by choosing one of the ten letters between A and J. They knew that these decisions influenced their own earnings and also possibly the amount of a donation to UNICEF, but they did not know the actual mapping from choices to payoffs. Each decision had the payoff structure in Table 3, but we varied which letter corresponded to the bad option. Clients received a mixture of recommendations with good advice, bad advice and no advice (corresponding to cases in which a firm was not able to hire a worker). Clients were only informed of the total payoffs at the end of their decisions.

Second, while their payment was determined and prepared, participants completed a survey in which they rated various firms and industries on a scale from 1 (*very immoral*) to 5 (*very moral*). For firms, clients also had the option to choose “I don’t know this organization.” The complete list of firms and industries is available in Appendix Tables A1 and A2. Instructions are available at [osf.io/ym657/](https://osf.io/ym657/).

---

<sup>23</sup> At the conclusion of the laboratory session, we also measured participants’ affect levels (Bradley and Lang, 1994), beliefs about whether the clients would or would not follow recommendations (not incentivized) and concerns for social image using the public self-consciousness scale by Leary et al. (2015).

## 5.2 Results

In our analysis, we focus on testing for the existence of an immorality premium and sorting by heterogeneous moral types. We first discuss how we construct our incentivized measure of concern for morality,  $\theta^{Exp}$ . Next, we study *behavior* in the labor market, and whether it can be predicted by  $\theta^{Exp}$ . We then study *outcomes* in the labor market and their connection to the behavioral measure of moral concern ( $\theta^{Exp}$ ).

### 5.2.1 Construction of $\theta^{Exp}$

We construct  $\theta^{Exp}$  based on choices in the behavioral task that subjects completed at the beginning of the laboratory session (Appendix Table A3 shows the distribution of choices). Let  $m_{ir}$  be the number that individual  $i$  reports if the actual die roll is  $r$ . We classify an individual as a  $\theta_L$  type if  $m_{ir} \geq r$  for all  $r \in \{1, 2, \dots, 6\}$  and  $m_{ir} > r$  for at least one  $r$ ; that is, participant  $i$  is classified as having a low concern for acting morally if he or she lies at least once for personal gain and never in a self-harmful manner. We classify the remaining participants as  $\theta_H$  types.<sup>24</sup> Based on this classification, we have 66 (27.5 percent)  $\theta_L$  types and 174 (72.5 percent)  $\theta_H$  types.<sup>25</sup> For convenience, given our interpretation of  $\theta$  we will often refer to  $\theta_L$  types as *immoral* and  $\theta_H$  types as *moral*. We next explore the differential behavior of the different types in the labor market and the consequences of this behavior.

### 5.2.2 Labor supply of moral and immoral types

Assuming that  $\theta^{Exp}$  measures a stable concern for morality that translates into labor-market choices, we should observe differential labor-market behavior between  $\theta_H$  and  $\theta_L$  types, but only when employment requires immoral work. The data confirm this. In Table 4, we report the results of a double-hurdle regression of the decision of whether to submit a wage request and, conditionally, the actual wage request. The key independent variable is a subject's type from the behavioral task at the beginning of the experiment. In the immoral work condition,  $\theta_H$  types opted to submit wage requests less frequently, by almost 30 percentage points, than  $\theta_L$  types (61.6 percent vs. 90.6 percent,  $p < 0.001$ ). By declining to submit a wage request, a

---

<sup>24</sup> A total of 13 subjects (5.4 percent, see Appendix Table A3) harmed themselves at least once with a lie ( $m_{ir} < r$ , e.g., reporting  $m_{ir} = 1$  when  $r = 2$ ). Since these subjects do not appear to be motivated by egoism, we classify them as  $\theta_H$ . The remaining 161 subjects classified as  $\theta_H$  always report the true number. Classifying subjects that lied in a self-harmful manner as  $\theta_H$  types is conservative in that they act less morally than the honest subjects (see Appendix Table A5). Results do not change if we drop these subjects or if we classify them instead as  $\theta_L$ .

<sup>25</sup> We find similar results if we use alternative classification approaches (see Appendix Table A5). For example, in principle, we could classify subjects into more than two categories—e.g., conditional on the number of lies or based on the expected payoff from lying ( $\frac{1}{6} \sum_{r=1}^6 m_{ir} - r$ ). Due to the low number of subjects with different lying-patterns (see Appendix Table A3), we opt for a binary classification for our main analysis.

subject indicates an unwillingness to do the work even at a wage of 50 CHF ( $\approx$  50 USD) per period, the highest possible wage request in our experiment. Furthermore, consistent with the model,  $\theta_L$  types submit conditional reservation wage requests that are approximately 0.49 CHF lower than those of  $\theta_H$  types ( $p=0.073$ ). These effects do not become weaker over time—if anything, the greater willingness of  $\theta_L$  types to participate in the immoral labor market becomes slightly stronger over time.<sup>26</sup> Moreover, 21.5 percent of the  $\theta_H$  workers never participated in the market (i.e., refused to submit wage requests in any of the 15 periods), but this is true of only 4.3 percent of  $\theta_L$  workers ( $t = -3.78$ ;  $p=0.001$ ). Hence, our behavioral measure of a subject’s moral type predicts their willingness to seek employment in an immoral job.

**Table 4: Relationship participation decision/reservation wage and  $\theta^{Exp}$**

Dependent variable:	Participation	Reservation wage	Reservation wage
	(1)	(2)	(3)
Low-theta ( $\theta_L^{Exp}$ )	1.024*** (4.64)	-0.494* (-1.79)	-0.018 (-0.22)
Constant	0.295** (2.41)	4.056*** (20.20)	2.909*** (43.79)
Sigma		2.64*** (7.72)	0.609*** (10.57)
Condition	Immoral work	Immoral work	Neutral work
N	2520	1755	1077
LL (pseudo)	-1427.9	-4194.1	-993.9

*Notes: Estimates from Craggs double-hurdle Model: (1) is a probit model; (2) and (3) are truncated linear regressions (truncated from above at 50 CHF). Models (1) and (2) use only data from the immoral work condition; model (3) uses only data from the neutral work condition. For neutral work, we focus on reservation wages as we have only 3 incidences in which a subject did not participate. Independent variables: Low-theta in {0, 1}. Standard errors clustered at market level; z-statistics in parentheses; \* -  $p < 0.1$ ; \*\* -  $p < 0.05$ ; \*\*\* -  $p < 0.01$ .*

In the neutral work condition, however, non-participation is virtually non-existent—there were only 3 cases in total, representing 0.28 percent of all observations. That is, there is nearly universal participation when the job does not involve immoral behavior. Moreover, the average wage requests of  $\theta_H$  (CHF 2.91) and  $\theta_L$  (CHF 2.89) types do not differ in magnitude or in statistical significance ( $p$ -value from truncated linear regression = 0.827).

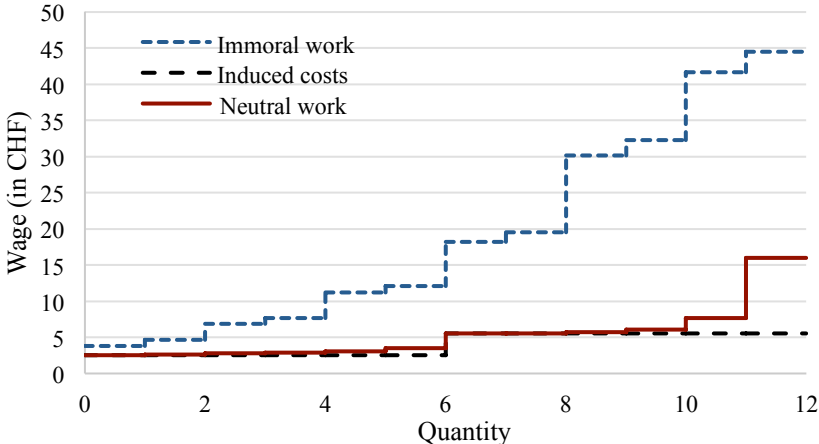
Thus, we observe that  $\theta_H$  types withdraw their participation and make higher wage requests when work requires immoral behavior. However, when the work activity is neutral, both types almost always participate and make similar wage requests. As a direct consequence of these observations, labor supply differs substantially between the two kinds of markets, as

<sup>26</sup> Specifically, if we add a linear time trend to the hurdle model and its interaction with  $\theta_L^{Exp}$  (see Appendix Table A4), we find that  $\theta_L$  types become slightly more likely to participate over time and provide lower reservation wages, relative to  $\theta_H$  types. However, both coefficients are small and statistically insignificant.

shown in Figure 4.<sup>27</sup> For neutral work, labor supply is fairly close to the induced costs. However, for any given wage, there is a substantially lower supply of labor in the condition requiring immoral work.

In the following sections, we explore the implications of the above heterogeneous behavior for labor market outcomes.

**Figure 4: Labor supply for neutral and immoral work in the laboratory**



Notes: Wage requests are ranked within each market period of each group. The figure shows the average wage request for each rank for both the immoral work and the neutral work conditions. Note that wage requests are censored at the maximal wage request that subjects could make, 50 CHF. For this figure, we set the wage requests of subjects who are not willing to participate to CHF 50. Therefore, the supply curve for the immoral work condition should be interpreted as a lower bound.

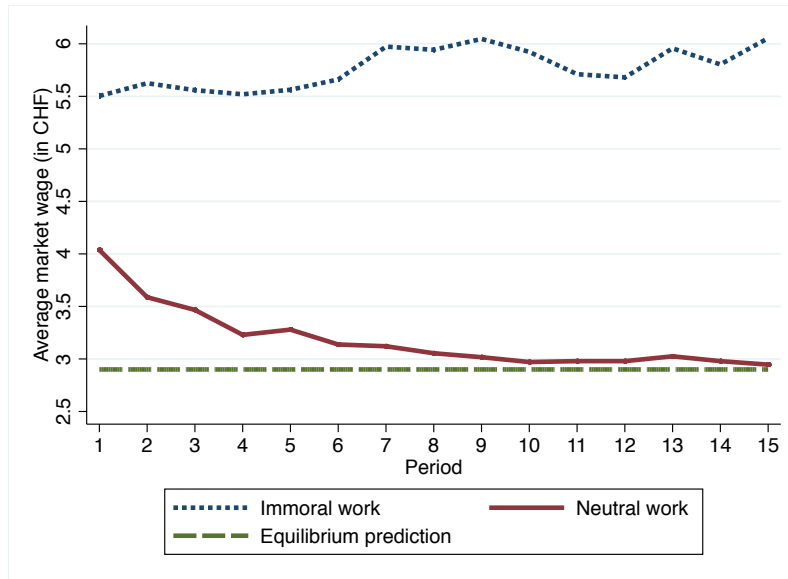
5.2.3 Wage premium for immoral work

We find a substantial immorality premium—between roughly CHF 1.5 and CHF 3 per period—as shown in Figure 5. This follows directly from the differential labor supply in Figure 4. While market wages in the neutral work condition converge toward the equilibrium prediction of CHF 2.90, the average market wage is persistently higher for immoral work and this difference is statistically significant in a t-test from a regression with standard errors clustered at the market-level (coefficient=2.581, t=6.00, p<0.001). Moreover, the difference grows across periods, suggesting that it is robust to repeated market competition and experience. Hence, our laboratory labor market shows causal evidence of a substantial and persistent wage premium for immoral work.<sup>28</sup>

<sup>27</sup> Appendix Figure A1 shows the labor supply if we only consider the last 5 periods. Appendix Figure A2 displays the labor supply in (simulated) labor markets with only low-theta or only high-theta types.

<sup>28</sup> A natural concern might be that the immorality premium results from the specific labor demand structure that we employ. However, as evident from the differences in labor supply in Figure 4, workers are sufficiently concerned with acting morally that wage premiums would obtain under a variety of demand specifications.

**Figure 5: Immorality wage premium in laboratory labor markets**



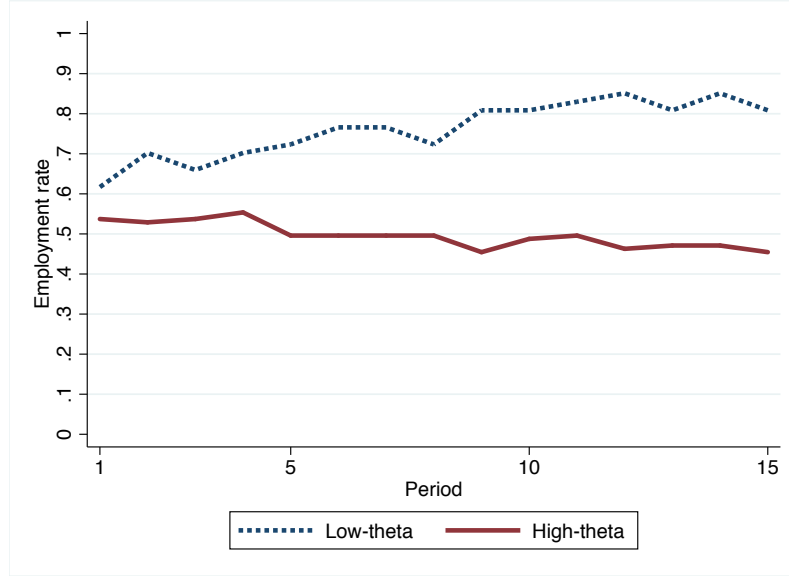
We also find persistent differences in employment levels in the two markets (Figure A3 in the Appendix). While markets for neutral work converge to the equilibrium prediction of 6, the average market quantity remains below 4 in the immoral work condition. This difference is significant in a t-test comparing the means (coefficient=-1.201,  $t=-6.30$ ,  $p<0.001$ ). Moreover, the trends in Figure 5 provide further evidence that the manifestation of  $\theta_H$  types' morality in labor market behavior does not erode over time. This persistence is remarkable given the high amount of social information provided at the end of each market period: in the immoral work condition, participants who forgo working see other participants less concerned with morality repeatedly earning high wages due to their own reluctance to act immorally.

#### 5.2.4 Sorting in immoral labor markets

We next turn to our second main prediction regarding sorting by different types into immoral work. Figure 6 shows that, indeed,  $\theta_H$  types are consistently employed less frequently in the immoral work condition. Table 5 shows that, on average,  $\theta_L$  types are 26.6 (=26.8 – 0.2) percentage points more likely to be employed than  $\theta_H$  types (column 1). This difference is statistically significant ( $p<0.001$ ) and robust to adding market fixed effects (column 2). The results are similar if we use, as a dependent variable, the number of work units provided (0, 1 or 2) rather than a binary measure of employment (columns 3 and 4). In the neutral work condition, we do not find a significant difference in employment rates between the two types (columns 1 to 4). This corroborates that the difference in hiring rates in the immoral work

condition is driven by differences in concerns for morality and not some other difference between  $\theta_H$  and  $\theta_L$  types.

**Figure 6: Employment rate by type in the immoral work condition**



**Table 5: Relationship between  $\theta^{Exp}$  and outcomes in the experimental labor markets**

Dependent variable:	Employment rate		Number of work units		Market income	
	(1)	(2)	(3)	(4)	(5)	(6)
<b>Low-theta (<math>\theta_L^{Exp}</math>)</b>	-0.002 (-0.05)	-0.034 (-0.90)	-0.002 (-0.05)	-0.034 (-0.90)	-1.33** (-2.40)	-0.56 (-1.39)
<b>Immoral work (IW)</b>	-0.331*** (-7.70)		-0.272*** (-7.11)		12.15*** (9.96)	
<b><math>\theta_L^{Exp} * IW</math></b>	0.268*** (4.54)	0.249*** (3.58)	0.256*** (4.10)	0.264*** (3.57)	8.42*** (2.71)	11.37*** (3.64)
<b>N</b>	240	240	240	240	240	240
<b>R<sup>2</sup></b>	0.179	0.319	0.103	0.174	0.138	0.243
<b>p-value: <math>\theta_L^{Exp} + \theta_L^{Exp} * IW = 0</math></b>	0.0000	0.0007	0.0000	0.0009	0.026	0.001
<b>Market FE</b>	No	Yes	No	Yes	No	Yes

Notes: Coefficient estimates of linear regression models. Independent variables: Low-theta in  $\{0, 1\}$ , Immoral work in  $\{0, 1\}$ . Standard errors clustered at market level; t-statistics in parentheses; \* -  $p < 0.1$ ; \*\* -  $p < 0.05$ ; \*\*\* -  $p < 0.01$ .

### 5.2.5 Market income in immoral labor markets

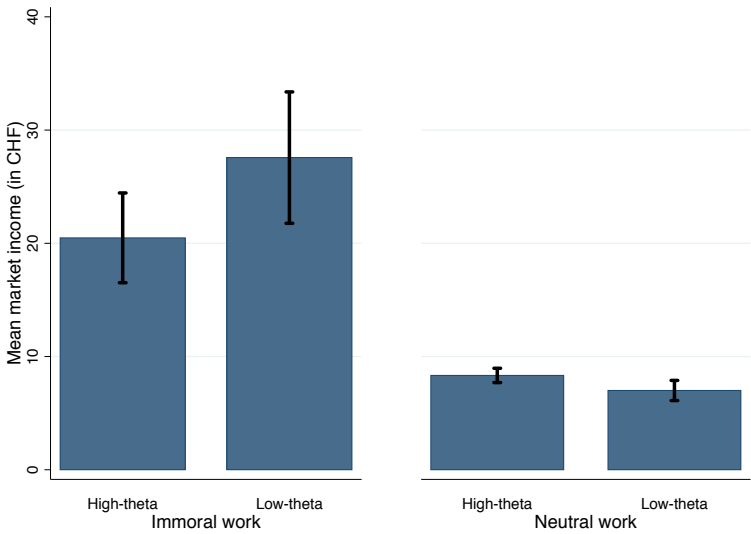
Two additional predictions address heterogeneous treatment effects on worker's utility. As a simple proxy for utility, we use the sum of all earnings accumulated by a worker over the 15 market periods.

