

Fairness and motivation

Fair treatment creates incentives, and is beneficial for workers and the firm

Keywords: fairness, wages, effort, reciprocity, backfiring incentives

ELEVATOR PITCH

How do firms motivate their employees to be productive? The conventional wisdom is that workers respond to monetary incentives—“Pay them more and they will work harder.” However, a large and growing body of empirical evidence from laboratory and field experiments, surveys, and observational data, as well as neuroeconomic research, suggests that workers’ perceptions of fairness and trust are also key drivers of their work effort. Treating employees with respect is not only ethically warranted, it can create positive economic outcomes for both the worker and the firm.

KEY FINDINGS

Pros

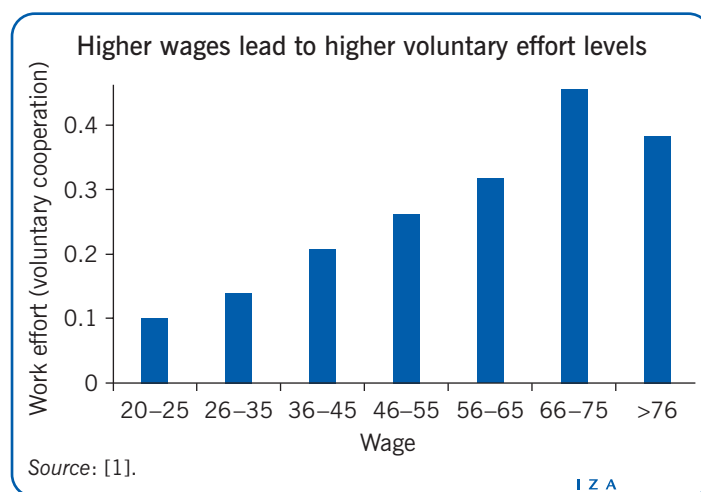
- + Kind and fair treatment can help employers to motivate their workers.
- + Offering sufficiently high wages may elicit effort levels above the contractually enforceable minimum.
- + Unfairness is punished, even if punishing is costly.
- + “Trust pays,” since many workers are intrinsically trustworthy and react negatively to controlling work environments.
- + Controlling devices in the workplace and particular forms of explicit performance incentives may entail “hidden costs.”

Cons

- Employment contracts are typically incomplete and cannot fully resolve the conflicting interests of workers and firms.
- Monetary incentives are less effective when offers are perceived as unfair.
- Treating agents in an unfair or unkind manner can evoke counterproductive activities such as shirking or sabotage.
- Unfair payment can lead to physiological stress and adverse health consequences.
- Controlling may give rise to a self-fulfilling prophecy of distrust between workers and firms.

AUTHOR’S MAIN MESSAGE

Firms have traditionally tried to motivate workers by setting explicit pay-for-performance incentives. However, new research has indicated that workers are motivated not just by wages per se, but also by their perceptions of how fair their wages are. Firms that emphasize the fair treatment of workers with regard to wages and other workplace conditions may produce better outcomes for both themselves and their workers.



MOTIVATION

One of the most important challenges that organizations and firms face is the problem of motivating employees. The goal to be achieved is for workers to be cooperative, take the initiative, identify and solve problems, and associate with the objectives of their organization. Organizational theorists, as well as personnel managers, often stress the importance of a worker's general job attitude. For example, it has been argued that "consummate cooperation"—defined as "an affirmative job attitude whereby gaps are filled, initiative is taken, and judgment is exercised in an instrumental way"—is of central importance to organizations [2]. Thus, solving the motivation problem and facilitating consummate cooperation are vital for the success of any firm or organization.

The motivation problem originates in the generic conflict of interest between workers and organizations. Firms want their employees to work hard and to pay them as little as possible, while employees want to maximize their pay and expend as little effort as possible. The conflict could easily be resolved by writing complete, performance-contingent contracts. However, employment contracts are typically incomplete, because the level of exerted effort is not necessarily verifiable by the employer and is therefore not third-party enforceable (for example, by arbitrators). As a consequence, labor contracts often just stipulate a wage payment and leave out many details that actually determine performance.

