Center for European, Governance and Economic Development Research
cege

Discussion Papers

Number 304 – February 2017

GENDER DIFFERENCES IN MOTIVATIONAL CROWDING OUT OF WORK PERFORMANCE

Volker Benndorf
Holger A. Rau
Christian Sölch

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

ISSN: 1439-2305

Gender Differences in Motivational Crowding Out of Work Performance*

Volker Benndorf[†1], Holger A. Rau[‡2] and Christian Sölch[§3]

[1]University of Frankfurt
[2]University of Mannheim, University of Göttingen
[3]University of Erlangen-Nuremberg

February 2017

Abstract

This paper studies motivational crowding-out effects after financial incentives are lowered. In a real-effort setting, workers receive a piece rate before financial incentives are substituted by a one-time payment. Under the fixed payment, effort is significantly lower only when preceded by piece-rate incentives. The decrease is driven by a fraction of men who reduce their effort by 12%, whereas women constantly perform well. We find that this motivational crowding-out effect disappears when men do not have prior experience of a piece rate. In a series of control treatments, we discard all alternative explanations besides from motivational crowding out.

JEL Classification: C91, J16, M54.

Keywords: Gender Differences, Incentives, Motivational Crowding Out, Real-Effort Task.

*We are grateful to Gary Charness, Catherine Eckel, Uri Gneezy, Veronika Grimm, Nagore Iriberri, Peter Katuscak, Julie Neilson, Muriel Niederle, Nikos Nikiforakis, Hans-Theo Normann, Jörg Oechssler, and Emmanuel Peterlé for helpful comments. We would especially like to thank Simon Gächter for his helpful support. We are also grateful to the participants of the European ESA Conference 2014 in Prague, the EFI Workshop 2014 in Nuremberg, the Workshop on Gender Issues in Experimental Economics 2014 in Alicante, the NIBS Workshop 2015 in Nottingham, the BEGender Workshop in Berlin 2015, and seminar audiences at the universities of Düsseldorf and of Göttingen. We thank the Emerging Fields Initative (EFI) at the University of Nuremberg and the Düsseldorf Institute for Competition Economics (DICE) for financial support.

[†]Grüneburgplatz 1, 60323 Frankfurt (Germany) E-Mail: *benndorf@dice.hhu.de*

[‡]Corresponding Author, Platz der Göttinger Sieben 3, 37073 Göttingen, E-Mail: *holger.rau@uni-goettingen.de*

[§]Lange Gasse 20, 90403 Nürnberg, E-Mail: *christian.soelch@fau.de*

1 Introduction

In economic crises or in times of financial distress, firms often pursue extensive restructuring plans to get back on track. Many plans target a substantial reduction of the labor costs. This is typically achieved by mass layoffs, but in some cases firms also take alternative routes such as cutting wages[1] or changing the company's payment schemes.[2] One could argue that changing the payment scheme is the least drastic measure among those listed above. However, when companies distance themselves from performance-based payments (e.g., bonuses) and increase the relative importance of the fixed wage, this may lead to unwanted side effects.

The psychological literature emphasizes that a "crowding out of intrinsic motivation" may occur when performance-based compensation schemes are reduced or discontinued (e.g., Titmuss, 1970; Deci, 1971).[3] When financial incentives are in the place, the intuition is that they leave a mark (Gneezy and Rustichini 2000a,b[4]), which affects the work motivation of employees. If, for example, the incentives are lowered or removed, workers may compare the new environment to the prior financial incentives and reduce efforts rather drastically. In other words, workers who work in the absence of performance-based remuneration may be less motivated if they had previously faced financial incentives compared to a situation without performance-based incentives where they have no prior experience of such a payment scheme. Hence, there may be durable motivational crowding-out effects when performance-based remuneration is changed to flat-rate pay (see Gneezy et al., 2011).[5]

So far, empirical research has repeatedly documented the importance of remuneration changes on incentive effects. For instance, field experiments find that productivity increases when hourly wages are altered to piece rates (e.g., Fernie and Metcalf, 1999; Lazear, 2000; Shearer, 2004), whereas the opposite is true when piece rates are removed (Freeman and Kleiner, 2005). Although field data demonstrates that remuneration changes may influence productivity, it is difficult to isolate their effects on workers' intrinsic motivation. In the field, workers are typically affected by a number of side effects which may complicate the measurement of motivational crowding out. One example is reputation effects. Employees in firms are exposed to the pressure of being fired and cannot readily reduce effort when incentives are lowered. Another issue is that firms who change the incentive schemes are often acting in a difficult economic

[1] Evidence is found in the finance sector (e.g., Wall Street Journal: `http://www.wsj.com/articles/SB10236802077044731 60`) or in the mining industry (e.g., Bloomberg Business: `http://bloom.bg/1L9SXye`).

[2] For example, in the wake of the 2008 financial crisis banks reduced their employees' bonus payments (for example: Financial Times: `http://on.ft.com/QGRu8v` or New York Times: `http://nyti.ms/18oiOSt`).

[3] The theoretical literature on rewards and punishment highlights that financial incentives may lead to motivational crowding-out effects (Frey 1994; Bénabou and Tirole 2003; Bénabou and Tirole 2006). Moreover, Frey and Oberholzer-Gee (1997) find that paying subjects to accept the construction of a nuclear power plant in their vicinity decreases their willingness to do so. Mellström and Johannesson (2008) report that paying subjects for blood donations may decrease their propensity to donate blood.

[4] Gneezy and Rustichini (2000b) find in an experiment that subjects compensated by a small piece rate achieve a lower performance in IQ tasks as compared to non-incentivized subjects.

[5] This is also demonstrated by experiments on institutional changes. For example, Falkinger et al. (2000) report that cooperation decreases in a public-good game once a subsidy mechanism is removed. The same holds in a gift-exchange game (Gächter et al., 2011) when contracts with bonuses are taken away. Both studies find that absent incentives, subjects contribute more when they had never experienced such incentives in the first place.

environment. Freeman and Kleiner (2005) analyze data of an American shoe manufacturer that changed the payment scheme from piece rates to hourly wages in the 1990s. The authors report that this change resulted in lower productivity. However, they also acknowledge that at this point in time, the domestic shoe production in the US had been in decline for decades. Since similar confounds may also exist in other field data, we run a controlled lab experiment to increase the understanding of motivational crowding-out effects in this context.

In this paper we present the data of a real-effort experiment where the same subjects subsequently work under different compensation schemes. Our main research question focuses on the impact of prior financial incentives on subjects' future work motivation when incentives are lowered. In other words, we study the influence of a prior piece rate on subjects' intrinsic work motivation under a subsequent flat wage. We therefore systematically vary the order of how monetary incentives are introduced and removed. In all of our treatments subjects work in the absence of employers, as we intend to rule out other forces such as reciprocal behavior,[6] reputation effects or strategic motives.[7] In our main treatment subjects first work under a piece rate. Afterwards the remuneration is changed to a flat wage. Our hypothesis is that workers show a significantly lower performance under a flat wage when there were previous financial incentives. The data find meaningful support for the occurrence of this motivational crowding-out effect. With the means of several control treatments, we discard all alternative explanations besides motivational crowding out.

The experiment unveils a substantial gender effect in crowding out of intrinsic motivation. We find that only male employees show the motivational crowding-out effects that we hypothesize. In the main treatment, a sizable share of men drastically reduce their effort by 12%. In stark contrast, women virtually show the same performance under the piece rate and the subsequent fixed payment. Altogether, women always perform well, irrespective of remuneration and gender differences vanish when workers do not experience prior incentives.

Our findings have implications for many aspects of labor economics. First, we provide controlled support that prior piece rates may have a negative impact on subjects' future performance. Second, we demonstrate that this is exclusively driven by male workers. The results suggest that when changing remuneration schemes it is not only the sequence of compensation that matters, but also the gender composition of the workplace. These findings may be of importance for managers when they have to restructure their companies. The result that incentive effects never harm women's performance has interesting implications for job-promotion decisions. When women perform well even under low incentives, it may explain why employers try to exploit this by not promoting them. This may potentially add new insights to the understanding of labor-market phenomena such as glass-ceiling effects or the gender wage gap.

[6] Irlenbusch and Sliwka (2005) show in an experimental principal-agent setting that reciprocal relationships may also induce motivational crowding out. This appears in a setting when principals can choose between a fixed payment which is supplemented by a piece rate and a fixed payment. By contrast, effort is higher when principals can only pay fixed payments.

[7] Schnedler and Vanberg (2014) theoretically point out that strategic motives may be an alternative explanation as to why workers decrease their effort when incentives are increased. That is, in the presence of financial incentives workers may lower effort to enforce a higher remuneration.

2 Experimental Design

In this section we describe the general framework, the treatments, and the procedures of our experiment. We then turn to our hypothesis section.