First, our theoretical analysis predicts that the immoral types benefit from an increase in the job's immorality (Proposition 3). Figure 7 and results from linear regressions (see Table 5, Columns 5 and 6) show that the  $\theta_L$  types earn more market income than the  $\theta_H$  types in the



immoral work condition. The difference of CHF 7.09 (8.42 – 1.33 in column 5) is statistically significant ( $p=0.026$ ). Note that the potential market income is constrained by the market wage. If we control for the market wage by adding market fixed effects (column 6), the  $\theta_L$  types are estimated to earn CHF 10.81 more than the  $\theta_H$  types ( $p=0.001$ ). Thus, as predicted, immoral types earn considerably more in a market for immoral work. There is no such difference in markets for neutral work; if anything,  $\theta_L$  types earn slightly less than the  $\theta_H$  types.

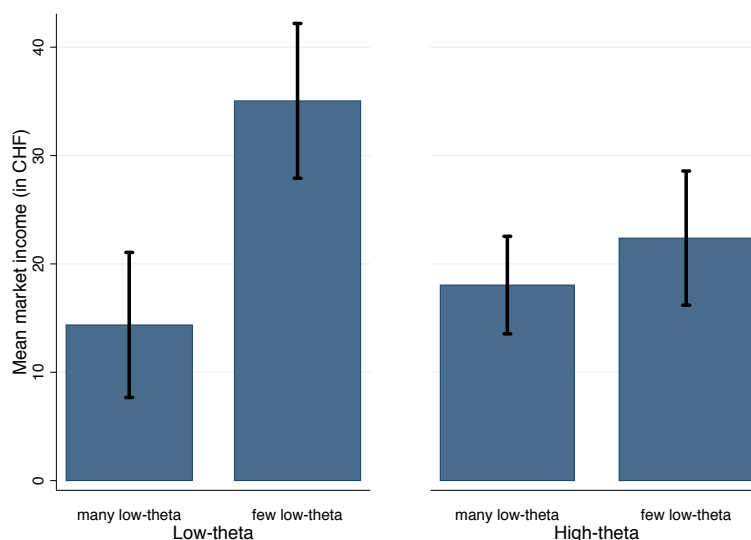
**Figure 7: Market income by moral type**



Notes: Average market income by treatment condition and moral type. Error bars show 95% confidence intervals.

Finally, our analysis predicts that, in the immoral work condition, immoral types have higher utility in the presence of more moral types (Proposition 4). To test this prediction, for each subject we count the number of other  $\theta_L$  workers in the market and split the sample based on the median of this measure. We thus classify a subject as being in a market with *few*  $\theta_L$  types if the number of other  $\theta_L$  workers is lower than 2, and as being in a market with *many*  $\theta_L$  types if the number is 2 or more. Using only the immoral work condition, this results in 70 subjects in the first category, and 98 in the second category. Figure 8 shows these subjects’ mean earnings, based on their own type and the composition of others’ types in their market. The income of  $\theta_H$  types is CHF 4.33 higher in a market with few  $\theta_L$  types than in one with many  $\theta_L$  types ( $t=2.05$ ,  $p=0.051$ , see Appendix Table A6). However, for  $\theta_L$  types, being in a market with few other  $\theta_L$  types increases income by an additional CHF 16.35 ( $t=3.53$ ,  $p=0.002$ ), resulting in a total difference of CHF 20.68 ( $t=4.29$ ,  $p<0.001$ ).

**Figure 8: Externalities of moral behavior for immoral types**



*Notes: Average market income by moral type and market composition; data only from immoral work condition. A subject is in a market with many low-theta workers if the number of (other) low-theta workers is 2 or more and in a market with few low-theta workers if the number of (other) low-theta workers is lower than 2. Error bars show the 95% confidence intervals.*

Our laboratory findings thus confirm all four predictions from our theoretical analysis. People with high concern for morality consistently refuse to do immoral jobs (or require a high wage for doing so), thereby decreasing labor supply and increasing wages and producing a persistent immorality premium. As a consequence, subjects with a low concern for morality are better off in markets for immoral work, particularly when in the presence of more moral subjects.

## 6. Stated real-world employment preferences and sorting

Several days (4-7) before the laboratory session, subjects completed an online survey. This survey includes questions designed to measure subjects' expectations of their own future labor market outcomes, including the willingness to work for different firms and industries and expected future wages, and questions eliciting subjects' concern for morality. The survey gives us a second, easily scalable, measure of moral concern ( $\theta$ ) and, more importantly, allows us to investigate whether subjects' moral types predict not only behavior in laboratory labor markets, but also preferences and expectations for labor market outcomes outside the laboratory.

We first construct an individual measure of concern for morality based on the answers to the psychological survey questions ( $\theta^{Sur}$ ). We then show that this second measure of

concern for morality correlates both with the comparable behavioral measure from the laboratory experiment ( $\theta^{Exp}$ ) and with outcomes in the laboratory labor market. This validation of  $\theta^{Sur}$  is useful for future research, as it is based solely on survey questions that are easier to collect than the incentivized laboratory measure. Moreover, the comparison of  $\theta^{Exp}$  and  $\theta^{Sur}$  provides some evidence on the stability of moral concerns across time and contexts, which is necessary for heterogeneous moral concerns to persistently influence labor market behavior. We then show that both  $\theta^{Sur}$  and  $\theta^{Exp}$  predict stated job preferences in labor markets outside the laboratory, consistent with the prediction regarding sorting (Proposition 2).

### 6.1 The online questionnaire

We first asked subjects several questions about their future labor-market expectations. Subjects were shown a list of 26 well-known companies in Switzerland and another list consisting of the 20 industries in Figure 1. Both lists are available in Tables A1 and A2 in the Appendix. Participants rated their willingness to work for each firm and industry (1: *not at all willing*; 5: *very much willing*). For firms, participants also had the option to select, “I don’t know this organization.” Note that these were the first questions participants answered, meaning that when they encountered them they had not been exposed to any references to morality or moral behavior.

Subjects next encountered several multi-item scales intended to measure an individual’s broad concern for morality and moral acts. These were:

- 1) **HEXACO-PI.** We administered 10 items from the short version of the HEXACO Personal Inventory (Ashton and Lee, 2009) related to the factor “Honesty-Humility”—consisting of the four traits, sincerity (3 items), fairness (3 items), greed avoidance (2 items) and modesty (2 items). Every item describes a thought that a moral or immoral person might have and participants indicate the extent to which each thought reflects their own opinions.
- 2) **Protected Values.** The Protected Values scale (Gibson et al., 2013) measures an individual’s position regarding values that can be seen as inviolable, and not substitutable against money, and that are usually central to the person’s identity. In our case, we adapted the Protected Values to a situation where a financial adviser can give bad investment advice to a client for personal benefit. First, 5 items assess the morality of this behavior (*Protected value 1*); second, 4 items examine how truthfulness matters in such a situation (*Protected value 2*).

- 3) Integrity and Work Ethics Test.** We used two items from an online test designed to allow firms to measure the integrity of job applicants (*Work ethics 1, Work ethics 2*). In each item, participants read fictitious dialogues between two characters with different opinions about a situation (e.g., calling in sick at work to enjoy a sunny day outside). Participants then rate with which character they have greater agreement.
- 4) Charity attitude index.** We used a 9-item scale developed by Brashear et al. (2000) in which participants rate statements regarding how important they perceive it is to help others in society and how positive and useful they perceive work done by charities.

In each case, subjects expressed agreement or disagreement with statements on either a 5-point or 7-point Likert scale. Thorough descriptions of these survey scales are provided in Online Appendix Table A7.<sup>29</sup> The online questionnaire was implemented with the Qualtrics software.

## 6.2 Constructing $\theta^{Sur}$

We next discuss how we construct a survey-based measure of concern for morality,  $\theta^{Sur}$ . Table 6 lists the 9 subscales from the morality measures (Appendix Table A7 provides summary statistics). To reduce multiple measures of (presumably) related individual characteristics to a single dimension, we employ factor analysis. We perform a principal-component factor analysis and select the factor with the highest eigenvalue (eigenvalue = 2.44) to represent  $\theta^{Sur}$ .<sup>30</sup> Column 1 in Table 6 presents the corresponding factor loadings. We normalized  $\theta^{Sur}$  such that it lies between 0 and 1; the resulting variable has a mean of 0.5 and a median of 0.5. Low values represent a low concern for morality. Appendix Figure A5 shows the distribution of  $\theta^{Sur}$ . Given our interpretation of  $\theta$ , we will often refer to subjects with a low  $\theta^{Sur}$  as *immoral types* and subjects with high  $\theta^{Sur}$  as *moral types*.

## 6.3 Does $\theta^{Sur}$ predict behavior in the laboratory?

To validate  $\theta^{Sur}$ , we investigate how it correlates with behavior, roughly one week later, in the laboratory. Table 6, column 2, shows the coefficients from independent simple regressions of a subject's type measured by the behavioral laboratory task,  $\theta^{Exp}$ , on each item comprising

---

<sup>29</sup> We also asked subjects to provide unstructured responses stating beliefs about their future career trajectories—specifically, what work they expected to do after their studies and how much they expected to earn at the age of 40—and whether several non-profit organizations (including UNICEF) are worth supporting. We also collected additional personal characteristics using a short version of the Big Five (Gosling et al., 2003) and several demographic characteristics.

<sup>30</sup> In the Appendix (Tables A8, A9 and A10), we show that our results are robust to different aggregation mechanisms. Specifically, we look at two alternatives: i) each of the nine survey measures is given equal weight and ii) the weight of the measures is determined by a regression of  $\theta^{Exp}$  on the survey measures.

$\theta^{Sur}$ . The dependent variable is binary, indicating that a subject is a  $\theta_H$  type according to  $\theta^{Exp}$ . The results show a significant positive correlation between  $\theta^{Exp}$  and all personality measures, except for *Work ethics 1* and *Work ethics 2*. Consistent with the positive relationship of the individual items, a regression of  $\theta^{Exp}$  on  $\theta^{Sur}$  shows a positive and significant relationship (coefficient=0.723,  $t=4.32$ ,  $p<0.001$ ); that is, a person who is characterized by a low concern for morality according to our survey-based measures is more likely to lie self-servingly in the behavioral measure in the experiment. These generally positive relationships suggest that  $\theta^{Exp}$  and our survey-based measures capture a stable individual characteristic.

**Table 6: Items comprising  $\theta^{Sur}$  and their relationship to  $\theta^{Exp}$**

	Factor loadings (weights for $\theta^{Sur}$ ) (1)	Regression coefficient of $\theta^{Exp}$ (2)
Protected value 1	0.664	0.540*** (3.66)
Protected value 2	0.708	0.352** (2.22)
Work ethics 1	0.213	-0.039 (-0.39)
Work ethics 2	0.252	0.042 (0.52)
HEXACO sincerity	0.482	0.362** (2.53)
HEXACO fairness	0.611	0.353*** (2.61)
HEXACO greed avoidance	0.477	0.225* (1.73)
HEXACO modesty	0.508	0.236* (1.79)
Charity attitude index	0.545	0.711*** (3.11)

Notes: Each subscale is constructed by taking averages over all items of the scale, and then normalized such that it lies between 0 and 1. (1): Factor loadings from principal-component factor analysis of survey measures on  $\theta^{Sur}$ . (2): Coefficient estimates of linear probability models.  $N = 237$  for each regression (3 subjects did not complete the online-survey and are excluded). Dependent variable: being a high-theta type according to  $\theta^{Exp}$ . Independent variables: survey measures in  $[0,1]$ , higher numbers indicate more morality. Robust standard errors;  $t$ -statistics in parentheses; \* -  $p < 0.1$ ; \*\* -  $p < 0.05$ ; \*\*\* -  $p < 0.01$ .

We next consider the extent to which  $\theta^{Sur}$  also predicts participants' behavior in the laboratory labor market, particularly in the immoral work condition.<sup>31</sup> Results from a linear regression of the employment rate on  $\theta^{Sur}$  indicate that those participants with the lowest concerns for morality (that is, participants with  $\theta^{Sur} = 0$ ) are 43.9 percentage points more

<sup>31</sup> Appendix Figure A6 displays the labor supply in a (simulated) labor market with only low  $\theta^{Sur}$  or only high  $\theta^{Sur}$  types.

likely to be hired in markets for immoral work than participants with the highest possible concern for morality (that is, with  $\theta^{Sur} = 1$ ). This difference is marginally statistically significant ( $p=0.057$ , see Appendix Table A8, column 1). A less noisy measure of subjects' market behavior is their actual choices. Results from a hurdle model indicate that those subjects with the lowest concerns for morality are 52.1 percentage points more likely to participate in markets for immoral work by submitting a wage request than individuals with the highest possible concerns ( $p=0.015$ , see Appendix Table A9). In the neutral work condition, as expected, we find no substantial differences in employment rates or in labor market behavior between the two types.

#### 6.4 Do $\theta^{Sur}$ and $\theta^{Exp}$ predict stated real-world labor market preferences?

Our study collects two measures of subjects' concern for morality ( $\theta^{Sur}$  and  $\theta^{Exp}$ ). The online survey also elicits the same subjects' willingness to work for several firms and industries, without any reference to morality. We also separately obtained independent ratings of the perceived immorality of these firms and industries, from the "clients" (see Section 5.1.4). In this section, we use all of this information to analyze how our measures of concern for morality connect to expectations about labor market outcomes outside the laboratory.

We create a measure of *perceived firm immorality* in the same way we created *perceived industry immorality*: by first scaling the ratings such that they lie between -1 (very moral) and +1 (very immoral), and then averaging them.<sup>32</sup> We use these variables as noisy measures of the immorality of work,  $I(j)$ , in industry (or, firm)  $j$ , a key component of the theoretical analysis motivating our research. The horizontal axes in Figures 9a and 9c plot the resulting normalized ratings for industries in our sample, the horizontal axes in Figures 9b and 9d plot the normalized ratings for firms (see also Appendix Tables A1 and A2).

Our focus in this section is to investigate whether these perceptions of industry and firm immorality interact with our subjects' measured concern for morality (either  $\theta^{Sur}$  or  $\theta^{Exp}$ ) to produce differential labor market preferences. For this purpose, we normalized subjects' stated willingness to work for firms and industries, such that they take values between 0 (*not at all willing*) and 1 (*very much willing*).<sup>33</sup>

---

<sup>32</sup> When clients rated firms, they also had the option to choose "I don't know this organization" instead of rating the firm. To calculate the perceived firm immorality, we exclude these observations. Alternatively, we could code these as neutral ratings. These two measures are highly correlated ( $\text{corr}=0.9854$ ). Our results do not change substantially if we use the alternative measure (see Appendix Table A11).

<sup>33</sup> Incidentally, the list of industries accidentally omitted five industries for five participants that participated in the first lab session, meaning we are missing these data. Other than these cases, all subjects completed the full questionnaire. We exclude all these missing observations from the analysis. When participants rated their

**Figure 9: Correlation between the difference in willingness to work between moral and immoral types and perceived immorality of industries/firms**



Source: Survey study (Perceived immorality), online survey (Willingness to work,  $\theta^{Sur}$ ), laboratory experiment ( $\theta^{Exp}$ )

Notes: Differences in willingness to work: Coefficient estimates of linear regression models of the participants' willingness to work for different industries (a and c) or firms (b and d) on  $\theta_H^{Exp}$  (a and b) or  $\theta_H^{Sur}$  (c and d).

Dependent variable: Willingness to work is in  $\{0, 0.25, 0.5, 0.75, 1\}$  where 0 means not at all willing to work, 0.5 means indifferent and 1 means really much willing to work. Observations where subjects did not know the firm ("I don't know this organization") or did not fill out the questionnaire are excluded. Independent variables: a and b use  $\theta^{Exp}$  to classify participants, where  $\theta_H^{Exp} = 0$  for low- theta types and  $\theta_H^{Exp} = 1$  for high-theta types, while c and d use  $\theta^{Sur}$  in  $[0, 1]$ . Perceived immorality is in  $[-1, 1]$  where -1 means very moral, 0 means neutral and 1 means very immoral.

The vertical axes of Figures 9a and 9c plot the difference in willingness to work for the industries between subjects who were classified as moral or immoral, according to  $\theta^{Exp}$  (Figure 9a) or  $\theta^{Sur}$  (Figure 9c). The strong negative relationship across all of these figures indicates that subjects classified as immoral using either our survey-based ( $\theta^{Sur}$ ) or

willingness to work for *firms*, they also had the option to choose "I don't know this organization." This option was chosen in 17.8 percent of all answers. We also exclude these observations. We obtain similar results if we classify such observations as "indifferent" or if we restrict our analysis to subjects that know all firms (see Appendix Table A12).

incentivized behavioral ( $\theta^{Exp}$ ) task are, on average, more willing to work for industries that others perceive as immoral.