The traditional economic approach to mitigate the motivation problem is to specify explicit pay-for-performance incentives. This approach is based on the premise that most people are self-interested and predominantly motivated by money. This paper discusses an alternative view that is based on the assumption that many people have "social preferences" and are, therefore, motivated by notions of fairness, reciprocity, and trust. A person exhibits social preferences if he cares not only about his own material resources and well-being, but also about those of others. A large and growing body of empirical evidence from laboratory and field experiments, studies based on survey and observational data, and neuroeconomic research suggests the importance of social preferences.

The main implication of this new evidence is that the traditional economic view—that people are predominantly selfish and interested in maximizing monetary payoffs—is too narrow and psychologically incomplete. Social factors such as fairness, trust, or reciprocity are key drivers of human motivation, implying two important consequences for the motivation problem. First, there exist alternative ways to motivate employees in addition to performance-based monetary incentives. Second, the existence of social preferences implies that monetary incentives that are optimal in a standard economic framework may be less effective, or may even backfire, in a real-world setting.

DISCUSSION OF PROS AND CONS

Reciprocity and motivation

Reciprocity means that people reward kind actions and punish unkind ones [3], [4]. In this sense, reciprocity is a contingent social preference, because, depending on the behavior of the employer, a reciprocal worker values the employer's material payoff either positively or negatively. More specifically, if the worker perceives the actions of the employer as kind, the worker values the employer's payoff positively. If, in contrast,

the employer's actions are perceived as hostile, the worker values the employer's payoff negatively. Whether an action is perceived as kind or hostile depends on the consequences and the fairness or unfairness of the intention underlying the action.

Intuitively, the existence of reciprocity can help the employer to elicit high effort levels (i.e. to motivate workers) if he treats workers in a kind and fair way. One way of testing the notion of reciprocity for motivating individuals is the so-called gift exchange game [5]. The first action in the gift exchange game is that the employer determines the wage level. The worker then chooses to accept the wage and decides how much effort to provide, or the worker chooses not to accept the wage and does not work for the employer. The employer wants the worker to provide a high level of effort, but the worker faces effort costs, which increase along with the level of effort provided.

The gift exchange game

In its most simple form, the gift exchange game implements a two-stage, two-player game with one subject in the role of a principal (employer) and one in the role of an agent (worker).

In the first stage, the principal can make a binding wage offer w to an agent, who can, but does not have to, accept it. If an agent accepts the offer, he/she has to determine the effort level e .

In the experiment, the choice of effort level is represented by the choice of a number. The higher the chosen number, the higher the effort level, and the higher the monetary effort costs to be borne by the agent. The higher e , the larger the material payoff for the principal, but the higher also the agent's effort costs $c(e)$, which are increasing and convex.

Material payoffs from an exchange are given by $ae-w$ for the principal and $w-c(e)$ for the agent, where a is a constant. If the principle and his/her agent do not manage to trade, they both earn zero.

The gift-exchange game is one way of testing the notion of reciprocity for motivating individuals (see Fehr et al., 1993).

Fehr, E., G. Kirchsteiger, and A. Riedl. "Does fairness prevent market clearing? An experimental investigation." *The Quarterly Journal of Economics* 108:2 (1993): 437-459.

Thus, the gift exchange game represents the conflict of interest described above: While employees (materially) prefer a high wage and a low level of effort, the reverse holds true for employers. If workers choose to provide the maximum level of effort, the payoff to employers is maximized. However, under the standard economic assumption that people maximize their own material payoffs, workers will provide the lowest possible level of effort, regardless of the wage paid and the level of effort desired by employers. Expecting this, employers have no incentive to offer above-minimum wages. This "selfish" equilibrium prediction therefore implies low payoffs for workers **and** employers, and low efficiency.