2.1 General framework

The current paper studies whether initial financial incentives lead to a crowding out of intrinsic motivation. We are particularly interested in whether performance under (non-contingent) fixed-payment settings is lower when subjects experienced (contingent) piece rates beforehand. To address this question we present a real-effort experiment where we study the performance of the same subjects under different compensation schemes.

All treatments consist of two parts with 10 periods each. Between the two parts we systematically vary the sequence of contingent and non-contingent incentives. The set-up applies extrinsic incentives by means of a piece rate. Whereas we employ a non-contingent fixed payment to study subjects' effort when compensation does not depend on work performance. Here, subjects receive a one-time lump sum payment. To isolate the strategic effects from motivational crowding out while minimizing side effects such as reciprocal or strategic behavior, we refrain from modeling employers.

In the experiment, we measure performance by means of a real-effort task which extends the word encryption task by Erkal et al. (2011).[8] Figure 1 presents our variant, where subjects are asked to encode random combinations of three letters into numbers (see the instructions in the appendix for a screenshot of our task). More precisely, each letter in the first row "word" has to be encrypted in a three-digit number. The "allocation table" of the task provides the subjects with the correct allocation of the letters and the corresponding three-digit number. The table always displays all 26 capital letters of the Latin alphabet. Subjects have to type in the correct three-digit number of each letter in the "code" row below the letter. After all three letters have been encoded the subjects press a button and are informed as to whether the answer was correct. They are also given their total number of correctly solved puzzles. While doing the task subjects receive no information on the total payoff they have earned so far.[9]

Since the task is applied in a within-subjects design, it aims to rule out the learning behavior of subjects.[10] This is why we apply the following procedure in our version of the encryption task. Whenever a subject enters a correct solution, the word to be encrypted changes. At the same time, the mapping from letters to numbers and the positions of the letters in the table are randomly rearranged. This double randomization is a special feature of our task and

[8] Word encryption tasks have been successfully applied in real-effort experiments on gender differences (e.g., Iriberri and Rey-Biel, 2014; Kuhn and Villeval, 2015; Masclet et al., 2015).

[9] We did not implement an outside option as our experiments do not focus on the impact of incentives on compliance behavior. Corgnet et al. (2015) show that under individual pay performance does not significantly differ in the presence of an outside option. By contrast, in team-production incentive schemes performance is lowered when an outside option exists.

[10] The piece-rate data of this experiment is part of another paper that introduces our new task as a workhorse to minimize learning behavior in repeated real-effort experiments (Benndorf et al., 2014).

word:	**Z**	**N**	**T**

code:	113	154	

allocation table:

B	T	R	S	U	Z	F	N	C	Y	V	X	H	Y	K
384	118	201	543	386	113	980	154	745	265	432	262	110	960	245

Figure 1: Example of a problem in the real-effort task. For reasons of space, only 15 allocations are presented in this example.

it is the main difference to the task introduced by Erkal et al. (2011). The idea is that the randomization processes prevent subjects from increasing their performance over time because of improving their ability in doing the task.[11] When subjects enter the wrong answer they are informed by the computer program. Then, the number allocations and the locations of the letters are not shuffled until subjects make a correct input.

2.2 Treatments

Our main treatment is called *Piece-to-Fix* and consists of two parts with 10 periods each. It aims to study whether prior financial incentives may crowd out subjects' future work motivation. The treatment analyzes the role of prior financial incentives in the following way. In part one subjects repeatedly work under a piece rate and receive €0.08 for each correctly-solved task. Subjects know of the existence of a second part, but have no information on its design. Before part two starts we change the piece-rate remuneration to a non-contingent payment which is paid before part two starts. That is, subjects are informed that the piece rate has been removed and that they will receive a one-time fixed payment of €8 for the next 10 periods. It is independent of their performance in the task and was calibrated such that earnings under the piece rate and the fixed payment are approximately the same.[12] The idea is to avoid wealth effects, i.e., subjects might become frustrated when realizing that they are earning much less under the fixed payment than under a piece rate. The treatment tests whether performance decreases when incentives are removed. A special interest is whether prior financial incentives lead to lower effort compared to the case when subjects did not experience an initial piece rate beforehand. To address the latter issue, we make use of a first control treatment.

In the control treatment, we again focus on the same real-effort setting with two parts of the same length. The treatment attempts to isolate whether an effort decrease between part one and two of *Piece-to-Fix* is reinforced by a motivational crowding-out effect. In this regard, we introduce a setting where subjects do not experience piece-rate incentives before they start

[11] Erkal et al. (2011) report that subjects in their task increased their performance by 72% over time.

[12] The €8 fixed payment was aligned to prior evidence of a pilot session where subjects in the task solved, on average, 10.19 puzzles.

working under the fixed payment in part two. As a consequence, subjects both in the first and the second part work under a fixed payment of €8. When working in the first part, subjects do not know the conditions of the second part. At the beginning of the second part they are informed that they will again receive a one-time fixed payment. We call this treatment: *Fix-to-Fix*.

Two further control treatments were run to rule out alternative drivers besides motivational crowding out. The treatments also consist of two parts with 10 periods each. This section only briefly introduces these treatments and their design. We will comment on the idea of these treatments in more detail when we present our main data in the results section. The first additional treatment is named: *Piece-to-Fix10*. This treatment is very similar to our main treatment (*Piece-to-Fix*). Here, subjects initially work under the same piece rate of €0.08. Then, prior to the start of part two they are informed that the remuneration has been changed to a non-contingent payment. A crucial difference is that the fixed payment is higher than in *Piece-to-Fix*, i.e., €10 instead of €8. The idea of this treatment is to address whether the size of the crowding-out effect depends on reference points, i.e., subjects may be concerned about income differences between the piece-rate and the fixed-payment part. Our last control condition is named *Piece-to-Piece*. In this scenario, subjects work under a €0.08 piece rate in both parts of the experiment. The goal is to rule out that subjects' performance in the second part does not decline due to fatigue at the end of the experiment. Table 1 presents an overview of our treatments. It displays the order of the remuneration schemes and the number of underlying observations.

	Remun. 1	Remun. 2	Both Gender	Males	Females
Main Treatment	periods 1-10	periods 11-20	obs.	obs.	obs.
Piece-to-Fix	€0.08 piece	€8 fixed	124	60	64
Controls					
Fix-to-Fix	€8 fixed	€8 fixed	64	32	32
Piece-to-Fix10	€0.08 piece	€10 fixed	63	32	31
Piece-to-Piece	€0.08 piece	€0.08 piece	32	16	16
	–	–	283	140	144

Table 1: Overview of the remuneration schemes (periods 1–10 and periods 11–20) and number of observations in our treatments.

2.3 Procedures

At the beginning of the experiment, subjects in all treatments first received a set of written instructions which explained the usage of the real-effort task. After all subjects had confirmed that they understood the functioning of the task, we started the trial period.[13] Subjects were asked to solve exactly 10 puzzles without payment. After all subjects had successfully completed

[13]The purpose of the trial period is to make subjects familiar with the task and to mitigate learning behavior.

the trial stage they were provided with a new set of instructions for the first part.[14] In the instructions it was explained that the experiment would consist of two parts and that the current instructions only covered the first part. Participants were informed that at the end of the first part they would receive a new set of instructions explaining part two. In all treatments, subjects perform the encryption task for 20 periods. One period lasts two minutes. At the end of a period the real-effort task automatically stops and inputs are no longer possible. Once a period has ended, subjects have a break of 10 seconds where they receive information on the total number of correct answers in the previous round. Then the next round starts. The experiments are divided into two parts where both parts comprise 10 periods. The treatments only differ in respect to whether one of the remuneration schemes was applied in the first, in the second or in both parts. After the end of period 20 we conducted a questionnaire on subjects' motivations for doing the task under the corresponding remuneration schemes. Finally, we asked questions on demographics. We use this information to get deeper insights into subjects' work motivation and as control variables in our regressions.

The experiments were programmed with z-Tree (Fischbacher, 2007) and conducted in the *LERN* laboratory of the University of Erlangen-Nuremberg. In all treatments we have sessions of 32 subjects.[15] In total, 283 subjects (143 women and 140 men) participated in our experiments (see Table 1). Subjects are from various fields and were invited with ORSEE (Greiner, 2015). On average, the sessions lasted 60 minutes. Subjects did not receive a show-up fee and the average earnings were €16.37.[16]

2.4 Hypothesis

Our main treatment (*Piece-to-Fix*) analyzes the impact of initial financial incentives on subjects' performance in the absence of contingent incentives. In this case, we expect that subjects will decrease their performance after the remuneration is changed from piece rate to a flat payment.

Hypothesis
Subjects in Piece-to-Fix will decrease their effort levels after the remuneration scheme is changed from a piece rate to a fixed payment.

There are several reasons why subjects should lower effort after the compensation is changed to a fixed payment. First, experiments on real-effort tasks report a positive relation between effort and incentives (e.g., Dickinson, 1999; Dohmen and Falk, 2011; Ederer and Manso, 2013).[17]

[14] When we started the trial period subjects were merely informed that the experiment would follow. They were also told that the subsequent experiment would consist of two parts.