Table 7, columns (1) to (4) provide statistical evidence of the relationships in Figures 9a and 9c. The dependent variable in the regressions is a subject's willingness to work for an industry, while the explanatory variables include the perceived industry immorality ( $I(j)$ , obtained from a separate group of respondents), the subject's concern for acting morally ( $\theta$ ) and the interaction of these two terms. While there is little evidence of a systematic difference in willingness to work for neutral industries between moral and immoral types, subjects' moral types have much stronger predictive power for their willingness to work in industries perceived as immoral. This pattern is significant at least at the 5%-level, holds for both measures of individual moral concerns,  $\theta^{Sur}$  and  $\theta^{Exp}$ , and is robust to controlling for subjects' gender, age, Swiss nationality, area of study, mean industry wages, industry size (number of employees), and industry sales.

**Table 7: Regressions of willingness to work for diverse industries and firms on perceived immorality and moral types**

Dependent variable:	Willingness to work for industry $j$				Willingness to work for firm $j$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Perceived immorality ( $I(j)$ )	-0.232*** (-3.99)	-0.226*** (-4.72)	-0.050 (-0.77)	-0.043 (-0.74)	-0.140** (-2.08)	-0.139** (-2.07)	0.114 (1.56)	0.110 (1.50)
Type from experiment ( $\theta_H^{Exp}$ )	-0.027 (-1.34)	-0.029 (-1.49)			-0.043 (-1.58)	-0.051** (-1.96)		
$\theta_H^{Exp} * I(j)$	-0.078*** (-2.56)	-0.078** (-2.51)			-0.154*** (-4.30)	-0.154*** (-4.38)		
Type from survey ( $\theta^{Sur}$ )			-0.101* (-1.71)	-0.107* (-1.72)			-0.173** (-2.28)	-0.211*** (-2.85)
$\theta^{Sur} * I(j)$			-0.479*** (-5.22)	-0.479*** (-5.24)			-0.731*** (-8.80)	-0.722*** (-8.53)
N	4715	4715	4715	4715	5064	5064	5064	5064
Control variables	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Coefficient estimates of linear regression models. Dependent variable: Willingness to work is in  $\{0, 0.25, 0.5, 0.75, 1\}$  where 0 means not at all willing to work, 0.5 means indifferent and 1 means really much willing to work. Observations where subjects did not know the firm ("I don't know this organization") or did not fill out the questionnaire are excluded. Independent variables: (1), (2), (5) and (6) use  $\theta_H^{Exp}$  to classify participants, where  $\theta_H^{Exp}=0$  for low- theta types and  $\theta_H^{Exp}=1$  for high-theta types, while (3), (4), (7) and (8) use  $\theta^{Sur}$  (in  $[0,1]$ ) instead. Perceived immorality is in  $[-1, 1]$  where -1 means very moral, 0 means neutral and 1 means very immoral. Control variables: age, gender, Swiss nationality, subject of study, average wage industry 2016 (SLFS; only for industries), industry size 2016 (STATENT; only for industries), industry sales 2015 (Value Added Tax Statistics; only for industries). Standard errors clustered at individual and industry/firm level (Cameron, Gelbach and Miller, 2011); z-statistics in parentheses; \*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$ .

We repeat the same analysis using data on subjects' willingness to work for our selection of well-known firms in Switzerland. The vertical axis of Figures 9b and 9d plot the difference in willingness to work for firms between subjects who were classified as moral and



immoral according to  $\theta^{Exp}$  (Figure 9b) or  $\theta^{Sur}$  (Figure 9d). Again, subjects classified as immoral are, on average, more willing to work for firms perceived as immoral. This relationship is confirmed by Table 7, columns (5) to (8): subjects less concerned with moral behavior are more willing to work for firms that other people rate as more immoral ( $p < 0.01$ ), which is again true for both measures of concern for morality ( $\theta^{Sur}$  and  $\theta^{Exp}$ ). Hence, we find that firms that are perceived as immoral more attractive to applicants with a lower concern for morality.<sup>34</sup>

The above analysis provides evidence, from outside the laboratory, supporting our second main prediction. Those who are least concerned with morality are significantly more willing to work in firms that are generally perceived as less moral. A limitation of this analysis is that it is based on hypothetical future choices. However, Wiswall and Zafar (2018) provide evidence that such stated preferences are predictive of ultimate employment. To further validate subjects' stated real-world labor market preferences, we can test whether the stated employment preferences correlate with individual employment rates in the immoral work condition of our laboratory experiment. Indeed, we find that people who are hired more often for the immoral job have a statistically significantly higher stated willingness to work in immoral industries outside the laboratory (see Appendix Table A13).<sup>35</sup>

## 7. Discussion and Conclusion

We investigate whether heterogeneity in individual preferences for avoiding immoral work and the perceived immorality of work influence the jobs that individuals select and individuals' earnings. Our study employs a laboratory experiment, surveys and administrative data to identify heterogeneity in concerns for morality and to create (or, measure) variation in the perceived immorality of different jobs. We use these different kinds of data to test two main hypotheses—first, that jobs generally perceived as immoral will yield a wage premium and, second, that individuals less concerned with moral behavior will be more likely to sort into such jobs.

---

<sup>34</sup> As we show in the Appendix, the differential willingness to work for immoral firms and industries by moral and immoral types does not depend on how we construct  $\theta^{Sur}$  (Table A10), nor on how we deal with missing observations (Table A11, Table A12).

<sup>35</sup> In the online survey, we also asked subjects to state how much they expect to earn when they reach the age of 40. A regression of expected future wages on  $\theta^{Sur}$  suggests a positive relationship such that the least moral types ( $\theta^{Sur} = 0$ ) report expected annual income that is 30,272 CHF higher, on average, than the expected income of the most moral types ( $\theta^{Sur} = 1$ ). This relationship is not statistically significant ( $p = 0.125$ ), which potentially reflects the noise in subjects' estimates of their earnings approximately 20 years in to the future.

In a laboratory setting, we find support for both our hypotheses in an environment where we cleanly vary only the immorality of work. Labor markets for immoral work yield significantly higher wages. Moreover, workers that we classify as “immoral” based on a separate behavioral task are significantly more likely to be hired for immoral work than are moral workers; but this relationship disappears in a labor market for neutral work. We also find that a market for immoral work benefits immoral types, particularly when they compete with fewer other immoral types.

We separately use survey responses to classify the perceived immorality of real-world firms and industries and show that industries classified as immoral pay higher wages. We also use surveys to obtain a separate measure of workers’ moral types. This individual characteristic is correlated with the moral type measured in the laboratory and predicts subjects’ behaviors in the laboratory labor market. Moreover, both the survey- and lab-based measures of morality also predict stated preferences for working in firms and industries outside the laboratory that vary in their perceived morality. Using either measure, workers less concerned with morality are more willing to work for firms and industries generally perceived as less moral.

Given the significance of many social ills produced by immoral work practices, such as deceptive marketing of socially harmful products, our study sheds important new light on the interaction between individual’s types, their willingness to do immoral work and the resulting labor-market outcomes. A limitation of our laboratory experiment is that we abstract from important features that occur in many labor markets outside the laboratory, such as search frictions and bargaining power. Future research can build on our experimental paradigm to investigate the importance of concerns for morality relative to such factors.

Our work has several potentially important policy implications. First, in industries such as tobacco and weapons manufacturing that have the potential to create substantial societal harm, social welfare may depend on workers’ voluntarily internalizing the negative impacts of their actions and forgoing potentially profitable opportunities. For instance, a weapons manufacturer may restrict sales to conflict areas if doing so is sufficiently aversive to top management. However, our evidence suggests that it is the *least* moral types who will sort into these industries, potentially making it less likely that such internalization will occur. This also creates an important contrast between our findings and related work on sorting by mission-oriented types into firms or jobs with a pro-social orientation. In such cases, sorting may often be beneficial for society, as those who care about a cause become the ones who impact it; in our case, however, those who care about doing “good” may—if they seek to

avoid socially stigmatized industries—avoid the opportunity to determine how much “bad” takes place. While it is beyond the scope of this paper to examine the societal consequences and the optimality of the sorting we have documented, we believe this is a fruitful and important avenue for future research.

Second, another implication of our empirical findings is that the perception that a firm, industry or type of work is immoral may be self-reinforcing. If, as our results indicate, the perception that work involves immoral acts leads people less concerned with acting morally to differentially opt into such work, then the end result of such sorting may be a workforce more likely to commit immoral acts. Even if some of the firms and industries that we study do not actually involve any inherently immoral activities in their line of work, the fact that they are disproportionately more attractive to people more willing to do immoral things may result in a self-confirming greater prevalence for immoral behavior. Thus, firms and industries that regularly confront the perception that they involve immoral work—such as financial services or pharmaceuticals—may need to be particularly attuned to such selection in their hiring.<sup>36</sup>

Finally, our theoretical analysis predicts—in line with our experimental data—that the least moral types are overcompensated by the immorality premium. This is in stark contrast to Mankiw’s (2010) “just deserts theory”—that is, everybody should receive his or her contribution to society. Our work suggests a perverse case in which those willing to do the most socially harmful acts may instead benefit from doing so. Moreover, this benefit is the direct result of the actions by others who are concerned with behaving morally. Indeed, we provide evidence that a shift in preferences toward a greater aversion to performing immoral work may reward those individuals who are least concerned with morality.

Of course, our work leaves open many important questions regarding the precise characteristics that lead some kinds of work to be differentially perceived as immoral and the specific nature of the preference underlying workers’ market behavior. Nevertheless, as the above examples make clear, the differential sorting by people more or less concerned with immoral behavior into different lines of work may have important implications for the extent to which market activity yields beneficial social outcomes.

---

<sup>36</sup> Problematic financial practices and scandals in the banking industry have caused serious economic damage and threaten the reputation of banks (Cohn, Fehr and Maréchal, 2014). Egan, Matvos and Seru (2019) document broad financial misconduct among financial advisers, resulting in substantial costs for the finance industry.

## 8. References

- Andreoni, J. (1989) "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence." *Journal of Political Economy*, 97(6): 1447–1458.
- Ariely, D., Bracha, A. and Meier, S. (2009) "Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially," *American Economic Review*, 99(1): 544-555.
- Arunachalam, R. and Shah, M. (2008) "Prostitutes and Brides?," *American Economic Review, Papers and Proceedings*, 98(2): 516-522.
- Ashraf, N., and Bandiera, O. (2017) "Altruistic Capital," *American Economic Review, Papers and Proceedings*, 107(5): 70-75.
- Ashraf, N., Bandiera, O., Davenport, E. and Lee, S. S. (2019) "Losing pro sociality in the quest for talent? Sorting, Selection and Performance in Public Service Delivery," working paper.
- Ashraf, N., Bandiera, O. and Delfino, A. (2020) "The distinctive values of bankers" *AEA Papers and Proceedings*, 110: 167-171.
- Ashton, M.C. and Lee, K. (2009) "The HEXACO-60: A Short Measure of the Major Dimensions of Personality," *Journal of Personality Assessment*, 91(4): 340-345.
- Barfort, S., Harmon, N. A., Hjorth, F. and Olsen, A. L. (2019) "Sustaining Honesty in Public Service: The Role of Selection." *American Economic Journal: Economic Policy*, 11 (4): 96-123.
- Bartling, B., Weber, R. A. and Yao, L. (2015) "Do Markets Erode Social Responsibility?," *Quarterly Journal of Economics*, 130(1): 219-66.
- Bates, C. and Rowell, A. (1998) "Tobacco explained: The truth about the tobacco industry ... in its own words." Scotland: ASH.
- Bénabou, R. and Tirole, J. (2006) "Incentives and Prosocial Behavior," *American Economic Review*, 96(5): 1652-1678.
- Bénabou, R., and Tirole, J. (2011) "Identity, morals, and taboos: Beliefs as assets," *Quarterly Journal of Economics*, 126(2): 805-855.
- Besley, T. and Ghatak, M. (2005) "Competition and Incentives with Motivated Agents," *American Economic Review*, 95(3): 616-636.
- Blitz, D. and Fabozzi, F. J. (2017) "Sin Stocks Revisited: Resolving the Sin Stock Anomaly," *Journal of Portfolio Management*, 44(1): 1-7.

- Bock, O., Baetge, I., and Nicklisch, A. (2014) „Hroot: Hamburg registration and organization on-line tool,” *European Economic Review*, 71: 117-120.
- Bradley M.M. and Lang, P.J. (1994) “Measuring emotion: the self-assessment manikin and the semantic differential,” *Journal of Behavioral Therapy and Experimental Psychiatry*, 25(1): 49-59.
- Brashear, G., Green, L. and Webb, J. (2000) “Development and validation of scales to measure attitudes influencing monetary donations to charitable organizations,” *Journal of the Academy of Marketing Science*, 28(2): 299-309.
- British American Tobacco (2015) “Annual Report 2015,” report.
- Cameron, A. C., Gelbach, J. G., and Miller, D. L. (2008) “Bootstrap-Based Improvements for Inference with Clustered Errors,” *Review of Economics and Statistics*, 90(3): 414-427.
- Cameron, A. C., Gelbach, J. B. and Miller, D. L. (2011) “Robust inference with multiway clustering,” *Journal of Business & Economic Statistics*, 29(2): 238-249.
- Carpenter, J., and Myers, C. K. (2010) “Why Volunteer? Evidence on the Role of Altruism, Image, and Incentives,” *Journal of Public Economics*, 94(11–12): 911–20.
- Carpenter, J. and Gong, E. (2016) “Motivating Agents: How Much Does the Mission Matter?,” *Journal of Labor Economics*, 34(1): 211-236.
- Carpenter, J., Matthews, P. and Robbett, A. (2017) “Compensating Differentials in Experimental Labor Markets,” *Journal of Behavioral and Experimental Economics*, 69: 50-60.
- Carter, J. R., and Irons, M. D. (1991) “Are Economists Different, and If So, Why?” *Journal of Economic Perspectives*, 5(2): 171-177.
- Cassar, L. and Meier, S. (2018) “Nonmonetary Incentives and the Implications of Work as a Source of Meaning,” *Journal of Economic Perspectives*, 32(3): 215-238.
- CNBC (2019) “Facebook has struggled to hire talent since the Cambridge Analytica scandal, according to recruiters who worked there,” accessed June 13, 2019, <https://www.cnbc.com/2019/05/16/facebook-has-struggled-to-recruit-since-cambridge-analytica-scandal.html>.
- Cohn, A., Fehr, E. and Maréchal, A. (2014) “Business culture and dishonesty in the banking industry,” *Nature*, 516: 86-89.
- Colonnello, S., Curatola, G. and Gioffré, A. (2019) “Pricing Sin Stocks: Ethical Preference vs. Risk Aversion,” *European Economic Review*, 118: 69-100.

- DellaVigna, S., List, J. A., Malmendier, U., and Rao, G. (2016) "Voting to tell others," *Review of Economic Studies*, 84(1): 143-181.
- Delfgaauw, J. and Dur, R. (2008) "Incentives and Workers' Motivation in the Public Sector," *Economic Journal*, 118(525): 171-191.
- Dufwenberg, M., Heidhues, P., Kirchsteiger, G., Riedel, F., and Sobel, J. (2011) "Other-regarding preferences in general equilibrium," *Review of Economic Studies*, 78(2), 613-639.
- Dur, R. and van Lent, M. (2019) "Socially Useless Jobs," *Industrial Relations*, 58(3), 543-546.
- Dur, R. and Zoutenbier, R. (2014) "Working for a Good Cause," *Public Administration Review*, 74(2): 144-155.
- Edlund L. and Korn, E. (2002) "A Theory of Prostitution," *Journal of Political Economy*, 110(1): 181-214.
- Egan, M., Matvos, G. and Seru, A. (2019) "The Market for Financial Adviser Misconduct," *Journal of Political Economy*, 127(1): 233-295.
- Falk A. and Szech, N. (2013) "Morals and Markets," *Science*, 340: 707-711.
- Fehr, E., Kirchler, E., Weichbold, A. and Gächter, S. (1998) "When social norms overpower competition: Gift exchange in experimental labor markets," *Journal of Labor Economics*, 16(2): 324-351.
- Fehrler, S. and Kosfeld, M. (2014) "Pro-Social Missions and Worker Motivation: An Experimental Study," *Journal of Economic Behavior & Organization*, 100: 99-110.
- Fischbacher, U. (2007) "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, 10(2), 171-178.
- Fisman, R., Jakiela, P., Kariv, S. and Markovits, D. (2015) "The distributional preferences of an elite," *Science*, 349(6254): 1300.
- Frank, R. H. (1996) "What prices the moral high ground?," *Southern Economic Journal*, 63(1): 1-17.
- Frank, R. H., Gilovich, T. and Regan, D. T. (1993) "Does Studying Economics Inhibit Cooperation?," *Journal of Economic Perspectives*, 7(2): 159-171.
- Friebel, G., Kosfeld, M. and Thielmann, G. (2019) "Trust the Police? Self-Selection of Motivated Agents into the German Police Force," *American Economic Journal: Microeconomics*, 11(4): 59-78.
- Gertler, P., Shah, M. and Bertozzi, S. M. (2005) "Risky Business: The Market for Unprotected Commercial Sex," *Journal of Political Economy*, 113(3): 518-550.