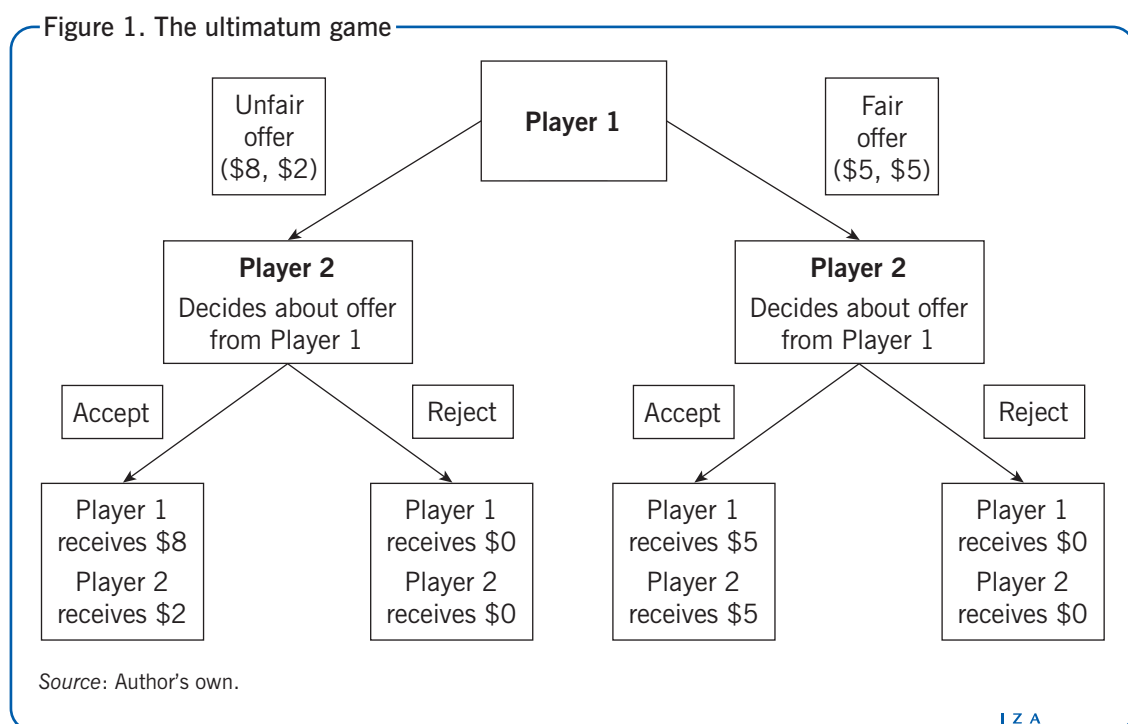
The predicted outcome in the presence of reciprocal motives is quite different. If workers perceive high wage payments as fair or kind, they will respond in kind by providing above-minimum levels of effort, even though providing above-minimal effort is costly

to them. Thus, by offering sufficiently high wages, employers may be capable of eliciting effort levels above the minimum. This “reciprocity” equilibrium prediction therefore implies higher payoffs for workers and employers, and greater efficiency.

The main finding of the gift exchange game supports the latter prediction: Most employers decide to offer relatively high wages to their workers, and on average workers reciprocate by providing an above-minimal level of effort. This gives rise to a positive wage–effort relation, or an efficiency wage: The effort exerted by reciprocal workers increases with the wage they receive [6]. The positive wage–effort relation has been replicated in dozens of gift exchange experiments using different parameters, protocols, and institutional set-ups. The wage–effort relation is particularly pronounced where there are repeated interactions between employers and workers, providing employers with the opportunity to leverage their reputation for paying above-minimum wages in order to attract workers who are willing to exert above-minimum effort.