[15] The only exception is one session of *Piece-to-Fix* where only 28 subjects showed up. Furthermore, we had to drop one female observation of *Piece-to-Fix10* as the participant's computer crashed.

[16] The average payments of the treatments were: €15.92 (*Piece-to-Fix*); €16 (*Fix-to-Fix*); €17.70 (*Piece-to-Fix10*); €16.20 (*Piece-to-Piece*).

[17] Although, subjects generally react to higher incentives, Pokorny (2008) finds that the performance under low piece-rate incentives may be higher than under high piece rates. The paper argues that subjects refer to reference

In the piece-rate part of *Piece-to-Fix* subjects receive a higher income when performing well. By contrast, exerting more effort does not increase earnings under a fixed payment. Thus, subjects should reduce effort in the fixed-payment part of *Piece-to-Fix*.

Second, subjects may be prone to motivational crowding-out effects. This should additionally decrease effort under the flat payment. We designed *Piece-to-Fix* in such a way as to entail motivational crowding-out effects under the fixed payment. The idea is that subjects experience financial incentives (i.e., a piece rate) before the remuneration is changed to a flat payment. The literature on motivational crowding out (Gneezy et al., 2011) and real-effort experiments (Sliwka and Werner, forthcoming) reports that prior financial incentives may work as remuneration benchmarks. It follows that subjects compare an initial piece rate to subsequent remuneration schemes. The consequence is that subjects may become demotivated and might lower effort once monetary rewards are removed. Evidence is documented by within-subjects experiments on institutional change in public-good and gift-exchange games (Falkinger et al., 2000; Gächter et al., 2011). The papers report that cooperation is lower when financial incentives have been removed compared to the case where they have never been applied. In our case, prior piece-rate incentives should attenuate subjects' intrinsic motivation when working in the absence of incentives. In this paper we establish the motivational crowding-out effects in the main treatment (*Piece-to-Fix*) where subjects in the first part do not face financial incentives before working under the flat payment in part two.

3 Results

In this section, we test our hypothesis and present the experimental results. Note that we always report two-sided p-values when applying non-parametric tests.

3.1 Main results

The left panel of Figure 2 depicts the development of subjects' average effort over time in our main treatment *Piece-to-Fix*.[18] The diagram depicts subjects' performance in the first (periods 1–10) and second part (periods 11–20) of the experiment. We report the data conditioned on males and females as well as pooled for both genders.[19]

The left panel demonstrates our main result: subjects' performance clearly shifts downwards

incomes normally earned in experiments, which can be early achieved under high incomes. As a consequence, subjects may reduce their effort after achieving their anticipated income. Note this should not matter in our case, as we do not vary the level of the piece rate. Instead, we substitute the piece-rate by a fixed payment which was calibrated to be similar in terms of the prior piece-rate earnings.

[18] To improve the readability of the diagram, we do not include standard deviations. See Table 4 in the appendix for more detailed information.

[19] In our main treatment, women achieve a significantly higher performance in both parts of *Piece-to-Fix* (two-sided Mann-Whitney test, $p < 0.001$). This is in line with the results of Majeres (1983) who reports a female advantage in symbol-digit encryption tasks. Our post-experimental questionnaire reveals that half of the men (50%) and women (54%) believe that no gender differences in performance exist. A similar fraction of men (25%) and women (23%) expect that either men or women perform better.

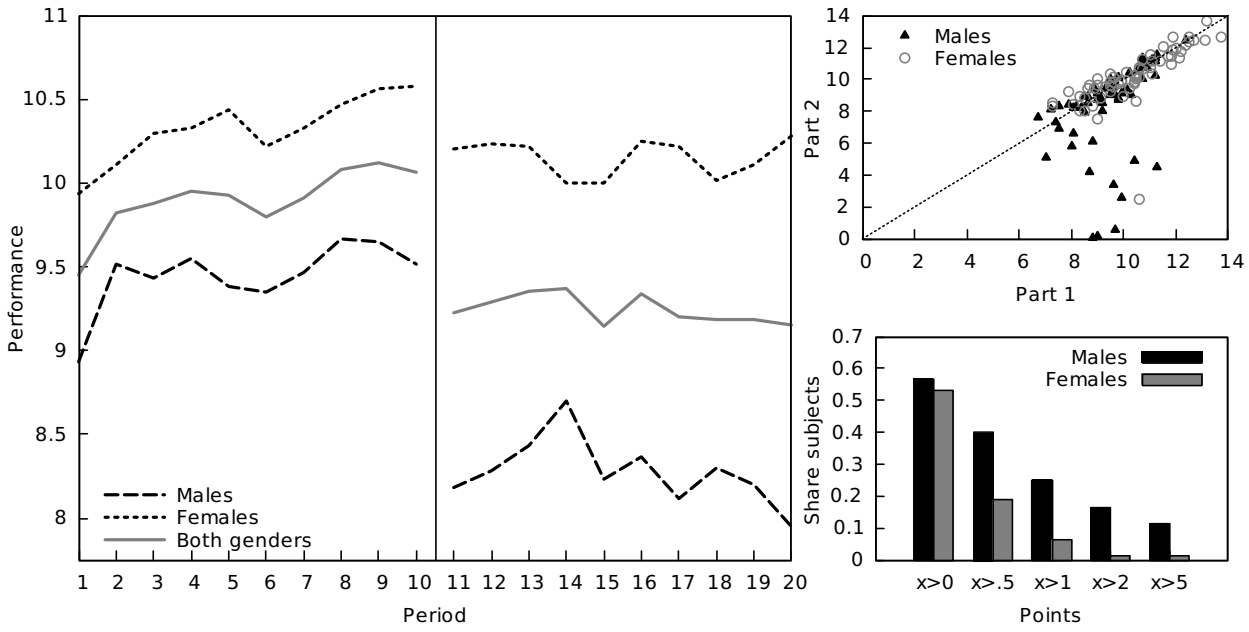


Figure 2: The left panel shows the development of the average performance over time. The upper-right panel compares subjects' average performance in the first and second part of the main treatment at the individual level. The lower-right panel plots the share of male and female subjects who reduce their effort by at least x points in the second part.

after the compensation scheme is changed from a piece rate to a fixed payment (see solid line). In the first part, when subjects are paid according to a piece rate, the average performance is 9.90. However, effort significantly decreases to 9.25 after the piece-rate incentives have been substituted by a fixed payment (two-sided Wilcoxon matched-pairs test, $p = 0.004$). In periods 11 to 20 workers' average fixed-payment performance is lower than each single period of the piece-rate part. We therefore find strong support for our hypothesis. The figure also shows that this decline is exclusively driven by men (see dashed line). Men significantly reduce their performance by 12% from 9.45 to 8.28 (two-sided Wilcoxon matched-pairs test, $p = 0.002$), whereas female subjects (see dotted line) do not respond to the remuneration change and show practically the same performance under the piece rate (10.33) and the fixed payment (10.15) (two-sided Wilcoxon matched-pairs test, $p = 0.420$). When comparing the changes in the performance of men and women, we find that men decrease their efforts more drastically than women (the average change for men is -1.17 points compared to -0.17 points for women). This difference across genders is statistically significant at the 5% level (two-sided Mann-Whitney test, $p = 0.047$).

Main Result

(a) Subjects significantly decrease their effort when the piece rate is changed to a fixed payment. (b) This result is entirely driven by men who reduce their effort substantially while women's performance remains practically unaffected.

A closer look at the data reveals that the drop in men's performance is driven by a minority

who drastically reduce their effort. This is illustrated by a scatter plot in the upper-right panel of Figure 2. Here, we compare the performance of individual subjects in the first and second part of the experiment. In the diagram, each black triangle represents one male subject and each gray circle represents one female subject. It can easily be seen that a sizable share of the male subjects reduce their efforts quite substantially in the second part (see black triangles below the 45-degree line) while the performance of female subjects is stable (see gray circles along the 45-degree line). In fact, there was only one female subject who reduced her effort by a considerable degree. The bar chart in the lower-right panel shows the shares of male and female subjects who reduced their performance by at least x points. More precisely, x is the average level of performance points and it is depicted on the horizontal axis. In this graph, there is further evidence that men reduce their effort more drastically than women. For example, the third black bar indicates that 25% of the male participants reduce their effort by at least one point compared to only about 6% of the female participants (see third gray bar).

3.2 Explaining the main result

The section above emphasized that men's performance declines in the second part of our main treatment. We argue that the removal of initial financial incentives triggers a motivational crowding out for male subjects. However, there may also be alternative explanations for the drop in performance we observe. This subsection therefore analyzes a variety of potential drivers that may explain our main result. Importantly, our analysis shows that all potential explanations can be discarded, except for the one suggesting motivational crowding-out effects for men.