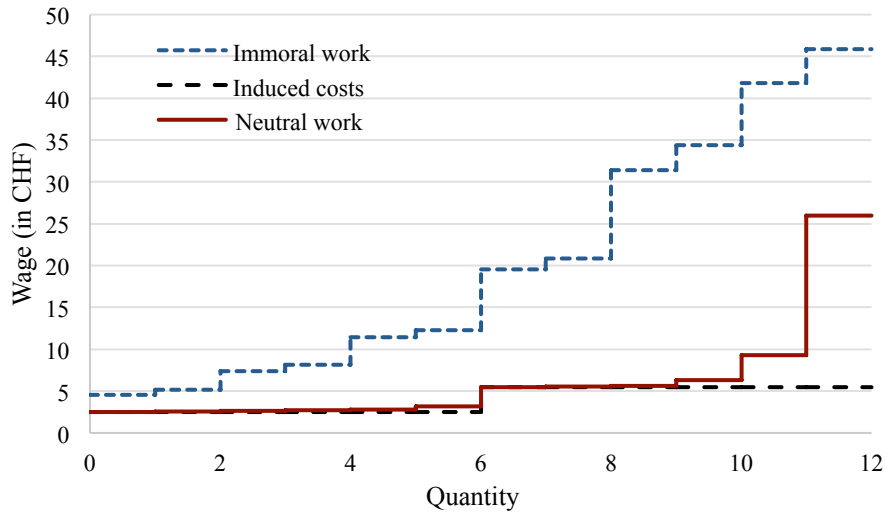
- Gibson, R., Tanner, C. and Wagner, A. (2013) "Preferences for truthfulness: Heterogeneity among and within individuals," *American Economic Review*, 103(1): 532-548.
- Gill, A., Heinz, M., Schumacher, H., and Sutter, M. (2020) "Trustworthiness in the financial industry," working paper.
- Gneezy, U., Rockenbach, B. and Serra-Garcia, M. (2013) "Measuring lying aversion," *Journal of Economic Behavior and Organization*, 93: 293-300.
- Gosling, S.D., Rentfrow, P.J., and Swann Jr. W.B. (2003) "A very brief measure of the Big-Five personality domains," *Journal of Research in Personality*, 37: 504-528.
- Hanna, R. and Wang, S. (2017) "Dishonesty and Selection into Public Service: Evidence from India," *American Economic Journal: Economic Policy*, 9 (3): 262-290.
- Heath, D. (2016) "Contesting The Science of Smoking," *The Atlantic*, <https://www.theatlantic.com/politics/archive/2016/05/low-tar-cigarettes/481116/>
- Heidhues, P., Köszegi, B. and Murooka, T. (2016) "Inferior products and profitable deception," *Review of Economic Studies*, 84(1): 323-356.
- Hill, C. A. (2012) "Bankers behaving badly? The limits of regulatory reform," *Review of Banking and Financial Law*, 31: 675-691.
- Hong, H. and Kacperczyk, M. (2009) "The Price of Sin: The Effects of Social Norms on Markets," *Journal of Financial Economics*, 93: 15-36.
- Jones, D. B. (2015) "The Supply and Demand of Motivated Labor: When Should We Expect to See Nonprofit Wage Gaps?," *Labour Economics* 32: 1-14.
- Kirchler, M., Huber, J., Stefan, M. and Sutter, M. (2016) "Market design and moral behavior," *Management Science*, 62: 2615-2625.
- Lazear, E. P., Malmendier, U. and Weber, R. A. (2012) "Sorting in Experiments with Application to Social Preferences," *American Economic Journal: Applied Economics*, 4(1): 136-163.
- Leary, M.R., Diebels, K. J. and Jongman-Sereno, K. P. (2015) "Measures of concerns with public image and social evaluation", in: *Measures of Personality and Social Psychological Constructs*, Elsevier Inc., 448-473. DOI: 10.1016/B978-0-12-386915-9.00016-4.
- Leete, L. (2001) "Whither the Nonprofit Wage Differential? Estimates from the 1990 Census," *Journal of Labor Economics* 19(1): 136-170.
- Levitt, S. D. and List, J. A. (2007) "What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World?," *Journal of Economic Perspectives*, 21(2): 153-174.

- Levitt, S. D. and List, J. A. (2008) "Homo economicus Evolves," *Science*, 319: 909-910.
- Mankiw, G. N. (2010) "Spreading the Wealth Around: Reflections Inspired by Joe the Plumber," *Eastern Economic Journal*, 36: 285-298.
- Mas, A., and Pallais, A. (2017) "Valuing Alternative Work Arrangements," *American Economic Review*, 107(12): 3722-3759.
- Mocan, N. H. and Tekin, E. (2003) "Nonprofit Sector and Part-Time Work: An Analysis of Employer-Employee Matched Data on Child Care Workers," *Review of Economics and Statistics* 85(1): 38-50.
- Moffatt, P. G. and Peters, S. A. (2004) "Pricing personal services: An empirical study of earnings in the UK prostitution industry," *Scottish Journal of Political Economy* 51(5): 675-690.
- Oh, S. (2021) "Does identity affect labor supply?," working paper.
- Okie, S. (2010) "A flood of opioids, a rising tide of deaths," *New England Journal of Medicine*, 363(21): 1981-1985.
- Philip Morris International Inc. (2015) "Form 10-K submitted to US Securities and Exchange Commission," report.
- Rosen, S. (1986) „The theory of equalizing differences,“ Ed: Ashenfelter, O. and Layard, R., in: *Handbook of Labor Economics*, Elsevier Science Publishers BV.
- Sausgruber, R. and Tyran, J-R. (2011) "Are we taxing ourselves? How deliberation and experience shape voting on taxes", *Journal of Public Economics*, 95: 164-176.
- Smith, A. (1776) "An Inquiry into the Nature and Causes of the Wealth of Nations," London: W. Strahan and T. Cadell.
- Smith, V.L., Williams A.W., Bratton W.K. and Vannoni, M.G. (1982) "Competitive market institutions: Double auctions vs. sealed bid-offer auctions," *American Economic Review*, 72(1): 58-77.
- Tonin, M. and Vlassopoulos, M. (2015) „Corporate Philanthropy and Productivity: Evidence from an Online Real Effort Experiment,“ *Management Science*, 61(8): 1795-1811.
- US Department of Justice, Office of Public Affairs (2016) "Goldman Sachs Agrees to Pay More than \$5 Billion in Connection with Its Sale of Residential Mortgage Backed Securities," <https://www.justice.gov/>, date accessed: 21.01.2017.
- Wiswall, M. and Zafar, B. (2018) "Preference for the Workplace, Investment in Human Capital, and Gender," *Quarterly Journal of Economics*, 133(1): 457-507.
- Ziegler, A., Romagnoli, G. and Offerman, T. (2020) "Morals in multi-unit markets," working paper.

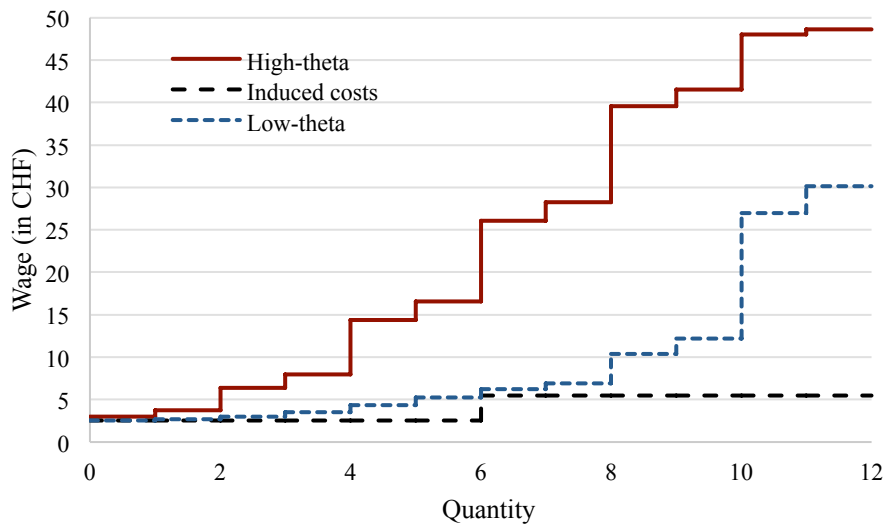


## Appendix A – Additional Analysis

**Figure A1: Labor supply for neutral and immoral work in the laboratory, last 5 periods**

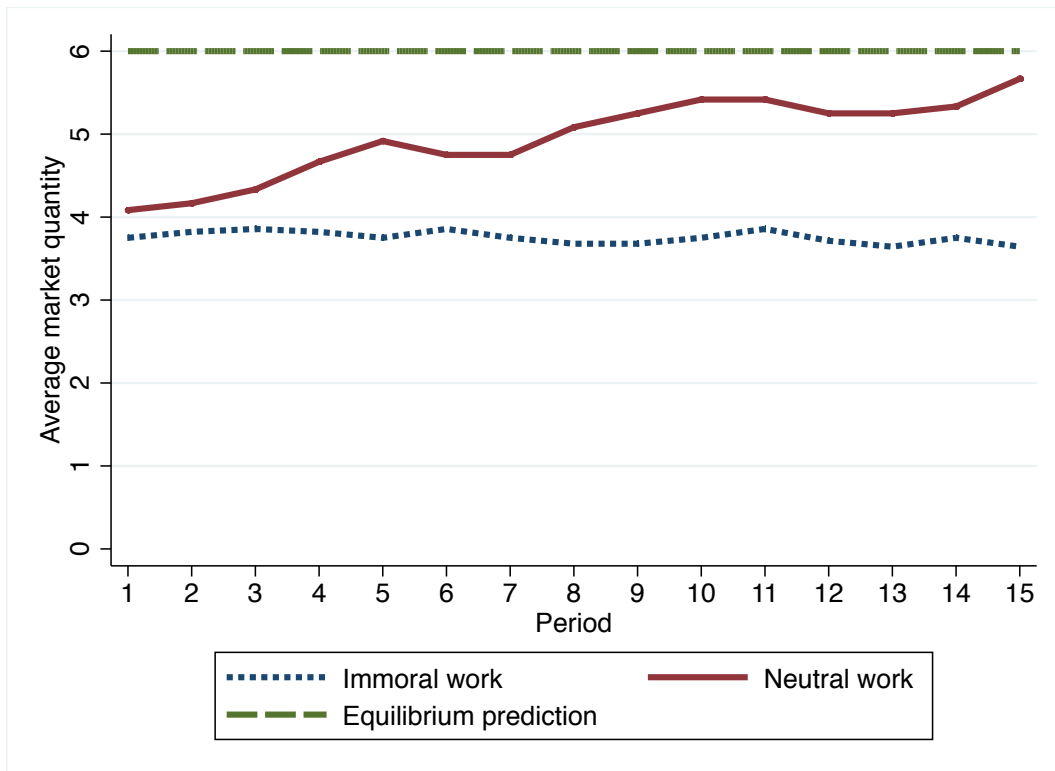


**Figure A2: Labor supply for immoral work in the laboratory for different types ( $\theta^{Exp}$ )**

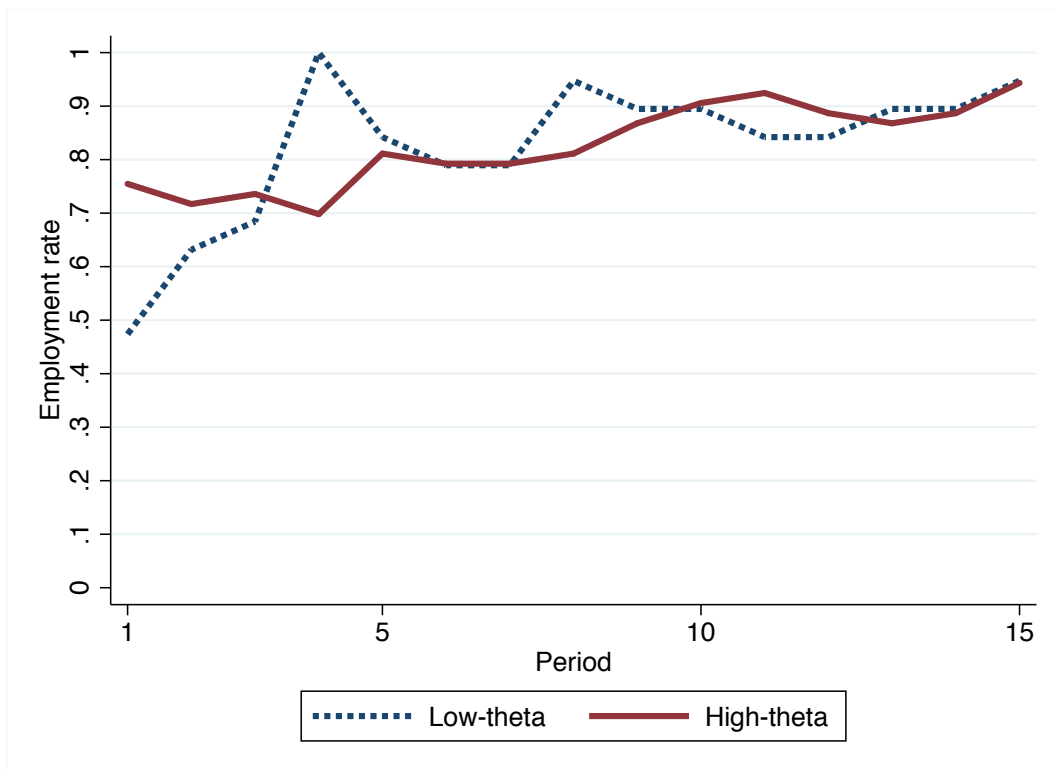


*Notes: Labor supplies conditional on types are calculated with a simulation: 6 labor market decisions (first and second wage request) of high-theta (or, low-theta) types are randomly drawn (without replacement) from our sample. We then calculate the labor supply for this group of people. We repeat this 1000 times (with replacement) and take the average of these 1000 individual labor supplies. This approach differs from the one we use in Figure 4 and Figure A1, where we take the average of the actual labor supplies in the different market groups and periods. Figure 4 does not change substantially if we simulate market composition instead of using actual composition.*

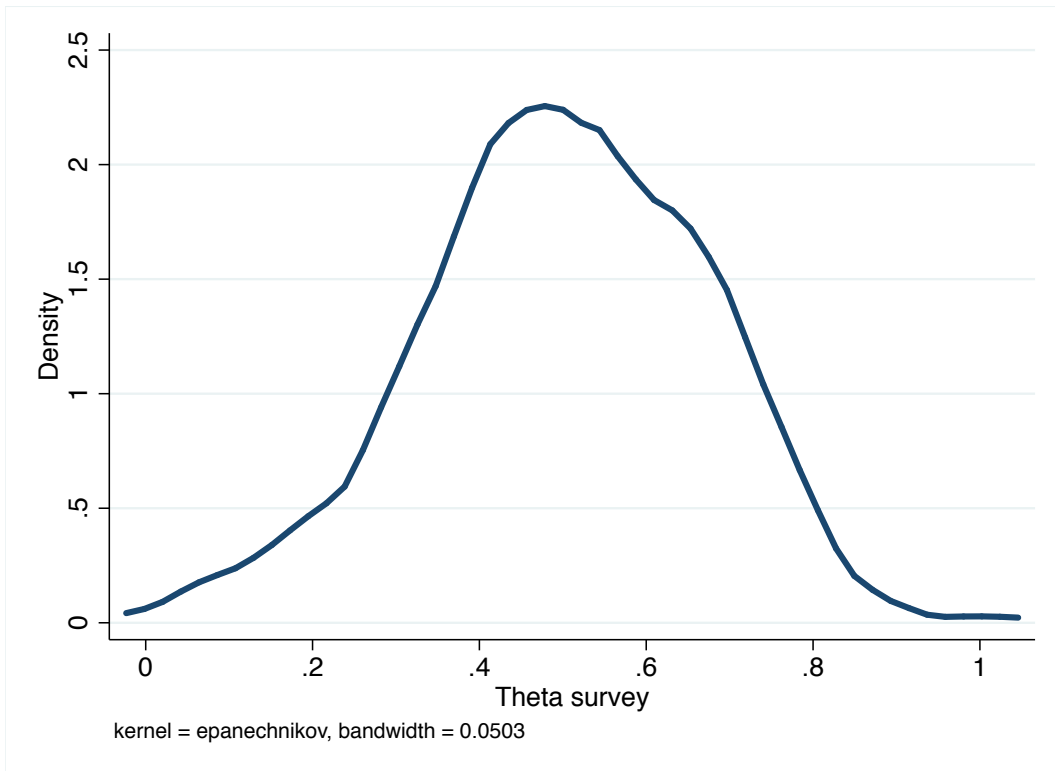
**Figure A3: Market quantities in laboratory labor markets**



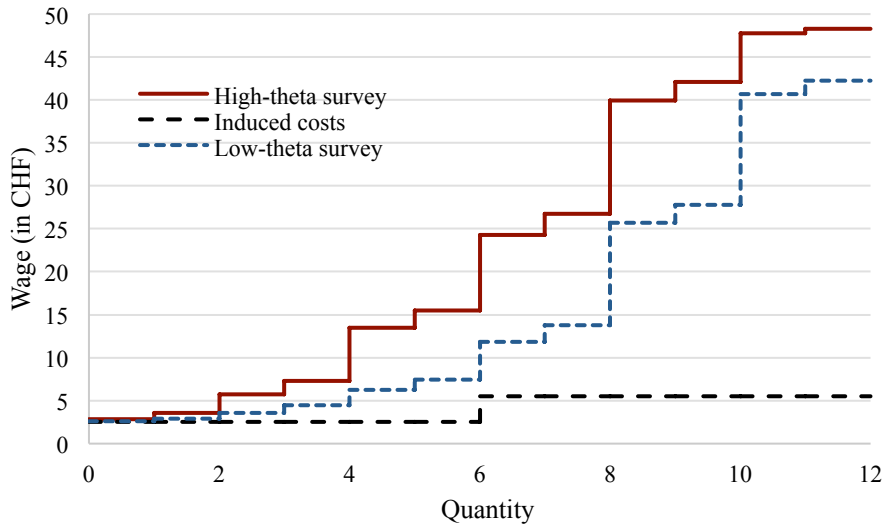
**Figure A4: Employment rate by the two types in the neutral work condition**



**Figure A5: Probability distribution of  $\theta^{Sur}$**



**Figure A6: Labor supply for immoral work in the laboratory for different types ( $\theta^{Sur}$ )**



Notes: High-theta survey is 1 if  $\theta^{Sur}$  is lower than the median of  $\theta^{Sur}$ . Labor supplies are calculated with simulations: 6 labor market decisions (first and second wage request) of high-theta survey (or, low-theta survey) types are randomly drawn (without replacement) from our sample. We then calculate the labor supply for this group of people. We repeat this 1000 times (with replacement) and take the average of these 1000 individual labor supplies.