These examples indicate that reciprocity is a powerful force in the relations between employers and workers, suggesting that fair treatment of workers has the potential to mitigate the problem of motivation. Fair treatment does of course not only comprise “fair” wages but also other components of compensation, features of the work environment, and social approval.

Reciprocity also implies the punishment of or use of sanctions against unfair or unkind actions. Evidence for negative reciprocity comes from numerous laboratory experiments, such as the ultimatum game, the public goods game with punishment options, or the gift exchange game with sabotage options. For example, in the ultimatum game, the first player is asked to allocate a given sum of money between his or herself and a second player. The second player can accept the offer (in which case both players get their respective fractions of the money according to the offer of the first player) or reject it (in which case both players receive nothing) (see Figure 1).



The main finding of the ultimatum game is that many second players reject offers when they perceive the offer as unfairly low. Thus, second players prefer to forgo money rather than accept what they perceive as an unfair outcome. Important applications of the ultimatum game are wage (re)negotiations: Employers should not expect workers to respond favorably toward **any** (increased) wage offer, but to do so only when the offer meets workers' perception of fairness.

Evidence for punishment due to negative reciprocity is also portrayed in a public goods experiment with a costly punishment option. In the public good game, the players face a social dilemma. They are endowed with a sum of money and can choose to keep it or to invest it in a public account. After all the players have decided on whether to contribute to the public good, the amount of money in the public account increases by a factor bigger than one. Then, the money in the public account is divided among all the players—among those who contributed and those who did not contribute. If all the players invested their total endowment in the public account, the outcome would be efficient, in the sense that the sum of earnings for all players would be maximized. However, each player has an incentive to free-ride on the contributions of the other players. The main finding of public good games with punishment is that players engage in costly sanctioning devices to punish free-riders (i.e. those who did not contribute to the public good) [7].

The previous examples illustrate how reciprocity shapes human behavior. While the discussion of the gift exchange game shows how positive reciprocity helps to mitigate the motivation problem, the findings of the ultimatum and public goods game portray how unfair actions (for example, low wage offers or free-riding) have the potential to trigger negative reciprocity, evoking counterproductive activities such as shirking or sabotage.

The potentially detrimental efficiency effect of unfair treatment is exemplified by a case involving faulty car tires sold in the US between 1991 and 2000 [8]. In 2000, Firestone had to recall and replace 14.4 million car tires because of manufacturing defects that caused the tread to separate from the tire on hot days. The US National Highway Traffic Safety Administration held Firestone responsible for 271 fatalities in accidents involving the tires. As a consequence, the stock market value of Bridgestone, the corporate parent of Firestone, fell from US\$16.7 to US\$7.5 billion. Bridgestone's top management was completely replaced, and the company considered dropping its use of the Firestone brand name.

This negative outcome can be understood as a consequence of unfair treatment of workers by Bridgestone's management. The factory in Decatur, Illinois, where the defective tires were manufactured, was the site of a long and bitter strike by United Rubber Workers that lasted from 1994 to 1996. Bridgestone wanted to reduce wages for new employees by 30%, increase shifts from 8 to 12 hours, and provide fewer paid holidays for older employees. Thus, Bridgestone had treated employees in a way that the employees perceived as "unfair." Workers negatively reciprocated by reducing their care in the production process. While, on an individual level, this does not imply that workers deliberately harmed Firestone's customers, in the aggregate it did result in the production of low-quality tires.

Other potentially detrimental efficiency consequences of unfair pay or treatment are adverse health effects. The physiological responses to perceptions of unfair pay have

been examined using an integrated approach exploiting complementarities between controlled lab and representative field data [9]. In a simple employer–worker experiment, workers produce revenue for firms by working on a tedious task. Employers decide how this revenue is allocated between themselves and their workers. In this environment, unfairness arises when a worker’s reward expectation is not met. Throughout the experiment, variations in the worker’s heart rate are recorded. Heart rate variability is an indication of elevated stress and has been shown to predict coronary heart diseases over the long term. The experiment provided evidence of a link between perceived unfairness and heart rate variability, with the higher the level of perceived unfairness (i.e. the lower the wage paid relative to total revenue created), the more stress the agents exhibited.