Lack of incentives

A possible explanation could be the lack of financial incentives in the second part of *Piece-to-Fix*. To test this, we consider a control treatment (*Fix-to-Fix*) where incentives are not lowered from piece rate to the fixed payment. Instead, subjects receive the fixed payment in both parts of this treatment. Hence, both fixed-payment parts of the control treatment are characterized as working conditions where prior financial incentives are absent. Figure 3 compares the fixed-payment performance of the main treatment to both parts of the control treatment. The diagram presents subjects' average performance and conditions on both genders. We find that in *Fix-to-Fix* men and women always exert a high performance, though no performance-based remuneration exists.[20]

If the lack of incentives was a crucial driver, men's performance in either part of *Fix-to-Fix* should be similarly low as in the second part of the main treatment. Yet, the second and the third bars (left panel: "Males") demonstrate that men exert a significantly higher effort in both parts of the control treatment compared to the second part of the main treatment (first

[20]This supports the lab findings of Azmat and Iriberri (2016), i.e., subjects generally exert positive effort under a flat payment.

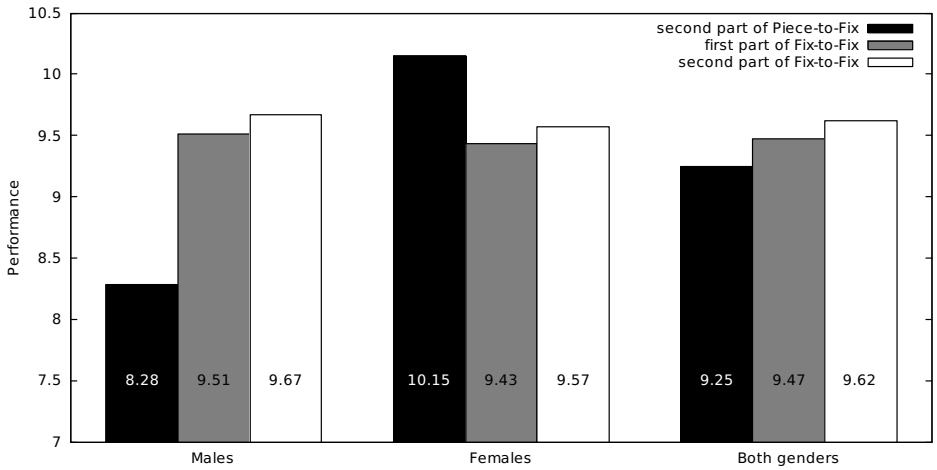


Figure 3: Subjects' fixed-payment performance in our main treatment and in the control treatment (*Fix-to-Fix*).

black bar) (two-sided Mann-Whitney tests, both parts: $p < 0.03$). Moreover, in *Fix-to-Fix*, the part one and part two performance of men does not significantly differ from the piece-rate performance in the main treatment (two-sided Mann-Whitney tests, both parts: $p > 0.234$). Hence, the lack of monetary incentives per se cannot explain the drop in men's performance when incentives are removed in the main treatment.

In the first part of *Fix-to-Fix* one might argue that there is an alternative explanation for the high performance of men: they do not know the remuneration conditions of part two. Intuitively, men might expect the performance in the first part to affect their prospective earnings in the second part. Even though our instructions did not contain any suggestion of such a scenario, we cannot completely rule out that some subjects had said expectations. However, our previous analyses demonstrate that this reasoning cannot explain the behavior in our main treatment. Note that if it was a crucial driver for our main result, men should immediately decrease effort in the second part of *Fix-to-Fix* when informed of the subsequent remuneration conditions.

Fatigue

Another possible driver of our main finding could be fatigue. When subjects in our main treatment start to work under a piece rate they do not know the payment conditions of part two. Thus, it is possible that subjects who drastically reduce their efforts may have spent too much energy on the previous incentivized task. As a consequence, they might be tired and cognitively depleted afterwards. The literature on competitiveness reports evidence of low-performing men being prone to such behavior (Cotton et al., 2013).

However, we demonstrate with another control treatment that this explanation can also be ruled out in our data. In our so-called *Piece-to-Piece* treatment, subjects receive the €0.08 piece rate in both parts of the experiment. If the reason for the drop in the performance was fatigue, we should see a comparable drop in performance of some subjects in this treatment.

This is, however, not the case. In *Piece-to-Piece* we even find the opposite: subjects of both genders manage to increase their performance (males from 9.83 to 10.06 and females from 10.04 to 10.58). The difference is significant for females and weakly significant for male subjects (two-sided Wilcoxon signed-rank test, $p = 0.001$ and $p = 0.060$, respectively). This clearly supports the notion that the drop in performance in the main treatment is caused by subjects who do not want to exert higher efforts and not by exhausted subjects.

Reference points

The possible impact of reference points might be another rationale for our main result. The intuition is as follows: in the main treatment, subjects receive an €0.08 piece rate in the first part and an €8 fixed payment in the second part. Thus, subjects who average below (above) 10 correctly-solved words per period earn less (more) than they would under the fixed payment they receive in the second part. This may affect individual behavior since subjects with significant income variations might perceive the fixed payment as a reference point. The consequence might be that these subjects react differently to the fixed payment.

The scatter plot[21] in Figure 2 shows that a significantly higher fraction of men (68.33%) achieve a part-one income below €8 (average performance <10) as compared to women (42.18%) ($\chi^2(1) = 8.548,\ p = 0.003$). The diagram suggests that men with a relatively low first-part performance are most likely to reduce their efforts. It could be argued that they interpret the fixed payment as a signal suggesting that their performance in the first part was poor and that they should try to solve more puzzles in the second part.[22] If the threshold of 10 words per period is perceived as unrealistically high, this might trigger frustration and induce some subjects to reduce their efforts.

The notion that our main result was caused by reference points is nevertheless wrong. We rule this reasoning out with a final treatment (*Piece-to-Fix10*) where we increase the probability of payoff differences between parts one and two. The treatment is identical to the main treatment except for the level of the fixed payment in the second part. That is, subjects receive a fixed payment of €10 instead of €8. It aims to increase the fraction of subjects who generate a part-one income below the fixed payment. If there was a causal relation between the performance in the second part and the income of part one, we should observe a pronounced decrease of subjects' average performance under the fixed payment. In this treatment we find that the fraction of men who average below the level of the subsequent fixed payment is significantly higher in *Piece-to-Fix10* (93.75%) than in *Piece-to-Fix* (68.33%) ($\chi^2(1) = 7.653,\ p = 0.006$). Nevertheless, we observe the same pattern as in the main treatment. In *Piece-to-Fix10*, men significantly decrease their performance from 9.16 to 8.19 when the fixed payment was introduced (two-sided Wilcoxon matched-pairs test, $p = 0.018$), whereas there is no significant change in

[21] Figure 3 in the appendix presents an overview of the scatter plots of all treatments.

[22] Bowles and Polania-Reyes (2012) argue that incentives may reveal information about the targeted behavior on agents' performance. This may affect the performance of subjects in work tasks.

women's performance (from 9.62 to 9.46) (two-sided Wilcoxon matched-pairs test, $p = 0.638$).[23] Comparing this control treatment to the main treatment, between parts one and two there is no difference in the performance change of men (*Piece-to-Fix10* : -14,5%; *Piece-to-Fix*: -12.3%) (two-sided Mann-Whitney test, $p = 0.765$). The pattern for women is also virtually unaffected by the increased fixed payment. In contrast to men, they hardly show a change in both treatments (*Piece-to-Fix10*: -1.6%; *Piece-to-Fix*: -1.3%; two-sided Mann-Whitney test, $p = 0.968$).

Post-experimental questionnaire

To gather further evidence on subjects' work motivations, we conducted a post-experimental questionnaire on participants' motives. The questionnaire was conducted in the last sessions (i.e., two *Piece-to-Fix10* sessions and one *Fix-to-Fix* session). In what follows, we present the answers to questions on the role of intrinsic motivation. We asked the following question for each of the two parts: *Did you try your best when doing the task in the first/second part?* In the piece-rate part when subjects state that they tried their best it may be that this was spurred by monetary incentives but also by intrinsic motivation. However, in the fixed-payment part, if subjects state that they tried their best this only signals that intrinsic motivation played a role as piece-rate incentives are absent. We further asked whether subjects concentrated on a performance goal. If they confirm that they had such a goal this would additionally emphasize the role of intrinsic motivation. Again, the reasoning would be that subjects did not receive a piece rate. We asked the following two questions for each of the two parts: *Did you have a performance goal when doing the task in the first/second part?*

	Yes, I tried my best		Yes, I had a goal	
	men	women	men	women
Fix-to-Fix pt. 1	0.94	1.00	0.69	0.88
Piece-to-Fix10 pt. 1	0.97	1.00	0.56	0.71
Fix-to-Fix pt. 2	0.89	0.89	0.88	0.56
Piece-to-Fix10 pt. 2	0.56	0.69	0.22	0.42

Table 2: Subjects' answers in the post-experimental questionnaire.