**Table A1: Perceived immortality of industries and summary of main variables from the Swiss Labor Force Survey**

Industries	N	Real gross hourly wage (in 2010 CHF)	Ln of the real gross hourly wage	Age	Male	Married	Edu high	Edu middle	Swiss	Experience	Full-time equivalent	Managerial duties	Industry size	Industry Sales	Perceived immortality I(j)
All industries	32,638	41.00 (20.11)	3.59 (0.54)	42.25 (11.16)	0.57 (0.49)	0.58 (0.49)	0.35 (0.48)	0.51 (0.50)	0.64 (0.48)	9.56 (9.21)	0.87 (0.24)	0.17 (0.37)	106.46 (87.54)	0.94 (1.27)	-0.18 (0.22)
Processing of tea and coffee	138	38.88 (14.00)	3.60 (0.35)	41.29 (10.34)	0.66 (0.47)	0.67 (0.47)	0.24 (0.43)	0.59 (0.49)	0.55 (0.50)	8.78 (8.94)	0.93 (0.20)	0.10 (0.30)	3.10 (0.43)	0.22 (0.02)	-0.11 (-0.41)
Manufacture of tobacco products	162	51.01 (20.20)	3.85 (0.44)	39.42 (8.26)	0.59 (0.49)	0.59 (0.49)	0.45 (0.50)	0.41 (0.49)	0.52 (0.50)	9.63 (8.84)	0.92 (0.10)	0.03 (0.16)	2.24 (0.15)	11.79 (0.97)	0.47 (-0.42)
Manufacture of paper and paperboard	96	42.33 (16.25)	3.69 (0.33)	46.00 (10.77)	0.78 (0.42)	0.77 (0.42)	0.35 (0.48)	0.50 (0.50)	0.57 (0.50)	16.80 (13.32)	0.96 (0.10)	0.09 (0.29)	1.58 (0.22)	2.11 (1.11)	-0.06 (-0.37)
Manufacture of weapons and ammunitions	104	53.26 (20.66)	3.91 (0.37)	48.53 (10.85)	0.95 (0.22)	0.58 (0.50)	0.57 (0.50)	0.37 (0.48)	0.83 (0.38)	16.36 (13.49)	0.96 (0.06)	0.10 (0.31)	1.44 (0.10)	0.46 (0.05)	0.71 (-0.40)
Manufacture of electronic components	1,133	44.79 (18.21)	3.72 (0.41)	43.27 (10.59)	0.68 (0.47)	0.61 (0.49)	0.51 (0.50)	0.37 (0.48)	0.53 (0.50)	10.47 (9.64)	0.93 (0.14)	0.06 (0.24)	22.08 (0.76)	0.66 (0.16)	-0.01 (-0.42)
Construction of buildings	3,600	37.87 (13.18)	3.57 (0.40)	43.21 (10.73)	0.91 (0.28)	0.70 (0.46)	0.21 (0.41)	0.48 (0.50)	0.48 (0.50)	11.16 (9.89)	0.95 (0.15)	0.14 (0.34)	82.97 (0.79)	0.43 (0.01)	-0.28 (-0.39)
Maintenance and repair of motor vehicles	2,664	34.64 (12.89)	3.48 (0.38)	41.38 (12.08)	0.82 (0.38)	0.61 (0.49)	0.19 (0.40)	0.71 (0.45)	0.66 (0.47)	11.47 (10.47)	0.92 (0.21)	0.26 (0.44)	58.65 (0.33)	0.54 (0.01)	-0.28 (-0.40)
Wholesale of tobacco products	87	58.75 (27.28)	3.96 (0.48)	45.10 (9.98)	0.55 (0.50)	0.70 (0.46)	0.51 (0.50)	0.49 (0.50)	0.65 (0.48)	11.05 (7.86)	0.93 (0.15)	0.13 (0.33)	1.72 (0.07)	1.49 (0.19)	0.44 (-0.38)
Wholesale of clothing and footwear	306	38.92 (18.97)	3.54 (0.51)	40.21 (10.73)	0.31 (0.46)	0.45 (0.50)	0.38 (0.49)	0.53 (0.50)	0.52 (0.50)	7.16 (6.70)	0.84 (0.24)	0.16 (0.37)	6.08 (0.18)	1.80 (0.22)	0.10 (-0.46)
Wholesale of perfume and cosmetics	353	53.47 (28.01)	3.84 (0.54)	39.78 (10.03)	0.34 (0.47)	0.60 (0.49)	0.58 (0.49)	0.36 (0.48)	0.40 (0.49)	7.80 (6.48)	0.89 (0.19)	0.15 (0.35)	6.19 (0.26)	1.89 (0.22)	0.12 (-0.37)

*See next page for the rest of the table.*

Industries	N	Real gross hourly wage (2010 CHF)	Ln of the real gross hourly wage	Age	Male	Married	Edu high	Edu middle	Swiss	Experience	Full-time equivalent	Managerial duties	Industry size	Industry Sales	Perceived immorality I(G)
Wholesale of watches and jewelry	158	44.16 (19.05)	3.71 (0.40)	43.89 (10.68)	0.42 (0.49)	0.58 (0.50)	0.46 (0.50)	0.48 (0.50)	0.55 (0.50)	8.09 (8.04)	0.88 (0.19)	0.20 (0.40)	2.65 (0.17)	2.13 (0.30)	0.04 (-0.41)
Hotels and similar accommodation	2,658	27.06 (10.58)	3.22 (0.48)	40.69 (11.24)	0.40 (0.49)	0.55 (0.50)	0.19 (0.39)	0.55 (0.50)	0.40 (0.49)	7.16 (7.59)	0.85 (0.26)	0.16 (0.36)	59.16 (0.75)	0.16 (0.00)	-0.34 (-0.37)
Restaurants and mobile food activities	5,560	25.12 (9.96)	3.14 (0.47)	40.55 (11.92)	0.43 (0.49)	0.56 (0.50)	0.14 (0.34)	0.57 (0.50)	0.46 (0.50)	6.58 (7.35)	0.79 (0.29)	0.21 (0.41)	94.45 (1.03)	0.15 (0.00)	-0.33 (-0.37)
Monetary intermediations	7,286	56.00 (21.87)	3.94 (0.44)	42.00 (10.37)	0.59 (0.49)	0.55 (0.50)	0.55 (0.50)	0.42 (0.49)	0.73 (0.44)	10.65 (9.42)	0.91 (0.18)	0.19 (0.40)	109.31 (2.05)	2.26 (0.88)	0.11 (-0.40)
Credit granting	79	53.62 (22.28)	3.90 (0.39)	40.00 (8.59)	0.46 (0.50)	0.53 (0.50)	0.51 (0.50)	0.48 (0.50)	0.67 (0.47)	5.45 (5.73)	0.89 (0.17)	0.11 (0.31)	1.23 (0.03)	1.50 (0.83)	0.15 (-0.41)
Non-life insurance	2,637	49.20 (18.82)	3.82 (0.39)	42.09 (11.02)	0.49 (0.50)	0.51 (0.50)	0.51 (0.50)	0.47 (0.50)	0.80 (0.40)	9.91 (9.07)	0.89 (0.20)	0.09 (0.29)	40.52 (0.97)	1.76 (0.76)	-0.13 (-0.44)
General public administration activities	4,563	44.42 (17.24)	3.70 (0.52)	45.36 (10.74)	0.48 (0.50)	0.60 (0.49)	0.42 (0.49)	0.52 (0.50)	0.91 (0.29)	10.71 (9.53)	0.80 (0.28)	0.14 (0.35)	311.69 (6.54)	0.02 (0.00)	-0.41 (-0.36)
Gambling and betting activities	153	38.96 (16.01)	3.59 (0.40)	41.83 (10.35)	0.61 (0.49)	0.50 (0.50)	0.31 (0.46)	0.52 (0.50)	0.45 (0.50)	7.51 (5.41)	0.89 (0.20)	0.12 (0.33)	2.69 (0.06)	0.77 (0.05)	0.42 (-0.41)
Organization and operation of sport facilities for indoor and outdoor sports events	502	34.03 (15.83)	3.40 (0.66)	44.03 (12.42)	0.50 (0.50)	0.53 (0.50)	0.27 (0.44)	0.57 (0.50)	0.70 (0.46)	7.63 (8.20)	0.72 (0.34)	0.11 (0.32)	3.69 (0.13)	0.24 (0.01)	-0.49 (-0.40)
Fitness facilities	399	30.76 (16.32)	3.29 (0.56)	40.82 (11.22)	0.28 (0.45)	0.58 (0.49)	0.28 (0.45)	0.62 (0.49)	0.67 (0.47)	6.06 (6.72)	0.54 (0.34)	0.19 (0.39)	4.61 (0.38)	0.15 (0.00)	-0.35 (-0.38)

Source: Weighed data from the SLFS, years 2010-2016 (wage and demographics), STATENT, years 2011-2016 (industry size, industry sales), Value Added Tax Statistics, years 2010-2016 (industry sales) and our own survey (perceived industry immorality). Notes: N=number of observations per industry in the SLFS dataset, Male in {0, 1}, Married in {0, 1}, Education high: higher vocational education and training or university/college, Education middle: apprenticeship, full-time vocational school, matura or pedagogical training, Education low (reference category): compulsory schooling or pre-vocational education, Swiss in {0, 1}, Experience = number of years in the firm, Full-time equivalent = (working hours /42), set to 1 for working hours >= 42, managerial duties in {0, 1}, Industry size = number employees in this industry / 1000 (2010 data is not available, we substitute it with 2011 data), Industry sales = Industry sales/number employees in this industry, Perceived immorality is in [-1, 1] where -1 means very moral, 0 means neutral and 1 means very immoral. Standard deviations in parentheses.

**Table A2: Perceived immorality of firms**

<b>Firms</b>	<b>Perceived immorality I(j)</b>	<b>Firms</b>	<b>Perceived immorality I(j)</b>
Marlboro	0.54	Swisscom	-0.07
Monsanto	0.52	Firmenich	-0.09
Glencore	0.46	Winterthur Assurance	-0.1
Philip Morris	0.46	Swiss Life	-0.13
Nestlé	0.39	Swatch	-0.17
Tamoil	0.37	Adecco	-0.18
Syngenta	0.23	ABB	-0.2
UBS	0.19	Migros	-0.38
Novartis	0.18	WWF	-0.66
Credit Suisse	0.17	Pro Juventute	-0.66
Roche	0.13	Pro Natura	-0.67
Holcim	0.03	UNICEF	-0.72
Ernst and Young	-0.05	Red cross	-0.81

*Source: own survey.*

*Notes: Perceived immorality is in [-1, 1] where -1 means very moral, 0 means neutral and 1 means very immoral.*

**Table A3: Distribution of behavior regarding the behavioral measure of concern for morality**

Number of lies	Reported number for state r:						Expected payoff lying	Frequency	Share	Classification
	1	2	3	4	5	6				
0 (Honest)	1	2	3	4	5	6	0	161	0.671	High-theta
1	6	2	3	4	5	6	0.83	6	0.038	Low-theta
	2	2	3	4	5	6	0.17	3		
2	6	6	3	4	5	6	1.5	12	0.104	Low-theta
	1	2	3	6	6	6	0.5	2		
	1	3	3	5	5	6	0.33	1		
	1	4	4	4	5	6	0.5	1		
	5	6	3	4	5	6	1.33	2		
	6	5	3	4	5	6	1.33	1		
	3	2	3	5	5	6	0.5	1		
	3	3	3	4	5	6	0.5	5		
3	6	6	6	4	5	6	2	11	0.050	Low-theta
	4	2	3	6	6	6	1	1		
4	6	6	6	6	5	6	2.33	3	0.017	Low-theta
	6	5	5	5	5	6	1.83	1		
5	6	6	6	6	6	6	2.5	15	0.067	Low-theta
	2	3	4	5	6	6	0.83	1		
Lied in a self-harmful manner	1	2	3	4	3	3	-0.83	1	0.054	High-theta
	1	2	3	4	4	4	-0.5	1		
	1	2	3	4	5	5	-0.17	1		
	1	3	2	5	4	6	0	1		
	1	4	2	4	5	6	0.17	1		
	1	4	6	3	5	6	0.67	1		
	2	1	3	4	5	6	0	1		
	3	4	5	4	6	2	0.5	1		
	5	1	3	6	4	2	0	1		
	5	2	3	4	1	6	0	1		
	5	4	6	4	6	5	1.5	1		
	6	2	5	5	1	3	0.17	1		
	6	6	6	6	6	5	2.33	1		

Notes: Expected payoff from lying =  $\frac{1}{6} \sum_{r=1}^6 (m_{ir} - r)$ , where  $m_{ir}$  is the number that individual  $i$  reports if the actual die roll is  $r$ .

**Table A4: Relationship participation decision/reservation wage and  $\theta^{Exp}$  (Hurdle model)**

Dependent variable:	Participate	Reservation wage	Reservation wage
	(1)	(2)	(3)
Low-theta ( $\theta_L^{Exp}$ )	0.925*** (4.64)	-0.362 (-0.99)	-0.060 (-0.39)
Period (t)	-0.019** (-2.40)	-0.047 (-1.64)	-0.050*** (-5.74)
Period * $\theta_L^{Exp}$	0.012 (1.14)	-0.015 (-0.40)	0.005 (0.54)
Constant	0.449*** (3.86)	4.425*** (14.76)	3.312*** (25.63)
Sigma		2.630*** (7.69)	0.571*** (10.58)
Condition	Immoral work	Immoral work	Neutral work
N	2520	1755	1077
p-value: $t + t * \theta_L^{Exp} = 0$	0.422	0.001	0.0000

Notes: Estimates from Craggs double-hurdle model: (1) is a probit model; (2) and (3) are truncated linear regressions (truncated from above at 50 CHF). Models (1) and (2) use only data from the immoral work condition; model (3) uses only data from the neutral work condition. For neutral work, we do not report the regression of market participation as we have only 3 incidences where a subject did not participate. Independent variables: Low-theta in  $\{0, 1\}$ , Period between 1 and 15. Standard errors clustered at market level; z-statistics in parentheses; \* -  $p < 0.1$ ; \*\* -  $p < 0.05$ ; \*\*\* -  $p < 0.01$ .

**Table A5: Relationship between the behavioral measures of concern for morality and outcomes in the experimental labor markets, robustness**

Dependent variable:	Employment rate							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<b>Number of lies</b>								
1 lie	0.201 (1.05)	0.128 (0.53)			0.004 (0.11)	-0.008 (-0.12)		
2 lies	0.220** (2.17)	0.182 (1.39)			-0.011 (-0.21)	-0.049 (-0.84)		
3 lies	0.398*** (5.29)	0.351*** (3.57)			-0.296*** (-12.89)	-0.326*** (-15.52)		
4 lies	0.392*** (4.94)	0.279** (2.61)			0.171*** (7.43)	0.167*** (7.18)		
5 lies	0.286*** (3.26)	0.233** (2.48)			0.037 (0.49)	0.018 (0.17)		
self-harmful lies	0.252** (2.53)	0.182 (1.37)			-0.0516 (-1.25)	-0.001 (-0.04)		
<b>Expected payoff lying</b>			0.127*** (5.04)	0.105*** (3.18)			0.006 (0.18)	-0.001 (-0.03)
<b>Constant</b>	0.475*** (11.6)	0.517*** (32.97)	0.510*** (13.79)	0.512*** (37.01)	0.829*** (36.11)	0.806*** (35.35)	0.824*** (35.65)	0.801*** (36.06)
<b>Condition</b>	Immoral	Immoral	Immoral	Immoral	Neutral	Neutral	Neutral	Neutral
<b>N</b>	168	168	168	168	72	72	72	72
<b>R<sup>2</sup></b>	0.121	0.263	0.0752	0.238	0.085	0.250	0.001	0.162
<b>Market FE</b>	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Coefficient estimates of linear regression models. Models (1)-(4) use only data from the immoral work condition, models (5)-(8) use only data from the neutral work condition. Expected payoff from lying =  $\frac{1}{6} \sum_{r=1}^6 m_{ir} - r$  ( $\in [-2.5, 2.5]$ ), where  $m_{ir}$  is the number that individual  $i$  reports if the actual die roll is  $r$ . Standard errors clustered at market level; t-statistics in parentheses; \* -  $p < 0.1$ ; \*\* -  $p < 0.05$ ; \*\*\* -  $p < 0.01$ .