Based on these findings in the laboratory, we would expect to observe bad health outcomes within a population in which individuals perceive their wages to be unfairly low. German socio-economic panel data provides the opportunity for an investigation into this conjecture, as it contains a survey of a representative sample of the German population on, among many other things, the perceived fairness concerning their wages and their health outcomes. An analysis of the survey suggests a strong and highly significant negative association between health outcomes (in particular, cardiovascular health) and perceptions of unfair pay. Thus, unfair treatment in the workplace affects the efficiency of labor relations not only by reducing worker morale, but also by potentially affecting the health status of the workforce.

Trust and control: When explicit incentives backfire

The existence of social preferences suggests not only alternative determinants of worker motivation, but also important implications for the standard set of tools used to motivate and incentivize workers. However, the use of controlling devices in the workplace, as well as particular forms of explicit performance incentives, may entail “hidden costs” that firms should take into account when designing employment contracts and workplace environments.

To illustrate this idea, researchers implemented a very simple employer–worker relation with conflicting material interests [10]. In the experiment, the worker decides on how much revenue she accumulates. While the income of her employer increases the more revenue she accumulates, her own income decreases with this revenue. The interesting feature of the experiment is that before the worker chooses how much revenue she accumulates, the employer decides on whether or not to implement a minimum revenue level that the worker is required to accumulate. If the employer implements a minimum revenue requirement, the worker’s choice of how much revenue she wants to accumulate is restricted; she cannot accumulate less than the minimum revenue level. (Such a restriction could be interpreted as a minimum presence requirement, a requirement for a minimum level of output or quality, or a “complete” incentive contract.) If the employer does not implement a minimum revenue requirement, the worker is free to choose how much revenue she wants to accumulate. Not enforcing the restriction can be interpreted as providing the worker with discretion and responsibility about working times, output choice, or the deliberate choice of an incomplete contract, leaving considerable freedom of choice. In other words, restricting the choice set of workers implies controlling the agent, while not restricting workers is a signal of trust.

Hidden cost of control

An agent (worker) chooses a productive activity x , which is costly to the agent, with $c(x)=x$. The productive activity x is beneficial to the principal (employer) with $b(x)=2x$.

Endowments are 120 for the agent and 0 for the principal, and payoffs are determined as follows: The principal receives $2x$ and the agent receives $120-x$. Before the agent chooses x , the principal determines the agent's choice set.

In particular, the principal can either enforce a minimum level of x equal to $x \geq 10$ or leave the decision about x completely to the agent's discretion. Trusting the agent implies the risk of earning less than $2 \cdot 10$ points, simply because the agent is free to act according to material self-interest.

The experiment illustrates the idea that the use of controlling devices in the workplace, as well as particular forms of explicit performance incentives, may entail "hidden costs." Firms should take these into account when designing employment contracts and workplace environments (see Falk and Kosfeld, 2006).

Falk, A., and M. Kosfeld. "The hidden cost of control." *American Economic Review* 96:5 (2006): 1611-1630.

Under the standard assumption that economic agents are interested only in their own material outcome, one should not expect a worker to accumulate any revenue (since the worker's own material payoff decreases as revenue is accumulated). Employers who are expecting this should control their workers by requiring them to accumulate some minimum level of revenue. If individuals exhibit social preferences, however, workers may perceive the controlling decision by their employer as a sign of distrust. Why should the employer restrict the worker's choice set if he does not expect her to accumulate revenue below the minimum level? If control is in fact perceived in a negative way, it may induce workers to provide merely the minimum level, even though they might have chosen to accumulate more if they had been trusted. Thus, controlling workers may give rise to a self-fulfilling prophecy of distrust, whereby an assumption of selfish behavior from workers (expressed in the controlling decision) actually promotes selfishness.