Table 2 presents the results of the questions. In the first parts of *Fix-to-Fix* and *Piece-to-Fix10* nearly all subjects stated that they had tried to solve as many puzzles as possible. Turning to the second part we find a similar pattern in *Fix-to-Fix* where subjects did not experience a piece rate in the beginning. This stands in contrast to the second part of *Piece-to-Fix10*. That is, fewer men (56%) than women (69%) stated that they tried their best under the fixed payment after experiencing a piece rate in part one. Hence, the replies are in line with the idea that initial financial incentives may have crowded out the intrinsic motivation of men in *Piece-to-Fix10*. In that case, the number of men who stated that they tried their best is lower compared to all other cases. We find additional support when focusing on the questions targeting performance goals. In the second part of *Piece-to-Fix10* almost twice as many women

[23] See Table 5 in the appendix for the periodical performance of the subjects.

(42%) as men (22%) stated that they had a performance goal. By contrast, in the second part of *Fix-to-Fix* the fraction of men with a performance goal increases to 88%. Focusing on parts one, we find that moderately more women claim that they had performance goals.

To summarize, the post-experimental questionnaire promotes the idea that our main result is induced by a crowding-out effect of intrinsic motivation. First, we find support for the idea that subjects were intrinsically motivated, as subjects in the first and second parts of *Fix-to-Fix* stated that they had tried their best. Second, we find evidence of a motivational crowding-out effect for male subjects only. That is, we only find for men that the fraction of subjects who state that they had tried their best decreases under *Piece-to-Fix10*.

3.3 Regression analysis

In our OLS regression analysis, we include the data of all four treatments.[24] The regressions apply clustered standard errors for 283 subjects. In the analysis we concentrate on three sources which may affect subjects: (i) the payment scheme; (ii) subjects' gender; (iii) the part of the experiment. In Regression (1) we incorporate the following dummy variables to test for these sources: *Piece* controls for the remuneration scheme (1: piece rate; 0: fixed payment), *Female* focuses on gender (1: female; 0: male), and *Part 2* controls for the part of the experiment (1: part 2; 0: part 1). In regressions (3) and (4) we consider the corresponding interaction terms of these dummies. Regressions (2) and (4) add control variables.

In regressions (1) and (2), we consider only the dummy variables (*Piece*, *Female*, and *Part 2*), and disregard possible interactions between these factors. These two regressions seem to suggest that women generally perform better in the word encryption task since the *Female* dummy is highly significant with a positive coefficient. Analogously, one might conclude that a piece rate will generally boost the performance of participants. However, when controlling for possible interactions of the dummy variables, it becomes apparent that this is driven by the performance decline of men who receive a fixed payment in the second part of the experiment. Regressions (3) and (4) apply the situation where men start to work under a fixed payment as a benchmark. Strikingly, the regressions only identify one significant departure from this benchmark if we control for the case where men receive the fixed payment in part two (Fix $\times$ Male $\times$ Part 2).[25] This confirms the main result and emphasizes that this is the most conspicuous finding.

Regressions (2) and (4) demonstrate that the finding is robust when controlling for further potential drivers. The models include additional control variables such as participants' age and a dummy indicating whether or not a subject is an economics student. However, the coefficients for these two variables are insignificant in both regressions. Apart from that, we also check for the influence of the variables *Fun* and *Period*.[26] *Fun* is a measure of how much participants

[24] *Piece-to-Fix*, *Piece-to-Fix10*, *Fix-to-Fix*, and *Piece-to-Piece*. We only control for the paid remuneration but not for the level of the fixed payment.

[25] There is one further potential effect (Piece $\times$ Female $\times$ Part 2), but it is only weakly significant.

[26] Note that *Fun* is based on subjects' self-assessment in a post-experimental questionnaire. Subjects were asked

	(1) Performance	(2) Performance	(3) Performance	(4) Performance
Dummy variables				
Piece	0.621***	0.595***		
	(0.220)	(0.220)		
Female	0.769***	0.505***		
	(0.206)	(0.194)		
Part 2	0.023	0.006		
	(0.158)	(0.157)		
Interaction terms				
Fix × Male × Part 1			baseline	baseline
			–	–
Fix × Male × Part 2			-0.892**	-0.953***
			(0.347)	(0.334)
Fix × Female × Part 1			-0.078	-0.328
			(0.444)	(0.411)
Fix × Female × Part 2			0.332	0.008
			(0.394)	(0.365)
Piece × Male × Part 1			0.047	-0.018
			(0.387)	(0.378)
Piece × Male × Part 2			0.556	0.591
			(0.578)	(0.559)
Piece × Female × Part 1			0.583	0.245
			(0.387)	(0.362)
Piece × Female × Part 2			1.069**	0.801*
			(0.491)	(0.468)
Control variables				
Age		-0.025		-0.026
		(0.026)		(0.026)
Fun		0.234***		0.237***
		(0.043)		(0.042)
Period		0.026***		0.026***
		(0.006)		(0.006)
Econ		-0.081		-0.062
		(0.226)		(0.224)
Constant	8.875***	8.140***	9.506***	8.809***
	(0.245)	(0.629)	(0.362)	(0.655)
Observations	5,660	5,660	5,660	5,660
R^2	0.048	0.120	0.062	0.135
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$				

Table 3: OLS regressions of subjects' performance in the real-effort task. Clustered standard errors in parentheses.

like the task, whereas *Period* controls for the number of the period within the current part. The variables reveal further insights, i.e., *Period* is significantly positive and remarkably stable in both regressions. This may be due to subjects who become more efficient at doing the task in later periods. *Fun* is also significantly positive in both regressions. One interpretation is that subjects who have a higher intrinsic motivation to do the task may perform better. The introduction of the control variables even emphasizes our main finding. The picture drawn by regression (4) is much clearer than the one from regression (3), and it once more documents that men who receive a fixed payment in the second part of the experiment exert a substantial lower performance induced by motivational crowding out.

4 Conclusion

In the current paper, we presented a real-effort experiment to isolate the effects of incentive changes on intrinsic motivation. The findings highlight the occurrence of considerable crowding-out effects when performance-based remuneration schemes were discontinued. We also show that these effects were exclusively driven by men. A sizable share of them reduced their efforts drastically after piece-rate incentives had been substituted by a fixed payment. By contrast, the behavior of women was virtually unaffected by such changes in the payment scheme.

Our identification strategy was as follows. The main treatment revealed that the average performance of male participants declined substantially when piece-rate incentives were substituted by a fixed payment. It was contemplated with three control treatments and a post-experimental questionnaire. These controls were implemented to isolate the drivers of our main result. Altogether, they clearly demonstrated that the decline of performance was caused by a decrease of subjects' intrinsic motivation. The alternative explanations could be ruled out using the control treatments and the questionnaire.

Our study makes three contributions. First, our research contributes to a better understanding of the economic literature on behavioral-incentive effects. Second, the diverse gender effects on remuneration changes may be of interest to managers and other decision-makers. Third, the study may shed new light on the broader topic of gender differences in the labor market and the determinants of the gender wage gap. In the following paragraphs, we comment on these aspects in more detail.

Our first contribution is that the present paper fills an important gap in the economic literature on incentives by connecting several streams of research. For instance, in economics, empirical work has documented that workers' productivity may decline if piece rates are replaced by hourly wages (e.g., Freeman and Kleiner, 2005). We add to this research by showing that such a decline may often be driven by a crowding out of intrinsic motivation and not just by the lack of extrinsic incentives per se. Similarly, prior work in behavioral economics has emphasized that paying higher incentives might often backfire and lower subjects' intrinsic motivation.

on a scale of 1–10 (1: no fun; 10: great fun) to assess the level of fun they had when doing the task. Also note that *Period* does not correlate with *Part 2* since it only counts the periods within a part.

Examples include: blood donations (Mellström and Johannesson 2008), the acceptance of nuclear sites in the vicinity (Frey 1994), and doing IQ tasks (Gneezy and Rustichini 2000b). Our contribution to this strand of literature is that we highlight that removing existing financial incentives may have equally adverse effects. The psychological literature already reports some evidence of this aspect (e.g., Deci, 1972). However, our study adds another important point, as it applies a controlled lab experiment to reveal the strong discrepancy between the reactions of men and women. The fact that women and men may react differently to certain incentive structures is well-known in the economics literature (e.g., Gneezy et al., 2003; Niederle and Vesterlund, 2007; Croson and Gneezy, 2009).[27] To the best of our knowledge, our paper is, however, the first to show that gender differences in the motivational crowding out of work performance exist. A potential explanation may be that the different remuneration schemes induce momentum effects for our workers. It is possible that receiving financial incentives is coded as "win", whereas the change to a fixed payment is coded as "loss." Recently, Cohen-Zada et al. (2017) show in natural experiment data that men's performance in Judo contests is significantly affected by psychological momentum, whereas women do not show this effect.[28] Put differently, the history of events matters, i.e., men who lost a previous match showed a lower performance than men who won it. By contrast, women were not prone to the history of events. In our case, changing piece-rate incentives to a fixed payment may have similar effects as a negative momentum. This may explain why only men decrease their performance.