**Table A6: Relationship between  $\theta^{Exp}$  and market income, depending on the behavior of other market participants**

Dependent variable:	Market income
Low-theta ( $\theta_L^{Exp}$ )	-3.679 (-1.05)
Many $\theta_H$ types	4.334* (2.05)
$\theta_L^{Exp}$ * Many $\theta_H$ types	16.348*** (3.53)
Constant	18.046*** (11.61)
N	168
R <sup>2</sup>	0.089
p-value: Many $\theta_H$ types + $\theta_L^{Exp}$ * Many $\theta_H$ types=0	0.0002

Notes: Coefficient estimates of linear regression models. Sample includes only subjects from the immoral work condition. Low-theta in {0, 1}, Many  $\theta_H$  types: 0=number of (other) low-theta type workers is lower than 2, 1=number of (other) low-theta type workers is 2 or more. Standard errors clustered at market level; t-statistics in parentheses; \* -  $p < 0.1$ ; \*\* -  $p < 0.05$ ; \*\*\* -  $p < 0.01$ .

**Table A7: Description and summary statistics of survey scales**

Variable	Number of items	Mean (Sd)	Interpretation
Protected value 1	5	0.75 (0.19)	1 = the person finds that the behavior of a banker who recommends sub-optimal assets to his clients because he has larger margins on them is: very outrageous, very blameworthy, very immoral, not at all acceptable and not at all praiseworthy.
Protected value 2	4	0.63 (0.18)	1 = the person thinks that truthfulness is a value that cannot be sacrificed.
Work ethics 1	1	0.38 (0.29)	1 = the person thinks that people are generally honest.
Work ethics 2	1	0.55 (0.36)	1 = the person thinks that calling sick to have a free day at work is really bad.
HEXACO sincerity	3	0.59 (0.19)	1 = the person is very sincere.
HEXACO fairness	3	0.68 (0.22)	1 = the person is very fair.
HEXACO greed avoidance	2	0.58 (0.23)	1 = the person is not at all greedy.
HEXACO modesty	2	0.66 (0.22)	1 = the person is very modest.
Charity attitude index	9	0.69 (0.13)	1 = the person's attitude towards charities is very positive.

Notes: Each subscale is constructed by taking averages over all items of the scale, and then normalized such that it lies between 0 and 1.

**Table A8: Relationship between  $\theta^{Sur}$  and outcomes in the experimental labor market**

Dependent variable:	Employment rate					
	(1)	(2)	(3)	(4)	(5)	(6)
Type survey ( $\theta^{Sur}$ )	-0.008 (-0.06)	0.036 (0.22)	0.104 (0.45)	0.150 (0.75)	0.0103 (0.06)	0.015 (0.06)
Immoral work (IW)	-0.046 (-0.37)		0.013 (0.07)		0.133 (0.59)	
$\theta^{Sur} * IW$	-0.431 (-1.64)	-0.366 (-1.14)	-0.558 (-1.53)	-0.537 (-1.26)	-0.648* (-1.69)	-0.505 (-1.05)
Aggregation $\theta^{Sur}$	Factor Analysis	Factor Analysis	Equal weight	Equal weight	Theta- Exp	Theta- Exp
N	237	237	237	237	237	237
R <sup>2</sup>	0.137	0.294	0.132	0.291	0.133	0.293
p-value: $\theta^{Sur} + \theta^{Sur} * Im = 0$	0.0569	0.237	0.0924	0.311	0.0708	0.252
Market FE	No	Yes	No	Yes	No	Yes

Notes: Coefficient estimates of linear regression models. Models differ in how we construct  $\theta^{Sur}$  from the nine psychological survey measures. Columns (1) and (2) report our main results, using factor analysis to aggregate the psychological measures. Columns (3) and (4) give the result if equal weight is given to each measure instead. Columns (5) and (6) give the results if weights are determined by a regression of the survey measures on  $\theta^{Exp}$ . Other independent variables: Immoral work is in  $\{0, 1\}$ ,  $\theta^{Sur}$  is in  $[0, 1]$ , where higher numbers indicate more morality. Standard errors clustered at market level; t-statistics in parentheses; \* -  $p < 0.1$ ; \*\* -  $p < 0.05$ ; \*\*\* -  $p < 0.01$ .

**Table A9: Relationship between participation decision/reservation wage and  $\theta^{Sur}$  (Hurdle model) in the immoral work condition**

Dependent variable:	Participate	Reservation wage	Participate	Reservation wage	Participate	Reservation wage
	(1)	(2)	(3)	(4)	(5)	(6)
Type survey ( $\theta^{Sur}$ )	-1.614** (-2.44)	0.432 (0.60)	-1.971** (-2.08)	0.571 (0.61)	-2.177** (-2.13)	0.373 (0.33)
Constant	1.326*** (3.98)	3.674*** (10.39)	1.806*** (2.99)	3.515*** (6.12)	1.841*** (3.01)	3.657*** (5.25)
Sigma		2.679*** (7.71)		2.679*** (7.72)		2.680*** (7.70)
Aggregation $\theta^{Sur}$	Factor Analysis	Factor Analysis	Equal weight	Equal weight	Theta- Exp	Theta- Exp
N	2475	1711	2475	1711	2475	1711
LL (pseudo)	-1478.3	-4114.1	-1488.2	-4114.2	-1490.9	-4114.6

Notes: Estimates from Craggs double-hurdle model: Regressions (1), (3) and (5) are probit models, regressions (2), (4) and (6) are truncated linear regressions (truncated from above at 50 CHF). Regressions differ in how  $\theta^{Sur}$  is constructed from the nine psychological survey measures. Columns (1) and (2) report our main results, using factor analysis to aggregate the psychological measures. Columns (3) and (4) give the result if equal weight is given to each measure instead. Columns (5) and (6) give the results if weights are determined by a regression of the survey measures on  $\theta^{Exp}$ .  $\theta^{Sur}$  is in  $[0, 1]$ , where higher numbers indicate more morality. Sample includes only subjects from the immoral work condition. (In the neutral work condition, the coefficient of  $\theta^{Sur}$  is not significant for any of the above specifications.) Standard errors clustered at market level; z-statistics in parentheses; \* -  $p < 0.1$ ; \*\* -  $p < 0.05$ ; \*\*\* -  $p < 0.01$ .

**Table A10: Regressions of willingness to work for diverse industries and firms on perceived immorality and moral types, robustness checks aggregation  $\theta^{Sur}$**

Dependent variable:	Willingness to work for industry $j$				Willingness to work for firm $j$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<b>Perceived immorality (<math>I(j)</math>)</b>	0.089 (1.00)	0.096 (1.17)	0.122 (1.33)	0.129 (1.43)	0.396*** (3.65)	0.391*** (3.58)	0.337*** (4.16)	0.327*** (4.00)
<b>Type survey (<math>\theta^{Sur}</math>)</b>	-0.154* (-1.93)	-0.149* (-1.77)	-0.131 (-1.49)	-0.158* (-1.68)	-0.200* (-1.84)	-0.267*** (-2.58)	-0.243** (-2.21)	-0.298*** (-2.69)
<b><math>\theta^{Sur} * I(j)</math></b>	-0.583*** (-5.05)	-0.583*** (-5.06)	-0.671*** (-4.95)	-0.671*** (-4.98)	-0.998*** (-7.83)	-0.990*** (-7.60)	-0.961*** (-8.54)	-0.944*** (-8.26)
<b>Aggregation <math>\theta^{Sur}</math></b>	Equal weight	Equal weight	Theta-Exp	Theta-Exp	Equal weight	Equal weight	Theta-Exp	Theta-Exp
<b>N</b>	4715	4715	4715	4715	5064	5064	5064	5064
<b>Control variables</b>	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Coefficient estimates of linear regression models. Observations where subjects did not know the firm (“I don’t know this organization”) or did not fill out the questionnaire are excluded. Independent variables: Models differ in how we construct  $\theta^{Sur}$  from the nine psychological survey measures. Columns (1), (2), (5) and (6) give the result if equal weight is given to each measure. Columns (3), (4), (7) and (8) give the results if weights are determined by a regression of the survey measures on  $\theta_L^{Exp}$ .  $\theta^{Sur}$  is in  $[0,1]$  where higher numbers indicate more morality. Willingness to work is in  $\{0, 0.25, 0.5, 0.75, 1\}$  where 0 means not at all willing to work, 0.5 means indifferent and 1 means really much willing to work. Perceived immorality is in  $[-1, 1]$  where -1 means very moral, 0 means neutral and 1 means very immoral. Control variables: age, gender, Swiss nationality, subject of study, average wage industry 2016 (SLFS; only for industries), industry size 2016 (STATENT; only for industries), industry sales 2015 (Value Added Tax Statistics; only for industries). Standard errors clustered at individual and industry/firm level (Cameron, Gelbach and Miller, 2011); z-statistics in parentheses; \*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$ .

**Table A11: Regressions of willingness to work for diverse industries and firms on perceived immorality and moral types, robustness checks classification firms immorality**

Dependent variable:	Willingness to work for firm $j$			
	(1)	(2)	(3)	(4)
<b>Perceived immorality (<math>I_{Alt}(j)</math>)</b>	0.141 (1.63)	0.138 (1.59)	-0.157** (-1.96)	-0.156** (-1.96)
<b>Type survey (<math>\theta^{Sur}</math>)</b>	-0.184** (-2.42)	-0.222*** (-2.99)		
<b><math>\theta^{Sur} * I_{Alt}(j)</math></b>	-0.848*** (-9.22)	-0.840*** (-9.10)		
<b>Type experiment (<math>\theta^{Exp}</math>)</b>			-0.045* (-1.65)	-0.053** (-2.04)
<b><math>\theta^{Exp} * I_{Alt}(j)</math></b>			-0.174*** (-4.06)	-0.175*** (-4.15)
<b>N</b>	5064	5064	5064	5064
<b>Control variables</b>	No	Yes	No	Yes

Notes: Coefficient estimates of linear regression models. Perceived immorality is calculated different then in our main analysis: Clients that choose “I don’t know this organization” are classified as giving neutral ratings. Dependent variable: Willingness to work is in  $\{0, 0.25, 0.5, 0.75, 1\}$  where 0 means not at all willing to work, 0.5 means indifferent and 1 means really much willing to work. Observations where subjects did not know the firm (“I don’t know this organization”) or did not fill out the questionnaire are excluded. Independent variables: (1) and (2) use  $\theta^{Sur}$  (in  $[0,1]$ ), while (3) and (4) use  $\theta^{Exp}$  to classify participants, where  $\theta^{Exp}=0$  for low-theta types and  $\theta^{Exp}=1$  for high-theta types. Perceived immorality is in  $[-1, 1]$  where -1 means very moral, 0 means neutral and 1 means very immoral. Control variables: age, gender, Swiss nationality, subject of study. Standard errors clustered at individual and industry/firm level (Cameron, Gelbach and Miller, 2011); z-statistics in parentheses; \*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$ .

**Table A12: Regressions of willingness to work for diverse industries and firms on perceived immorality and moral types, robustness checks for “I don’t know this organization”**

Dependent variable:	Willingness to work for firm $j$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<b>Perceived immorality (I(j))</b>	0.088 (1.48)	0.088 (1.47)	0.159*** (2.70)	0.159*** (2.71)	-0.123** (-2.21)	-0.123** (-2.20)	-0.092 (-1.49)	-0.092 (-1.46)
<b>Type survey (<math>\theta^{Sur}</math>)</b>	-0.127** (-2.01)	-0.165*** (-2.58)	-0.156 (-1.11)	-0.204 (-1.15)				
<b><math>\theta^{Sur} * I(j)</math></b>	-0.613*** (-7.46)	-0.613*** (-7.39)	-0.792*** (-5.49)	-0.792*** (-5.46)				
<b>Type experiment (<math>\theta^{Exp}</math>)</b>					-0.034 (-1.54)	-0.040* (-1.91)	-0.039 (-0.58)	-0.056 (-0.90)
<b><math>\theta^{Exp} * I(j)</math></b>					-0.131*** (-4.47)	-0.131*** (-4.43)	-0.158*** (-2.98)	-0.158*** (-2.87)
<b>N</b>	6162	6162	1352	1352	6162	6162	1352	1352
<b>Control variables</b>	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Coefficient estimates of linear regression models. Dependent variable: Willingness to work is in  $\{0, 0.25, 0.5, 0.75, 1\}$  where 0 means not at all willing to work, 0.5 means indifferent and 1 means really much willing to work. Observations where subjects did not fill out the questionnaire are excluded. Columns (1), (2), (5) and (6): Observations where subjects did not know the firm (“I don’t know this organization”) are classified as having willingness to work of 0.5. Columns (3), (4), (7) and (8): only participants that did know all firms ( $N=52$ ) are included. Independent variables: (1) - (4) use  $\theta^{Sur}$  (in  $[0,1]$ ), while (5) - (8) use  $\theta^{Exp}$  to classify participants, where  $\theta^{Exp}=0$  for low- theta types and  $\theta^{Exp}=1$  for high-theta types. Perceived immorality is in  $[-1, 1]$  where -1 means very moral, 0 means neutral and 1 means very immoral. Control variables: age, gender, Swiss nationality, subject of study. Standard errors clustered at individual and industry/firm level (Cameron, Gelbach and Miller, 2011); z-statistics in parentheses; \*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$ .

**Table A13: Regressions of willingness to work for diverse industries and firms on employment rate in the immoral work condition**

Dependent variable:	Willingness to work for industry $j$		Willingness to work for firm $j$	
	(1)	(2)	(5)	(6)
<b>Perceived immorality (I(j))</b>	-0.332*** (-6.64)	-0.399*** (-6.74)	-0.311*** (-6.42)	-0.310*** (-6.33)
<b>Employment rate (E)</b>	0.029 (1.07)	0.044* (1.79)	0.048 (1.36)	0.055 (1.54)
<b>E * I(j)</b>	0.073** (2.03)	0.073** (2.00)	0.114** (2.49)	0.114*** (2.44)
<b>N</b>	3275	3275	3561	3561
<b>Control variables</b>	No	Yes	No	Yes

Notes: Coefficient estimates of linear regression models. Sample includes only subjects from the immoral work condition. Dependent variable: Willingness to work is in  $\{0, 0.25, 0.5, 0.75, 1\}$  where 0 means not at all willing to work, 0.5 means indifferent and 1 means really much willing to work. Observations where subjects did not know the firm (“I don’t know this organization”) or did not fill out the questionnaire are excluded. Independent variables: Employment rate is the share of market periods in which the worker was employed. Perceived immorality is in  $[-1, 1]$  where -1 means very moral, 0 means neutral and 1 means very immoral. Control variables: age, gender, Swiss nationality, subject of study, average wage industry 2016 (SLFS; only for industries), industry size 2016 (STATENT; only for industries), industry sales 2015 (Value Added Tax Statistics; only for industries). Standard errors clustered at individual and industry/firm level (Cameron, Gelbach and Miller, 2011); z-statistics in parentheses; \*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$ .

## Appendix B – Robustness checks for analysis using Swiss Labor Force Survey

A potential critique of our analysis of wage premiums using the Swiss Labor Force Survey data in Section 4 is that we selected the potential immoral industries and the control industries ourselves. In this Appendix, we report results from two robustness checks that address this issue.

### *1) We use all industries in the Swiss Labor Force survey as control industries*

In a first robustness check, we use all non-immoral industries in the Swiss Labor Force survey as control industries. We do not have a measure of the perceived industry immorality,  $I(j)$ , for most industries. Instead of relying on such a measure, we define a set of industries as “immoral industries” and calculate wage premiums (or, wage discounts) for these industries, controlling for worker, job and industry characteristics. This approach is commonly used in the literature that studies stock returns for “sin industries” (e.g., Hong and Kacperczyk, 2009; Blitz and Fabozzi, 2017; Colonnello, Curatola and Gioffré, 2019). We expect immoral industries to pay a positive wage premium, a compensating differential for the immoral nature of the work.

*Set of immoral industries:* We select the set of immoral industries based on the industry ratings. Most participants that rated the immorality of the industries agreed that it is immoral to work in the following four industries: manufacture of weapons and ammunitions, manufacture of tobacco products, wholesale of tobacco products and gambling and betting activities (see Table A1). These four industries are also typically considered to be “sin industries” in the literature on sin stocks. We focus on these four industries.<sup>37</sup> Given that we use the industry ratings to select the immoral industries, this robustness check still depends on our selection of potentially immoral industries in Section 4. Note, however, that we use the entire dataset as control industries and therefore do not rely on our selection of control industries.

*Results:* Figure B1 (i) shows the average gross hourly wages for each of the four immoral industries and the average wage in all other industries. Table B1 reports regressions of the natural logarithm of real gross hourly wages on the dummies for working in each of the four

---

<sup>37</sup> In Section 4, we select two other industries, monetary intermediations and credit granting, that might potentially be perceived as immoral. As there was disagreement on the immorality of these two industries among survey respondents (see Table A1), we do not include them in the set of immoral industries. Note, however, that both industries pay substantial wage premiums. If we add dummies for working in these industries to Model 3, coefficients are 0.263 ( $t=11.12$ ,  $p<0.001$ ) for monetary intermediation and 0.282 ( $t=20.51$ ,  $p<0.001$ ) for credit granting.

immoral industries, along with several additional control variables (Model 1, 2 and 3). All four immoral industries pay substantial wage premiums, in line with an immorality premium for immoral work. According to Model 3, individuals working in the immoral industries have (geometric) mean hourly earnings of between 12 percent and 29 percent higher than people working in other industries.