The main finding of this simple experiment is that median levels of revenue are equal to the minimum level if workers are controlled and are much higher if the employer trusts them. This contrasts sharply with the standard economic model's prediction of no revenue accumulation when the employer trusts (i.e. does not control) the worker. Instead, trust pays in the experiment, because many workers are intrinsically trustworthy and react negatively to being controlled.

The intuition behind this argument is nicely captured by David Packard, the cofounder of computer company Hewlett-Packard, in his 1995 memoir [11]:

In the late 1930s, when I was working for General Electric...the company was making a big thing of plant security...GE was especially zealous about guarding its tool and parts bins to make sure employees didn't steal anything. Faced with this obvious display of distrust, many employees set out to prove it justified, walking off with tools and parts whenever they could...When HP got under way, the GE memories were still strong and I determined that our parts bins and

storerooms should always be open...Keeping storerooms and parts bins open was advantageous to HP in two important ways. From a practical standpoint, the easy access to parts and tools helped product designers and others who wanted to work out new ideas at home or on weekends. A second reason, less tangible but important, is that the open bins and storerooms were a symbol of trust, a trust that is central to the way HP does business.

The experiment suggests that elements in the labor contract that can be perceived as signals of distrust and control (such as minimum performance requirements) may harm more than they help. Similarly, characteristics of the workplace environment that limit freedom of choice and signal low expectations (such as high levels of monitoring and surveillance) may be equally counterproductive. The findings do not suggest, of course, that not controlling is universally optimal. The trade-offs between benefits and hidden costs of control will largely depend on:

- the personal characteristics of workers (in particular their trustworthiness);
- available control options;
- the risk preferences of employers; and
- the way the introduction of a controlling environment is communicated.

They do, however, suggest that, when designing incentive schemes and workplace environments, employers should be aware of the fact that, in choosing particular tools, they implicitly signal expectations about their workers' types, with the potential to reduce motivation.

Another well-known example of how explicit incentives can backfire is the field experiment involving Israeli day care centers [12]. The problem in many day care centers is that parents arrive late to pick up their child. A standard economic solution to this problem is to implement a fine for late parents. To assess the effectiveness of this solution, researchers divided several day care centers into a control group and a treatment group. The control group of day care centers did not impose fines on late parents. The treatment group did not initially impose fines, but then introduced them for several weeks before removing them again.

The standard economic model would predict that lateness should not increase in the treatment group (relative to the control group) after the introduction of fines. In contrast, the experiment found that lateness not only increased after the introduction of fines, but also settled at a level nearly twice as high as previous levels. Interestingly, the number of late-arriving parents did not fall back to its initial level when the fines were later removed.

A possible explanation for this surprising result is that before fines were introduced, late parents may have experienced a guilty conscience because they had violated the social norm that they should always pick up their child on time. The introduction of a pecuniary fine transformed a social relationship into an economic one. Given that fines were relatively low (and were invariant to the number of times the parent was late), being late became a more attractive option.

The two examples discussed above show that the psychology of incentives is much more complex than it appears in the traditional economics view. Studies have shown that

incentives may undermine motivation because they provide new information regarding the importance or the cost of the task, because they insult the agent, or because they are in conflict with notions of fairness and cooperation.

LIMITATIONS AND GAPS

The literature discussed above deals mainly with the fairness concerns of workers regarding the actions and intentions of their employers. Future research should address how the fairness concerns of workers regarding outcomes or the treatment of other workers affect their motivation. While the importance of social comparisons for the well-being of individuals is well documented, the implication for how relative concerns among workers may affect their performance at work is not well understood as yet. While social comparisons may induce workers with below-average pay to increase their work effort in order to keep up with their coworkers, pay inequality may also discourage below-average-earning workers. Therefore, an interesting follow-up question is how people select their reference group (i.e. why people feel inclined to compare their outcomes or treatment with some co-workers, but not with others).