Second, our results may be of interest to managers and other decision-makers in business. Of course, our setup is stylized and cannot be directly applied to firms. However, by using this abstract design, we were able to show that a change in the payment scheme may have very adverse effects on workers' intrinsic motivation and that the gender composition may play a key role. Both aspects are potentially interesting to managers and other decision-makers when they are in the process of restructuring their companies. Notably, the insights of the paper suggest that changes to the payment scheme may come with a side effect that counteracts their original purpose. The data also point out that the gender composition of the corresponding firm matters.[29] Thus, if a restructuring plan involves a change from performance pay to hourly wages, the success of this plan may crucially depend on the number of male and female workers affected by this change. Our research suggests that such plans are more likely to be successful if there are many female workers while they may be harmful in male-dominated branches.

Our results may also shed new light on broader topics of gender differences in the labor market like the gender pay gap or so-called glass-ceiling effects. Naturally, our stylized setup cannot account for the entire complexity of these issues, but we believe that it may add some valuable insights. We demonstrate that, in contrast to men, women always show high per-

[27] Wozniak et al. (2014) show that performance feedback may mitigate gender gaps in the willingness to compete.

[28] The authors point out that this is in line with the biological literature on testosterone. Here, it is shown that higher testosterone levels after positive (negative) events increase (decrease) performance for men only (e.g., Mazur et al. 1992; Bateup et al. 2002).

[29] See also Apesteguia et al. (2012) who highlight the importance of gender-composition effects in a data set of a large business game. Focusing on MBA students, the authors show that mixed gender teams perform best as general managers.

formance when working in non-incentivized environments. This finding may be part of an explanation of glass-ceiling effects, i.e., the phenomenon that women are rarely promoted to top positions.[30] If superiors realize that women perform well absent of high incentives, while men substantially reduce effort if they are not promoted, this may bias promotion decisions toward male applicants.

References

Apesteguia, J., Azmat, G., Iriberri, N., 2012. The impact of gender composition on team performance and decision making: Evidence from the field. Management Science 58 (1), 78–93.

Azmat, G., Iriberri, N., 2016. The provision of relative performance feedback: An analysis of performance and satisfaction. Journal of Economics and Management Strategy 25 (1), 77–110.

Bateup, H. S., Booth, A., Shirtcliff, E. A., Granger, D. A., 2002. Testosterone, cortisol, and women's competition. Evolution and Human Behavior 23 (3), 181–192.

Bénabou, R., Tirole, J., 2003. Intrinsic and extrinsic motivation. Review of Economic Studies 70 (3), 489–520.

Bénabou, R., Tirole, J., 2006. Incentives and prosocial behavior. American Economic Review 96 (5), 1652–1678.

Benndorf, V., Rau, H. A., Sölch, C., 2014. Minimizing learning behavior in experiments with repeated real-effort tasks. Available at SSRN 2503029.

Blackaby, D., Booth, A. L., Frank, J., 2005. Outside offers and the gender pay gap: Empirical evidence from the UK academic labour market. Economic Journal 115 (501), F81–F107.

Bowles, S., Polania-Reyes, S., 2012. Economic incentives and social preferences: Substitutes or complements? Journal of Economic Literature 50 (2), 368–425.

Cohen-Zada, D., Krumer, A., Shtudiner, Z., 2017. Psychological momentum and gender. Journal of Economic Behavior & Organization 135, 66–81.

Corgnet, B., Hernán-González, R., Schniter, E., 2015. Why real leisure really matters: Incentive effects on real effort in the laboratory. Experimental Economics 18 (2), 284–301.

Cotton, C., McIntyre, F., Price, J., 2013. Gender differences in repeated competition: Evidence from school math contests. Journal of Economic Behavior & Organization 86, 52–66.

[30] See, for example, Blackaby et al. (2005) who report a significant gender promotion gap in the UK academic labor market. Similarly, Newton and Simutin (2015) report empirical evidence demonstrating that female officers receive smaller raises if the firm is headed by male CEOs.

Croson, R., Gneezy, U., 2009. Gender differences in preferences. Journal of Economic Literature 47 (2), 448–474.

Deci, E. L., 1971. Effects of externally mediated rewards on intrinsic motivation. Journal of Personality and Social Psychology 18 (1), 105–115.

Deci, E. L., 1972. Intrinsic motivation, extrinsic reinforcement, and inequity. Journal of Personality and Social Psychology 22 (1), 113–120.

Dickinson, D. L., 1999. An experimental examination of labor supply and work intensities. Journal of Labor Economics 17 (4), 638–670.

Dohmen, T., Falk, A., 2011. Performance pay and multidimensional sorting: Productivity, preferences, and gender. American Economic Review 101 (2), 556–590.

Ederer, F., Manso, G., 2013. Is pay for performance detrimental to innovation? Management Science 59 (7), 1496–1513.

Erkal, N., Gangadharan, L., Nikiforakis, N., 2011. Relative earnings and giving in a real-effort experiment. American Economic Review 101 (7), 3330–3348.

Falkinger, J., Fehr, E., Gächter, S., Winter-Ebmer, R., 2000. A simple mechanism for the efficient provision of public goods: experimental evidence. American Economic Review 90 (1), 247–264.

Fernie, S., Metcalf, D., 1999. It's not what you pay it's the way that you pay it and that's what gets results: Jockeys' pay and performance. Labour 13 (2), 385–411.

Fischbacher, U., 2007. z-tree: Zurich toolbox for ready-made economic experiments. Experimental Economics 10 (2), 171–178.

Freeman, R. B., Kleiner, M. M., 2005. The last american shoe manufacturers: Decreasing productivity and increasing profits in the shift from piece rates to continuous flow production. Industrial Relations: A Journal of Economy and Society 44 (2), 307–330.

Frey, B. S., 1994. How intrinsic motivation is crowded out and in. Rationality and Society 6 (3), 334–352.

Frey, B. S., Oberholzer-Gee, F., 1997. The cost of price incentives: An empirical analysis of motivation crowding-out. American Economic Review 87 (4), 746–755.

Gächter, S., Kessler, E., Königstein, M., 2011. The roles of incentives and voluntary cooperation for contractual compliance. IZA Discussion Paper No. 5774.

Gneezy, U., Meier, S., Rey-Biel, P., 2011. When and why incentives (don't) work to modify behavior. Journal of Economic Perspectives 25 (4), 191–210.

Gneezy, U., Niederle, M., Rustichini, A., 2003. Performance in competitive environments: Gender differences. Quarterly Journal of Economics 118 (3), 1049–1074.

Gneezy, U., Rustichini, A., 2000a. A fine is a price. Journal of Legal Studies 29 (1), 1–17.

Gneezy, U., Rustichini, A., 2000b. Pay enough or don't pay at all. Quarterly Journal of Economics 115 (3), 791–810.

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

Iriberri, N., Rey-Biel, P., 2014. On women's underperformance in competitive environments: When and why. Working Paper.

Irlenbusch, B., Sliwka, D., 2005. Incentives, decision frames, and motivation crowding out-an experimental investigation. IZA Discussion Paper No. 1758.

Kuhn, P., Villeval, M. C., 2015. Are women more attracted to co-operation than men? Economic Journal 125 (582), 115–140.

Lazear, E. P., 2000. The power of incentives. American Economic Review 90 (2), 410–414.

Majeres, R. L., 1983. Sex differences in symbol-digit substitution and speeded matching. Intelligence 7 (4), 313–327.

Masclet, D., Peterle, E., Larribeau, S., 2015. Gender differences in tournament and flat-wage schemes: An experimental study. Journal of Economic Psychology 47, 103–115.

Mazur, A., Booth, A., Dabbs Jr, J. M., 1992. Testosterone and chess competition. Social Psychology Quarterly, 70–77.

Mellström, C., Johannesson, M., 2008. Crowding out in blood donation: Was Titmuss right? Journal of the European Economic Association 6 (4), 845–863.

Newton, D., Simutin, M., 2015. Of age, sex, and money: Insights from corporate officer compensation on the wage inequality between genders. Management Science 61 (10), 2355–2375.

Niederle, M., Vesterlund, L., 2007. Do women shy away from competition? Do men compete too much? Quarterly Journal of Economics 122 (3), 1067–1101.

Pokorny, K., 2008. Pay - but do not pay too much: an experimental study on the impact of incentives. Journal of Economic Behavior & Organization 66 (2), 251–264.

Schnedler, W., Vanberg, C., 2014. Playing 'hard to get': An economic rationale for crowding out of intrinsically motivated behavior. European Economic Review 68, 106–115.