2) *We elicited perceived industry immorality for a second set of industries*

As a second robustness check, we elicit the perceived industry immorality for a second set of industries, but in this case we do not select the industries. Instead, we ask research assistants unaware of our hypotheses to select moral and immoral industries. We then provide evidence for an immorality premium in this second set of industries.

*Set of industries:* We created a list with all industries that had at least 50 observations in the Swiss Labor Force Survey. This resulted in a list of 394 industries. We then asked five research assistants to select up to ten industries in which they think it is most immoral to work and up to ten industries in which they think it is most moral to work. They ranked the selected industries from most immoral (moral) to least immoral (moral). The research assistants were unaware of our research question. The five industries that the research assistants thought to be the most immoral and the five industries that they thought to be the most moral are included in the set of industries used in this robustness test.<sup>38</sup> In addition to these ten industries, we randomly selected a set of 40 other industries that had at least 50 observations in the Swiss Labor Force Survey. Table B2 gives all selected industries.

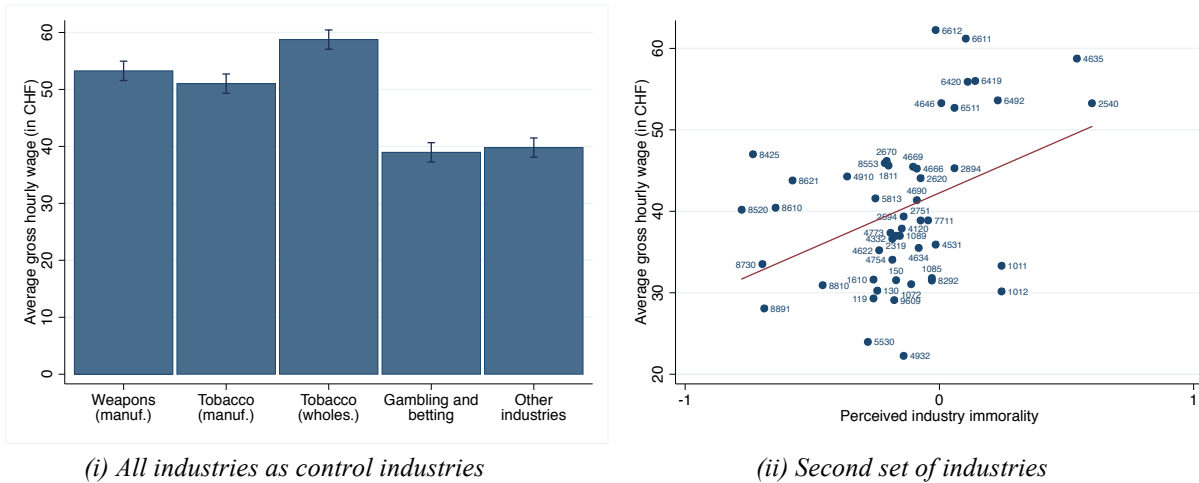
*Industry ratings:* We elicited a measure of perceived immorality for the set of 50 industries. We recruited 45 participants drawn from the same subject pool from which we recruit participants for our laboratory experiment (but that did not participate in our experiment). These participants rated how immoral they think it is to work for each of the 50 industries on a 7-point Likert scale ranging from very immoral to very moral. We re-scaled the responses to lie on the -1 to 1 interval. Table B2 gives the ratings for all industries.

---

<sup>38</sup> We calculated the average rank of each industry as follows. We first allocated points to each industry according to its rank: if an industry was rated the most immoral industry it received 10 points, the second most immoral industry received 9 points, and so on. We then added up points for every industry and selected the five immoral and the five moral industries with the highest number of points.

*Results:* Figure B1 (ii) shows the correlation between average industry wages and the new collected measure of perceived industry immorality. Table B1 reports regressions of the natural logarithm of real gross hourly wages on the new collected measure of perceived industry immorality, along with several additional control variables (Model 4 to 6).<sup>39</sup> We find a substantial and statistically significant immorality wage premium. According to Model 6, individuals working in an industry as immoral as manufacture of weapons and ammunition (i.e., Perceived immorality = 0.47) have (geometric) mean hourly earnings approximately 17 percent higher than people working in an industry with the median industry immorality rating (i.e., Perceived immorality = -0.14).<sup>40</sup>

**Figure B1: Correlation between wages and perceived industry immorality, robustness**



*Source:* Weighted data from the SLFS, years 2010-2016 (wage) and our own survey for the second set of industries (panel (ii), perceived industry immorality). *Notes:* Real gross hourly wage in 2010 CHF. Panel (i) gives the average gross hourly wages for Manufacture weapons and ammunitions (Weapons (manuf.)), manufacture tobacco (Tobacco (manuf.)), Wholesale tobacco (Tobacco (wholes.)), Gambling and betting and all other industries.  $N = 239,313$ . Panel (ii) gives the correlation between average industry wages and perceived industry immorality. Perceived immorality is in  $[-1, 1]$  where  $-1$  means very moral,  $0$  means neutral and  $1$  means very immoral. The numbers in the figure give the industry code of each dot (see Table B2).  $N = 47,935$ .

<sup>39</sup> Numbers of observations differ substantially between industries. If we weight observations by industry size (instead of using survey weights), estimates for perceived immorality are similar (Model 4: 0.342, Model 5: 0.281, Model 6: 0.266) and statistically significant ( $t=2.83$ ,  $t=3.21$ ,  $t=3.39$ , respectively).

<sup>40</sup> We obtain this number by doing the following calculation:  $e^{0.209 \cdot 0.74} - 1 \approx 0.167$ .

**Table B1: Relationship between wages and industry immorality, robustness**

Dependent variable: ln of real gross hourly wage (in 2010 CHF)						
	(1)	(2)	(3)	(4)	(5)	(6)
Manufacture weapons and ammunitions	0.345*** (14.98)	0.154*** (10.31)	0.147*** (8.24)			
Manufacture tobacco	0.285*** (12.37)	0.237*** (7.61)	0.288*** (9.16)			
Wholesale tobacco	0.402*** (17.47)	0.327*** (23.49)	0.280*** (16.29)			
Gambling and betting	0.025 (1.08)	0.087*** (7.26)	0.123*** (9.44)			
Perceived industry immorality				0.327** (2.34)	0.242*** (2.94)	0.209*** (2.77)
Age		0.006*** (12.66)	0.006*** (12.58)		0.008*** (7.20)	0.007*** (6.94)
Male		0.194*** (17.26)	0.195*** (17.81)		0.152*** (7.50)	0.155*** (7.62)
Married		0.035*** (4.57)	0.039*** (5.00)		0.030 (1.26)	0.032 (1.33)
Education high		0.586*** (31.24)	0.564*** (32.32)		0.602*** (13.69)	0.580*** (12.91)
Education middle		0.247*** (17.07)	0.241*** (16.99)		0.297*** (6.93)	0.288*** (6.60)
Swiss		0.0024 (0.20)	0.011 (1.01)		-0.011 (-0.63)	-0.002 (-0.11)
Experience		0.004*** (5.75)	0.005*** (6.24)		0.004** (2.49)	0.005*** (2.77)
Full-time equivalent		-0.076*** (-2.59)	-0.078*** (-2.73)		-0.166** (-2.35)	-0.168** (-2.44)
Managerial duties		-0.025 (-0.91)	-0.020 (-0.74)		0.063 (0.85)	0.066 (0.93)
Industry sales		0.006** (2.15)	0.005** (2.04)		0.020 (1.17)	0.021 (1.22)
Industry size (employees)		0.0004 (1.59)	0.0004* (1.83)		0.001*** (3.80)	0.001*** (4.08)
Constant	3.561*** (154.57)	2.832*** (79.63)		3.737*** (52.83)	2.876*** (32.09)	
N	239,313	236,625	236,625	47,935	47,935	47,935
Adjusted R <sup>2</sup>	0.001	0.206	0.221	0.041	0.248	0.263
Year and Region FE	No	No	Yes	No	No	Yes

Source: Weighed data from the SLFS, years 2010-2016 (wage and demographics), STATENT, years 2011-2016 (industry size, industry sales), Value Added Tax Statistics, years 2010-2016 (industry sales) and our own online-survey (perceived industry immorality). Notes: Manufacture weapons and ammunitions, Manufacture tobacco and Gambling and betting are binary variables where 1 means that the individual works in the respective industry. Perceived immorality is in [-1, 1] where -1 means very moral, 0 means neutral and 1 means very immoral. Control variables: Male in {0, 1}, Married in {0, 1}, Education high: higher vocational education and training or university/college, Education middle: apprenticeship, full-time vocational school, matura or pedagogical training, Education low (reference category): compulsory schooling or pre-vocational education, Swiss in {0, 1}, Experience = number of years in the firm, Full-time equivalent = (working hours /42), set to 1 for working hours >= 42, managerial duties in {0, 1}, Industry size = number employees in this industry / 1000 (2010 data is not available, we substitute it with 2011 data), Industry sales = Industry sales/number employees in this industry. Model 3 and 6 control for company region fixed effects (26 Swiss cantons) and year fixed effects (2010-2016). Standard errors clustered at the industry level, t-statistics in parentheses; \* p < 0.1; \*\* p < 0.05; \*\*\* p < 0.01.



**Table B2: Set of industries and industry ratings**

	Industry	Industry Code	Perceived immorality
Immoral industries	Manufacture of weapons and ammunition	2540	0.60
	Wholesale of tobacco products	4635	0.54
	Processing and preserving of meat (except poultry meat)	1011	0.24
	Credit granting	6492	0.23
	Processing and preserving of poultry meat	1012	0.24
Moral industries	Social work activities without accommodation for the elderly and disabled	8810	-0.46
	Residential care activities for the elderly and disabled	8730	-0.70
	Fire service activities	8425	-0.73
	Primary education	8520	-0.78
	Hospital activities	8610	-0.64
Other Industries	Manufacture of prepared meals and dishes	1085	-0.03
	Wholesale of office machinery and equipment, except computers and computer peripheral equipment	4666	-0.09
	Publishing of newspapers	5813	-0.25
	Monetary intermediation (cantonal banks, commercial banks, stock exchange banks, private bankers; banks with a special field of business; regional banks; Raiffeisen banks; Foreign-controlled banks)	6419	0.14
	Passenger rail transport	4910	-0.36
	Support activities for crop production (preparation of fields; establishing a crop; treatment of crops; crop spraying; trimming of fruit trees and vines; transplanting of rice; thinning of beets; harvesting; pest control; provision of agricultural machinery with operators and crew)	1610	-0.26
	Driving school	8553	-0.21
	Security and commodity contracts brokerage	6612	-0.01
	Printing of newspapers	1811	-0.2
	Growing of other non-perennial crops (growing of swedes, mangolds, fodder roots, clover, alfalfa, sainfoin, fodder maize and other grasses; buckwheat; potted and bedding plants; beet seeds (excluding sugar beet seeds); seeds of forage plants and flower seeds; forage kale and similar forage products; production of cut flowers)	119	-0.26
	Plant propagation	130	-0.24
	Packaging activities (bottling of liquids; packaging of solids; security packaging of pharmaceutical preparations; labelling, stamping and imprinting; parcel-packing and gift-wrapping)	8292	-0.03
	Manufacture of fasteners and screw machine products	2594	-0.14
	Construction of residential and non-residential buildings	4120	-0.15
	Manufacture of machinery for textile, apparel and leather production	2894	0.06
	General medical practice activities	8621	-0.58
	Activities of holding companies	6420	0.11

*See next page for the rest of the table.*

	Industry	Industry Code	Perceived immorality
Other Industries	Retail sale of electrical household appliances in specialised stores	4754	-0.19
	Wholesale trade of motor vehicle parts and accessories	4531	-0.01
	Wholesale of flowers and plants	4622	-0.24
	Dispensing chemist in specialised stores	4773	-0.19
	Administration of financial markets	6611	0.10
	Manufacture of other food products (soups and broths; artificial honey and caramel; perishable prepared foods; food supplements; yeast; extracts and juices; non-dairy milk and cheese substitutes; egg products; artificial concentrates)	1089	-0.16
	Other personal service activities (astrological and spiritualists' activities; social activities; pet care services; genealogical organisations; tattooing and piercing studios; shoe shiners; porters; valet car parkers; concession operation of coin-operated personal service machines)	9609	-0.18
	Manufacture of computers and peripheral equipment	2620	-0.07
	Joinery installation	4332	-0.19
	Wholesale of beverages	4634	-0.08
	Camping grounds, recreational vehicle parks and trailer parks	5530	-0.28
	Manufacture of optical instruments and photographic equipment	2670	-0.21
	Renting and leasing of cars and light motor vehicles	7711	-0.04
	Wholesale of other machinery and equipment (transport equipment except motor vehicles; production-line robots; wires and switches; other electrical material; machinery for use in trade, navigation and industry [except mining, construction, civil engineering and textile industry]; measuring instruments and equipment)	4669	-0.10
	Non-specialised wholesale trade	4690	-0.09
	Mixed Farming	150	-0.17
	Manufacture of electric domestic appliances	2751	-0.07
	Life insurance	6511	0.06
	Manufacture and processing of other glass, including technical glassware (laboratory, hygienic or pharmaceutical glassware; clock or watch glasses, optical glass and optical elements not optically worked; glassware used in imitation jewellery; glass insulators and glass insulating fittings; glass envelopes for lamps; glass figurines; glass paving blocks; glass in rods or tubes)	2319	-0.17
	Wholesale of pharmaceutical goods	4646	0.01
Taxi operation	4932	-0.14	
Child day-care activities	8891	-0.69	
Manufacture of rusks and biscuits; manufacture of preserved pastry goods and cakes	1072	-0.11	

*Notes: Immoral (Moral) industries are the industries that the research assistants selected as the most immoral (moral) industries. Other industries are 40 randomly selected industries that have at least 50 observations in the Swiss Labor Force Survey.*

## Appendix C – A simple model of heterogeneous moral concern in labor markets

In this section, we introduce a simple stylized model of labor markets with varying degrees of perceived job immorality and heterogeneity in concern for moral behavior among workers.

We examine a single labor market for a job,  $j \in J$ , which might involve doing immoral work. Firms decide whether to hire a worker to do  $j$  at the market wage,  $w$ , and workers decide whether to accept work  $j$  for the market wage. Workers differ in their concerns for morality. We investigate how the equilibrium wage and selection in this labor market change with the immorality of  $j$ . Our framework is a simplification of the theoretical literature on compensating wage differentials (see, e.g., Rosen, 1986).<sup>41</sup> We do not seek to expand this literature, but rather to apply it to a context in which the relevant job dimension is immorality.

The immorality of  $j$  is measured by a function  $I: J \rightarrow [0, \infty)$ , where  $I(j') > I(j)$  means that job  $j'$  is more immoral than job  $j$ , and  $I(j) = 0$  means that  $j$  involves no immoral acts. The set of immoral jobs is  $J^{IM} = \{j \in J: I(j) > 0\}$ .

Labor demand is represented by an interval of firms,  $k \in [0,1]$ . Firms' behavior is given by the labor demand function,  $D: \mathbb{R} \times J \rightarrow [0,1]$ , with  $\lim_{w \rightarrow \infty} D(w, j) = 0$ ,  $D(w, j) = 1$  for  $w \leq 0$ ,  $D$  continuous in  $w$  and  $D$  strictly decreasing in  $w$  on  $[0, \infty)$ . In addition, we assume that an increase in the immorality of the job does not decrease the profitability of employing labor, that is,  $I(j') > I(j)$  implies  $D(w, j') \geq D(w, j)$  for all  $w$ .<sup>42</sup>

Labor supply consists of an interval of workers,  $i \in [0,1]$ . Each worker has reservation utility  $\underline{u} \geq 0$ , and the utility of accepting job  $j$  for a worker of type  $i$  is given by:<sup>43</sup>

$$u_i^{accept}(j, w) = w - c - \theta_i * I(j),$$

---

<sup>41</sup> Unlike most models of compensating wage differentials, we do not have multiple labor markets, rather one and a fixed outside option. In our laboratory experiment, we also assign subjects to one labor market. This abstraction simplifies both the theory and the experiment. However, we show in Appendix D that the model allows for an interpretation with two jobs, an immoral job and a neutral job. Our results also apply to such a context.

<sup>42</sup> In our experiment, we vary the immorality of the job, but fix labor demand, that is,  $D(w, j') = D(w, j)$  for all  $w$  and all  $j, j' \in J$ . If an increase in immorality were to decrease profitability, there would be no incentives for firms to operate in immoral industries. Heidhues, Kőszegi and Murooka (2017) provide a basis for why deceptively marketed socially harmful products may be more profitable in the presence of naïve consumers. In Appendix D, we provide a behavioral foundation for the labor demand.