It is equally important to improve our understanding of what constitutes the perception of fairness in general. While social comparisons are a promising concept for addressing the question of how individuals determine whether the wages they receive or the treatment they are exposed to is fair or not, more research is needed to figure out where our perceptions of fairness originate, how they develop over time, and why they differ across contexts and cultures.

SUMMARY AND POLICY ADVICE

This paper has discussed evidence on fairness and motivation. In contrast to the standard assumption in economics, according to which people are predominantly motivated by maximizing material payoffs, a large and growing body of evidence suggests that social preferences are also key drivers of human motivation. As a consequence, motivation should not exclusively rely on extrinsic incentives and money. Importantly, the psychology of incentives suggests that explicit performance incentives may actually backfire, implying that policy advice built only on simplistic economic assumptions may be severely counterproductive.

Acknowledgments

The author thanks an anonymous referee and the IZA World of Labor editors for their helpful comments on an earlier draft.

Competing interests

The IZA World of Labor project is committed to the *IZA Guiding Principles of Research Integrity*. The author declares to have observed these principles.

© Armin Falk

REFERENCES

Further reading

Falk, A., E. Fehr, and C. Zehnder. "Fairness perceptions and reservation wages: The behavioral effects of minimum wage laws." *Quarterly Journal of Economics* 121:4 (2006): 1347–1381.

Fehr, E., and A. Falk. "Psychological foundations of incentives." *European Economic Review* 46:4–5 (2002): 687–724.

Fliessbach, K., B. Weber, P. Trautner, T. Dohmen, U. Sunde, C. E. Elger, and A. Falk. "Social comparison affects reward-related brain activity in the human ventral striatum." *Science* 318:5854 (2007): 1305–1308.

Key references

- [1] Fehr, E., and A. Falk. 1999. "Wage rigidity in a competitive incomplete contract market." *Journal of Political Economy* 10 (1999): 106–134.
- [2] Williamson, O. *The Economic Institutions of Capitalism*. New York: Free Press, 1985.
- [3] Falk, A., and U. Fischbacher. "A theory of reciprocity." *Games and Economic Behavior* 54:2 (2006): 293–315.
- [4] Fehr, E., and S. Gächter. "Fairness and retaliation: The economics of reciprocity." *Journal of Economic Perspectives* 14:3 (2000): 159–181.
- [5] Fehr, E., G. Kirchsteiger, and A. Riedl. "Does fairness prevent market clearing? An experimental investigation." *The Quarterly Journal of Economics* 108:2 (1993): 437–459.
- [6] Akerlof, G. "Labor contracts as partial gift exchange." *Quarterly Journal of Economics* 97 (1982): 543–569.
- [7] Fehr, E., and S. Gächter. "Cooperation and punishment in public goods experiments." *American Economic Review* 90:4 (2000): 341–369.
- [8] Krueger, A., and A. Mas. "Strikes, scabs, and tread separations: Labor strife and the production of defective Bridgestone/Firestone tires." *Journal of Political Economy* 112:2 (2004): 253–289.
- [9] Falk, A., F. Kosse, I. Menrath, P. E. Verde, and J. Siegrist. *Unfair Pay and Health*. Working Paper, 2013. Online at: <http://www.cens.uni-bonn.de/team/board/armin-falk/publications/unfair-pay-and-health>
- [10] Falk, A., and M. Kosfeld. "The hidden cost of control." *American Economic Review* 96:5 (2006): 1611–1630.
- [11] Packard, D. *The HP Way: How Bill Hewlett and I Built our Company*. New York: HarperBusiness, 1995.
- [12] Gneezy, U., and A. Rustichini. "A fine is a price." *Journal of Legal Studies* 29:1 (2000): 1–17.

The full reference list for this article is available from the IZA World of Labor website (<http://wol.iza.org/articles/fairness-and-motivation>).