Shearer, B., 2004. Piece rates, fixed wages and incentives: Evidence from a field experiment. Review of Economic Studies 71 (2), 513–534.

Sliwka, D., Werner, P., forthcoming. How do agents react to dynamic wage increases? An experimental study. Journal of Labor Economics.

Titmuss, R. M., 1970. The gift relationship: From human blood to social policy. Allen & Unwin London.

Wozniak, D., Harbaugh, W. T., Mayr, U., 2014. The menstrual cycle and performance feedback alter gender differences in competitive choices. Journal of Labor Economics 32 (1), 161–198.

A Additional tables (for online appendix)

		Women				Men			
		Part 1		Part 2		Part 1		Part 2	
Treatment	Per.	mean	sd	mean	sd	mean	sd	mean	sd
Piece-to-Fix	1	9.94	1.75	10.20	1.51	8.93	1.79	8.18	2.78
	2	10.11	1.59	10.23	1.67	9.52	1.62	8.28	3.10
	3	10.30	1.69	10.22	1.67	9.43	1.65	8.43	2.73
	4	10.33	1.55	10.00	2.08	9.55	1.48	8.70	2.98
	5	10.44	1.62	10.00	2.07	9.38	1.45	8.23	3.06
	6	10.22	1.84	10.25	1.79	9.35	1.61	8.37	2.92
	7	10.33	1.87	10.22	2.14	9.47	1.50	8.12	3.38
	8	10.47	1.75	10.02	2.10	9.67	1.39	8.30	3.23
	9	10.56	1.74	10.11	2.22	9.65	1.54	8.20	2.86
	10	10.58	1.95	10.28	1.94	9.52	1.36	7.95	2.93
Fix-to-Fix	1	9.22	1.74	9.31	1.86	9.28	1.65	9.31	2.19
	2	9.09	1.77	9.69	1.77	9.31	2.13	9.53	2.59
	3	9.22	1.81	9.50	2.13	9.25	2.37	9.72	2.68
	4	9.44	1.83	9.81	2.49	9.41	2.38	9.91	2.39
	5	9.56	1.58	9.72	2.00	9.50	2.27	9.88	2.69
	6	9.34	1.73	9.38	2.45	9.59	2.45	9.91	2.25
	7	9.38	1.72	9.34	2.52	9.34	2.36	9.72	2.41
	8	9.62	1.56	9.66	2.38	9.81	2.51	9.69	2.48
	9	9.56	1.54	9.59	2.45	9.81	2.46	9.59	2.60
	10	9.84	1.69	9.72	2.14	9.75	2.46	9.47	2.72
Piece-to-Piece	1	9.56	1.36	10.69	1.08	9.12	2.25	9.94	2.46
	2	10.06	1.65	10.69	1.66	9.31	2.21	9.94	2.08
	3	10.25	1.57	10.56	1.82	9.69	2.06	10.19	1.76
	4	10.00	1.32	10.56	1.41	9.75	1.73	9.94	1.69
	5	9.88	1.75	10.44	1.79	10.06	1.98	10.12	1.75
	6	10.00	1.10	10.44	1.75	10.25	1.73	9.94	2.21
	7	9.94	2.02	10.50	1.41	10.00	1.79	10.19	2.01
	8	10.50	1.51	10.50	1.79	10.00	2.16	10.19	1.94
	9	9.94	1.69	10.56	1.75	9.88	2.16	10.31	2.21
	10	10.31	1.49	10.81	1.38	10.19	1.56	9.88	2.06
Piece-to-Fix10	1	9.00	1.73	9.39	1.71	9.19	1.51	8.25	2.83
	2	9.52	1.59	9.13	1.54	9.56	1.81	8.50	3.20
	3	9.45	1.61	9.52	2.10	9.59	1.85	8.06	3.44
	4	9.58	1.63	9.61	1.82	9.69	1.55	8.31	3.13
	5	9.65	1.68	9.52	2.35	9.50	2.49	8.44	3.36
	6	9.77	1.54	9.55	2.17	9.66	2.06	8.00	3.45
	7	9.84	1.46	9.55	1.71	9.56	1.97	7.94	3.45
	8	9.61	1.63	9.71	1.85	9.78	1.54	8.16	3.76
	9	9.74	1.24	9.19	1.87	9.69	1.55	8.28	3.34
	10	10.06	1.44	9.48	2.11	9.94	1.66	7.97	3.47

Table 4: Subjects' average performance an standard deviations in the treatments.

B Additional figures (intended as an on-line appendix)

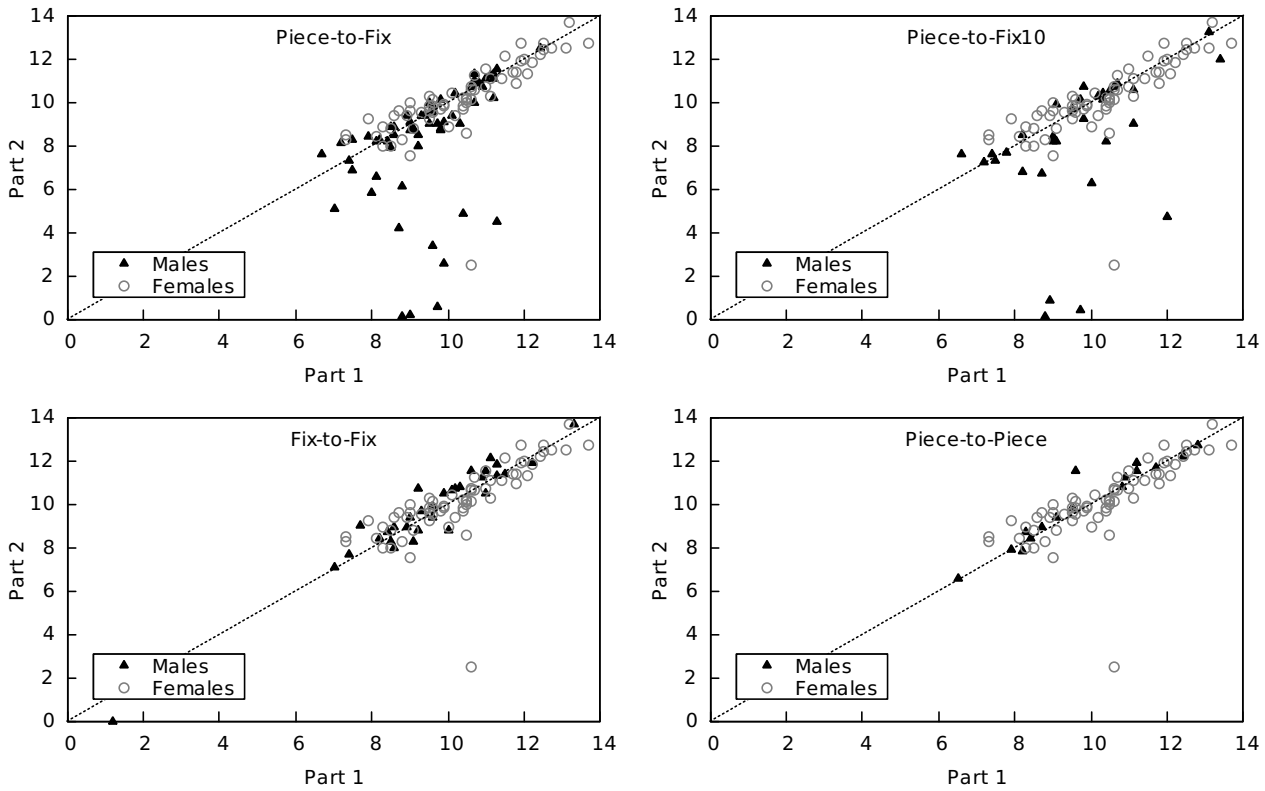


Figure 4: Individual performance of men and women in parts 1 and 2 for all treatments.