<sup>43</sup> Models about “mission-oriented” employees commonly assume very similar additive utility functions (e.g. Cassar and Meier, 2018), with the main difference that  $-I(j) * \theta_i$  is replaced by a positive term, the “meaningfulness of work” multiplied by how much the individual cares about meaning.

where  $c \geq 0$  is the worker's cost of effort, which is independent of  $j$ . The parameter  $\theta_i \geq 0$  measures the worker's aversion to immoral work and is distributed according to a cumulative density function  $F \in \mathcal{F}_\theta$ . The set  $\mathcal{F}_\theta$  consists of all density functions  $F$  that are continuous, strictly increasing on  $[0, \infty)$ , and with  $F(0) = 0$ , meaning that no worker likes immoral work.<sup>44</sup> The indirect utility of a worker of type  $i$  is then given by  $v_i(j, w) = \max\{\underline{u}, u_i^{accept}(j, w)\}$ . Workers' behavior determines the labor supply,  $S: \mathbb{R} \times J \rightarrow [0, 1]$ . If  $j \in J^{IM}$ , every worker with  $\theta_i \leq \frac{w - \underline{u} - c}{I(j)}$  accepts the job. Labor supply is therefore  $S(w, j) = F\left(\frac{w - \underline{u} - c}{I(j)}\right)$ .<sup>45</sup>

We next consider the equilibrium properties of this type of market. The equilibrium wage,  $w^*(j)$ , is implicitly defined by  $S(w^*(j), j) - D(w^*(j), j) = 0$ .<sup>46</sup> The following Lemma states that for every  $j \in J$ ,  $w^*(j)$  exists and is unique.

**Lemma.** *For all  $j \in J^{IM}$ ,  $w^*(j)$  exists, is unique and is in  $(\underline{u} + c, \infty)$ . For all  $j \in J \setminus J^{IM}$ ,  $w^*(j) = \underline{u} + c$ .*

**Proof.** Suppose  $j \in J^{IM}$ . *Existence:* Define  $f(w, j) = S(w, j) - D(w, j)$ . Note that  $f(\underline{u} + c, j) = 0 - (+) < 0$ ,  $\lim_{w \rightarrow \infty} f(w, j) = 1 - 0 = 1$  and  $f(w, j)$  is continuous in  $w$ . By the intermediate value theorem there exists  $w^*(j) \in (\underline{u} + c, \infty)$  such that  $f(w^*(j), j) = 0$ .

*Uniqueness:* Follows from  $f(w, j)$  being strictly increasing in  $w$  on  $[0, \infty)$ .

Suppose  $j \in J \setminus J^{IM}$ . Then,  $S(w, j) = \begin{cases} 0, & w < \underline{u} + c \\ [0, 1], & w = \underline{u} + c \\ 1, & w > \underline{u} + c \end{cases}$ . Note that for any  $w < \underline{u} + c$ , we have

$D(w, j) > 0$  but  $S(w, j) = 0$ , and for any  $w > \underline{u} + c$  we have  $D(w, j) < 1$ , but  $S(w, j) = 1$ . For  $w = \underline{u} + c$ ,  $S(w, j) = [0, 1]$  and  $D(w, j) \in [0, 1]$ , so  $D(w, j) \in S(w, j)$ . ■

In the following, we derive four properties of labor markets with immoral jobs. While straightforward, we use these results to make predictions for our empirical work. In particular, the first two propositions derive the primary hypotheses that we test across all of our analysis.

<sup>44</sup> Note that  $F(0) = 0$  implies that no worker likes to do immoral jobs.

<sup>45</sup> The assumptions on  $F$  (together with the properties of a cdf) imply that  $S$  is continuous and strictly increasing in  $w$  on  $[\underline{u} + c, \infty)$ ,  $\lim_{w \rightarrow \infty} S(w, j) = 1$ , and  $S(w, j) = 0$  for all  $w \leq \underline{u} + c$ .

<sup>46</sup> Note that for  $j \in J \setminus J^{IM}$ ,  $S$  is a correspondence. For this case,  $w^*(j)$  is defined by  $D(w^*(j), j) \in S(w^*(j), j)$ . Moreover,  $w^*(j)$  depends on  $F$ . When necessary (Corollary, Proposition 4) we will make this explicit by writing  $w^*(j, F)$  instead of  $w^*(j)$ .

Proposition 1 shows that there is an immorality premium for immoral jobs: an increase in the immorality of a job decreases supply and therefore increases the equilibrium wage.

**Proposition 1. (Immorality premium)** *For all  $j, j' \in J$  with  $I(j) < I(j')$ ,  $w^*(j) < w^*(j')$ .*

**Proof.**  $w^*(j) < w^*(j')$ : Suppose  $I(j) = 0$ . Then,  $w^*(j) = \underline{u} + c$  and  $w^*(j') > \underline{u} + c$  (see Lemma). Suppose  $I(j) > 0$ . Suppose that  $w^*(j) \geq w^*(j')$ . Then  $S(w^*(j), j) > S(w^*(j'), j')$  and  $D(w^*(j), j) \leq D(w^*(j'), j)$  because  $F$  and  $-D$  are strictly increasing in  $w$  on  $[0, \infty)$ ,  $I(j) < I(j')$  and therefore  $D(w, j') \geq D(w, j)$  for all  $w$ . So  $S(w^*(j), j) - S(w^*(j'), j') + D(w^*(j'), j') - D(w^*(j), j) > 0$ , a contradiction to the definition of  $w^*(j)$  and  $w^*(j')$ . ■

The following Corollary further shows that this wage premium will be insignificant if workers do not sufficiently care about morality.

**Corollary.** *For all  $j, j' \in J$  with  $I(j) < I(j')$  and  $\varepsilon > 0$ , there exists  $G \in \mathcal{F}_\theta$  such that  $w^*(j', G) - w^*(j, G) \leq \varepsilon$ .*

**Proof.** Suppose that there exists a  $G \in \mathcal{F}_\theta$  such that

$$G\left(\frac{\varepsilon}{I(j')}\right) = D(\underline{u} + c + \varepsilon, j').$$

Then,  $w^*(j', G) = \underline{u} + c + \varepsilon$ . The Lemma and Proposition 1 then imply that  $w^*(j, G) \in [\underline{u} + c, \underline{u} + c + \varepsilon)$ , and, as a result,  $w^*(j', G) - w^*(j, G) \leq \varepsilon$ .

To prove that such a  $G \in \mathcal{F}_\theta$  exist, take any  $H \in \mathcal{F}_\theta$  and construct  $G$  as follows:

$$G(x) = \begin{cases} 0 & \text{if } x < 0 \\ x \frac{I(j')}{\varepsilon} D(\underline{u} + c + \varepsilon, j') & \text{if } x \in [0, \frac{\varepsilon}{I(j')}] \\ D(\underline{u} + c + \varepsilon, j') + \left(1 - D(\underline{u} + c + \varepsilon, j')\right) H\left(x - \frac{\varepsilon}{I(j')}\right) & \text{if } x > \frac{\varepsilon}{I(j')} \end{cases}$$

The assumptions on  $D$  imply that  $D(\underline{u} + c + \varepsilon, j') \in (0, 1)$ . Note that  $G$  is continuous, strictly increasing on  $[0, \infty)$ , and with  $G(0) = 0$ . Therefore  $G \in \mathcal{F}_\theta$ . ■

Formally, the Corollary shows that there are distributions of moral types such that the wage differentials are arbitrary small.

Second, the types that care least about the immorality of a job ( $\theta_i \leq \frac{w^*(j)-\underline{u}-c}{I(j)}$ ), sort into accepting immoral jobs, while those more concerned with morality ( $\theta_i > \frac{w^*(j)-\underline{u}-c}{I(j)}$ ), refuse to do the job for the equilibrium wage.<sup>47</sup> This is formally shown in Proposition 2.

**Proposition 2. (Sorting)** For all  $j \in J^{IM}$ , worker  $i$  is hired iff  $\theta_i \leq \frac{w^*(j)-\underline{u}-c}{I(j)} \equiv \underline{\theta}(j) > 0$ .

**Proof.** A worker accepts job  $j$  iff  $u_i^{accept} = w^*(j) - c - \theta_i * I(j) \geq \underline{u} \Leftrightarrow \theta_i \leq \frac{w^*(j)-c-\underline{u}}{I(j)}$ .  
 $\underline{\theta}(j) > 0$ : Follows from  $w^*(j) > \underline{u} + c$  (see Lemma). ■

Proposition 2 is critical to the notion that immorality wage premiums are driven by those who find immoral work most distasteful opting out of such jobs. This important relationship has, to our knowledge, not been previously empirically tested.

Our next two propositions are less central, but nevertheless provide some useful and testable insights into behavior in immoral labor markets. Proposition 3 shows that the most immoral types profit from an increase in the immorality of work.

**Proposition 3.** For all  $j, j' \in J$  with  $I(j) < I(j')$ , there exists  $\tilde{\theta}(j, j') > 0$  such that  $v_i(j', w^*(j')) > v_i(j, w^*(j))$  iff  $\theta_i < \tilde{\theta}(j, j')$ .

**Proof.** Suppose  $I(j) = 0$ . Then  $v_i(j, w^*(j)) = v_i(j, \underline{u} + c) = \underline{u}$  (see Lemma). Note that  $u_i^{accept}(j', w^*(j')) > \underline{u}$  iff  $\theta_i < \frac{w^*(j')-\underline{u}-c}{I(j')} = \underline{\theta}(j')$ , so  $v_i(j', w^*(j')) = u_i^{accept}(j', w^*(j')) > \underline{u} = v_i(j, w^*(j))$  iff  $\theta_i < \underline{\theta}(j') \equiv \tilde{\theta}(j, j')$ . Note that  $\tilde{\theta}(j, j') > 0$  by Proposition 2.

Suppose  $I(j) > 0$ . Note that  $v_i(j', w^*(j')) > v_i(j, w^*(j)) = \max\{u_i^{accept}(j, w^*(j)), \underline{u}\}$  if and only if:

- i)  $u_i^{accept}(j', w^*(j')) > u_i^{accept}(j, w^*(j))$
- ii)  $u_i^{accept}(j', w^*(j')) > \underline{u}$

<sup>47</sup> This perfect sorting according to  $\theta$  is an extreme case. Heterogeneity in the costs of effort, reservation utility or productivity implies partial sorting according to  $\theta$  (see Garen (1988) and Hwang, Reed and Hubbard (1992) for related examples). We incorporate only one dimension of heterogeneity in the model for simplicity.

Inequality i) holds iff  $\theta_i < \frac{w^*(j')-w^*(j)}{I(j')-I(j)}$  and inequality ii) holds iff  $\theta_i < \underline{\theta}(j')$ . Therefore, we have

$$\tilde{\theta}(j, j') = \min \left\{ \frac{w^*(j')-w^*(j)}{I(j')-I(j)}, \underline{\theta}(j') \right\}. \text{ Note that } \tilde{\theta}(j, j') > 0 \text{ (see Proposition 1 and}$$

Proposition 2). ■

More precisely, there are always some types who are sufficiently unconcerned with morality who are hired in a market for immoral work and are overcompensated by the immorality premium.

Finally, Proposition 4 shows that moral behavior (refusing to do the immoral job) can have positive externalities for the most immoral types. If the distribution of types shifts toward more concern for morality—in the sense of stochastic dominance—then the supply at any given wage decreases, thereby increasing the equilibrium wage and the utility of those least concerned with morality. For instance, any completely immoral types (i.e.,  $\theta_i = 0$ ) will always benefit from a higher wage produced by increased moral concerns.

**Proposition 4.** *For all  $j \in J^{IM}$  and  $F, G \in \mathcal{F}_\theta$  with  $F(x) < G(x)$  for all  $x > 0$ , there exists  $\hat{\theta}(j, F) > 0$  such that  $v_i(j, w^*(j, F)) > v_i(j, w^*(j, G))$  iff  $\theta_i < \hat{\theta}(j, F)$ .*

**Proof.** First, we will proof that  $w^*(j, G) \leq w^*(j, F)$ . Suppose not, then  $w^*(j, G) > w^*(j, F)$ . But then  $S(w^*(j, G), j, G) > S(w^*(j, F), j, F)$  and  $D(w^*(j, G), j) < D(w^*(j, F), j)$  because  $F, G$  and  $-D$  are strictly increasing in  $w$  on  $[0, \infty)$ , and  $F$  first-order stochastically dominates  $G$ . But then  $S(w^*(j, G), j, G) - S(w^*(j, F), j, F) + D(w^*(j, F), j) - D(w^*(j, G), j) > 0$ , a contradiction to the definition of  $w^*$ .

Second, define  $\hat{\theta}(j, F) = \frac{w^*(j, F) - \underline{u} - c}{I(j)} \equiv \underline{\theta}(j, F)$ , and note that under this definition  $w^*(j, F) - c - S(j) * \theta_i > \underline{u}$  iff  $\theta_i < \hat{\theta}(j, F)$ . ( $\hat{\theta}(j, F) > 0$  follows from Proposition 2.)

To finish the proof, note that  $v_i(j, w^*(j, F)) = w^*(j, F) - c - I(j) * \theta_i > \max\{\underline{u}, w^*(j, G) - c - I(j) * \theta_i\} = v_i(j, w^*(j, G))$  for all  $\theta_i < \hat{\theta}(j, F)$ , and  $v_i(j, w^*(j, F)) = \underline{u} = v_i(j, w^*(j, G))$  for all  $\theta_i \geq \hat{\theta}(j, F)$ . ■

## Appendix D – Alternative model interpretation

The results in Section 3 and Appendix C also apply for a context with 2 jobs, a neutral job  $j^N$  ( $I(j^N) = 0$ ) and an immoral job  $j^{IM} \in J^{IM}$  ( $I(j^{IM}) > 0$ ). In the following, we show that, under some assumptions, labor demand and labor supply correspond to their counterparts in Appendix C. Therefore, all results derived in Appendix C also hold in the context with 2 jobs.

*Labor supply:* Labor supply consists of an interval of workers,  $i \in [0,1]$ . As in Appendix C, we assume that the utility of accepting job  $j$  of a worker of type  $i$  is given by:

$$u_i(j, w(j)) = w(j) - c - \theta_i * I(j),$$

where the parameter  $\theta_i$  is distributed according to a distribution with cdf,  $F \in \mathcal{F}_\theta$ . The set  $\mathcal{F}_\theta$  consists of all density functions  $F$  that are continuous, strictly increasing on  $[0, \infty)$ , and with  $F(0) = 0$ . Workers choose between the neutral and the immoral job. Note that every worker with  $\theta_i \leq \frac{w(j^{IM}) - w(j^N)}{I(j^{IM})}$  chooses the immoral job. The labor supply for the immoral job is then given by  $S(w, j^{IM}) = F(\frac{w}{I(j^{IM})})$ , where  $w = w(j^{IM}) - w(j^N)$  is the immorality premium. Note that the labor supply for the immoral job corresponds to the labor supply in Appendix C with  $\underline{u} = c = 0$ .

*Labor demand:* Labor demand consists of an interval of firms,  $k \in [0,1]$ . Each firm can either produce a neutral product or an immoral product. Firms that produce immoral products offer immoral jobs; firms that produce neutral products offer neutral jobs. Firms' profits are:

$$\pi_k(j, w) = a_k(j) - w(j),$$

where  $a_k(j)$  measures firm  $k$ 's earnings when producing good  $j$ . Firms choose to produce the immoral product if  $\Delta a_k(j^{IM}) = a_k(j^{IM}) - a_k(j^N) \geq w$ . This term,  $\Delta a_k(j^{IM})$ , is distributed according to a distribution with cdf  $G_{j^{IM}}$ . An increase in immorality of the job does not decrease firms earnings,<sup>48</sup> that is, i)  $G_j(0) = 0$  for all  $j \in J^{IM}$ , and ii) for all  $j, j' \in J^{IM}$  with  $I(j') > I(j)$ ,  $G_{j'}(x) \leq G_j(x)$  for all  $x \in \mathbb{R}$ . In addition,  $G_{j^{IM}}$  is continuous and strictly increasing on  $[0, \infty)$ . The labor demand for the immoral job is then given by  $D(w, j^{IM}) = 1 - G_{j^{IM}}(w)$ . Note that  $D$  is

---

<sup>48</sup> One interpretation is, for example, that  $I(j)$  measures negative externalities in production. Avoiding these externalities is costly; decreasing the immorality therefore increases production costs (see also Rosen, 1986).



continuous in  $w$ , strictly decreasing in  $w$  on  $[0, \infty)$ , with  $\lim_{w \rightarrow \infty} D(w, j^{IM}) = 0$  and  $D(w, j^{IM}) = 1$  for  $w \leq 0$ . In addition,  $I(j') > I(j)$  implies  $D(w, j') \geq D(w, j)$  for all  $w \in \mathbb{R}$ . Note that the labor demand satisfies all assumptions made for the labor demand in Appendix C.

The equilibrium wage,  $w^*(j^{IM})$ , is implicitly defined by  $S(w^*(j^{IM}), j^{IM}) - D(w^*(j^{IM}), j^{IM}) = 0$ .<sup>49</sup> As both labor demand and labor supply satisfy all assumptions made in Appendix C, the Lemma, Corollary and Proposition 1 to 4 apply (with  $j \in J^{IM}$ ). In particular,  $w^*(j^{IM})$  is strictly positive (Lemma), so there is an immorality premium, and this immorality premium is increasing in the immorality of  $j^{IM}$ ,  $I(j^{IM})$  (Proposition 1). The immoral types sort into accepting the immoral jobs, while the moral types sort into accepting the neutral jobs (Proposition 2).

---

<sup>49</sup> Note that market clearance in the immoral job market implies market clearance in the neutral job market,  $(1 - S(w^*(j^{IM}), j^{IM})) - (1 - D(w^*(j^{IM}), j^{IM})) = 0$ .