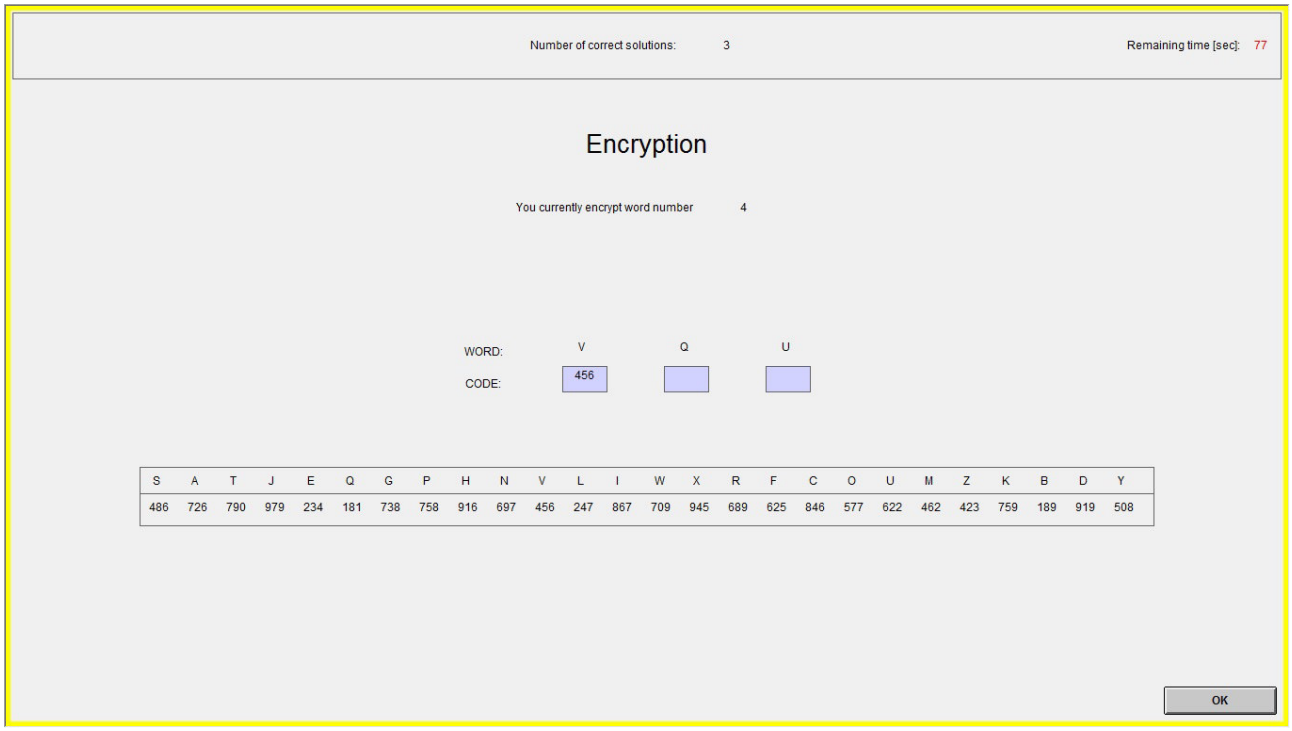
C **Instructions** (translated from German; not intended for publication)

C.1 Instructions describing the task and the trial period

In the following experiment you have the opportunity to earn money depending on your behavior. Please turn off your mobile phone and do not talk to other participants in the experiment. It is very important that you follow these rules. If you have any question while reading these instructions or during the experiment itself, we ask you to raise your hand. We will immediately come to your desk and answer your question individually.

1. General structure of the experiment

During the experiment you have the opportunity to do a task. The task consists of encoding combinations of letters (words) into numbers. In the task, three capital letters always yield a "word". You have to allocate a number to each capital letter. The encryption code can be found in a table below the corresponding letter. For that purpose, please consider the following screenshot:



In this example the participant has already encrypted three words correctly (see centered field: above). Here, the three capital letters: "V", "Q" and "U" have to be encoded. The solution follows immediately from the table:

- For "V" applies: 456 (see the current entry of the participant)
- For "Q" applies: 181
- For "U" applies: 622

To make an input please click on the blue box below the first capital letter.

Furthermore, the screen (see screenshot) provides the following information:

- "Number of correct solutions" = number of correctly encrypted words.
- "Remaining time [sec]" = remaining time in the current period.
- "You currently encrypt word number" = current word to encrypt.

If all 3 numbers have been entered, please click the "OK". The computer then checks whether all capital letters haven been encoded correctly. Only then the word is counted as correctly solved. Thereafter a new word (again consisting of three capital letters) is randomly drawn.

Furthermore, a new encryption table is randomly generated in two steps:

- The computer program randomly selects in the table a new set of three-digit numbers to be used for the encoding of the capital letters.
- Additionally, the computer program shuffles the position of the capital letters in the table. Please note that the program always uses all 26 capital letters of the German alphabet. Please note that if a new word appears, you have to click with your mouse on the first of the three blue boxes. Otherwise no input is possible!

The computer will mark (in red font) wrong inputs after pressing the "OK" button.

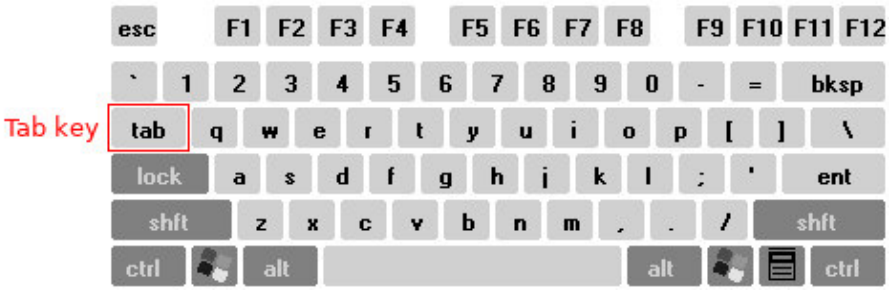
Bear in mind:

- After wrong inputs the current word to encode will not change until a correct input was made.
- However, your previous inputs (in the 3 boxes below the capital letters) will all be deleted.
- Furthermore, the table stays unaltered, meaning that the allocated numbers remain identical. Also the position of the capital letters in the table does not change.

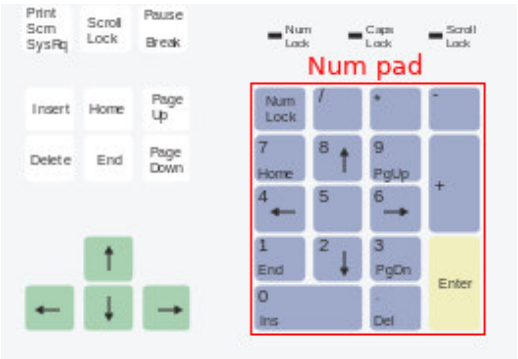
Important hints:

Please note that after having entered the three-digit number you can easily switch to the next blue box by using the tabulator key on your keyboard.

In the following picture you can see the position of the tabulator key on your keyboard:



The input of the numbers can be performed faster by using the numpad (on the right) of your keyboard. In the following picture you can see the position of the numpad on your keyboard:



2. Trial period:

The experiment starts with a trial period in which each participant has to encrypt exactly 10 words. Please note: Correct solutions do not lead to payments within the trial period. The general idea of the trial period is to make you as familiar as possible with the task before the actual experiment begins. Therefore you should take the trial period serious and try to solve the ten words as fast as possible!

Please raise your hand if you still have further questions. We will come to your desk and answer them individually.

C.2 Instructions for parts with piece rate

Please note: The experiment consists of 2 parts, whereas this part encompasses 10 periods. In each of these periods you have the possibility to do the same task: encoding words to numbers.

- Each period will last exactly 2 minutes.
- After the time has expired the corresponding period will be finished. Then the program only counts the words you correctly solved to the result of the period.

In this part of the experiment, your payoff only depends on the number of correctly-solved words. You will receive 8 Cent for each correctly solved word.

It applies that:

- After each period you will see an information screen presenting the number of correctly solved words of the past period.
- The next period will automatically start 10 seconds later.

This part of the experiment will end after the end of the 10 periods. Then you will be informed on the payoff you earned in this part. Please remain seated after the end of this part. We will inform you about the further process within a short time.

Please raise your hand if you should have any further questions. We will then come to your desk and answer it individually.

C.3 Instructions for parts with the €8 fixed payment

Please note: The experiment consists of 2 parts, whereas this part encompasses 10 periods. In each of these periods you have the possibility to do the same task: encoding words to numbers.

- Each period will last exactly 2 minutes.
- After the time has expired the corresponding period will be finished. Then the program only counts the words you correctly solved to the result of the period.

In this part of the experiment you will receive a lump sum of 8 Euro. Your payoff here does not depend on the number of correctly solved words.

It applies that:

- After each period you will see an information screen presenting the number of correctly solved words of the past period.
- The next period will automatically start 10 seconds later.

This part of the experiment will end after the end of the 10 periods. Then you will be informed on the payoff you earned in this part. Please remain seated after the end of this part. We will inform you about the further process within a short time.

Please raise your hand if you should have any further questions. We will then come to your desk and answer it individually.

C.4 Instructions for parts with the €10 fixed payment

Please note: The experiment consists of 2 parts, whereas this part encompasses 10 periods. In each of these periods you have the possibility to do the same task: encoding words to numbers.

- Each period will last exactly 2 minutes.
- After the time has expired the corresponding period will be finished. Then the program only counts the words you correctly solved to the result of the period.

In this part of the experiment you will receive a lump sum of 10 Euro. Your payoff here does

It applies that:

- After each period you will see an information screen presenting the number of correctly solved words of the past period.
- The next period will automatically start 10 seconds later.

This part of the experiment will end after the end of the 10 periods. Then you will be informed on the payoff you earned in this part. Please remain seated after the end of this part. We will inform you about the further process within a short time.

Please raise your hand if you should have any further questions. We will then come to your desk and answer it individually.