Güth, Werner; Paul, Christian

Working Paper

Downsizing the labor force by low and high profit firms: an experimental analysis

Jena Economic Research Papers, No. 2008,087

Provided in Cooperation with:
Max Planck Institute of Economics

Suggested Citation: Güth, Werner; Paul, Christian (2008) : Downsizing the labor force by low and high profit firms: an experimental analysis, Jena Economic Research Papers, No. 2008,087, Friedrich Schiller University Jena and Max Planck Institute of Economics, Jena

This Version is available at:
https://hdl.handle.net/10419/31719

JENA ECONOMIC RESEARCH PAPERS

2008 – 087

Downsizing the Labor Force by Low and High Profit Firms – An Experimental Analysis

by

Werner Güth
Christian Paul

www.jenecon.de

ISSN 1864-7057

The JENA ECONOMIC RESEARCH PAPERS is a joint publication of the Friedrich Schiller University and the Max Planck Institute of Economics, Jena, Germany. For editorial correspondence please contact markus.pasche@uni-jena.de.

Impressum:

Friedrich Schiller University Jena
Carl-Zeiss-Str. 3
D-07743 Jena
www.uni-jena.de

Max Planck Institute of Economics
Kahlaische Str. 10
D-07745 Jena
www.econ.mpg.de

Jena Economic Research Papers 2008 - 087

Downsizing the Labor Force by Low and High Profit Firms – An Experimental Analysis[1]

Werner Güth[2] and Christian Paul[3]

November 10, 2008

[1]We are grateful to Karl-Martin Ehrhart and Ralf Löschel for helpful comments.

[2]Max Planck Institute of Economics, Kahlaische Straße 10, D-07745 Jena, Germany (e-mail: gueth@econ.mpg.de).

[3]Institut für Wirtschaftstheorie und Operations Research, University of Karlsruhe, D-76128 Karlsruhe, Germany (e-mail: paul@wiwi.uni-karlsruhe.de).

Abstract

One may hope to capture the behavioral and emotional effects of downsizing the labor force in rather abstract settings as an ultimatum game (see Fischer et al. (2008)), or try to explore downsizing in its more natural principal-agent scenario with a labor market background. We pursue the latter approach and test experimentally whether downsizing occurs whenever (game) theoretically predicted and whether effort reactions question its profitability. Our main findings are that downsizing seems to happen less often than predicted and that its frequency does not depend on whether, theoretically, its gains are rather large or small. Interestingly, we also find strong evidence that piece-rate offers are used in a suboptimal way.

JEL classification: C72, C91, D21, J01

Keywords: downsizing, experimental economics, principal-agent model, labor economics

1 Introduction

Downsizing is understood as laying off a large group of workers and has become a synonym for neoliberalism in popular (science) literature.[1] More broadly, downsizing could refer to all situations where a subgroup of interacting parties within an organization would have better success or survival prospects than the whole group.[2] One reaction to such challenges could be volunteers offering to be excluded. But such "*hero volunteers*" are probably a rare species. What one realistically has to expect are attempts of some parties to exclude others against their will. This does not only invoke material aspects but also raises deep moral and emotional concerns not only of those who suffer, e.g., by being excluded, but also of those who exclude others.

One can hope to capture crucial aspects of such conflicts by employing rather abstract scenarios like ultimatum games, e.g., with one proposer and several responders, some of whom can be excluded (Fischer et al. (2008)). But then questions like *"Why can the proposer and some responders exclude other responders?"* or *"Why can the remaining players share more, and how is that related to what all would receive?"* would naturally arise. These are less troublesome when considering situations of which downsizing is typical, namely a firm which tries to reduce its labor force, although it is prospering.

In a principal-agent setting, the initiative for downsizing naturally rests on the owners or respectively their delegates, e.g., CEOs. What can be gained by downsizing is implied by the economic, technological, and legal environment. A smaller labor force may be the only chance to avoid bankruptcy, resembling the example of a lifeboat whose supplies do not suffice for all. One may derive important insights by investigating such situations where either all suffer or some survive. But here we focus on situations where downsizing is not a matter of immediate necessity but one of profitability.

Such – allegedly unfair – downsizing announcements regularly alert the public all over the world (NYTimes (2008)), e.g., in Germany in the recent past. In January 2008, the mobile phone producer Nokia announced to shut down a factory in Bochum, Germany, and lay off 4,300 full-time employees and temporary workers, although internal accounting showed a profit of 134

[1] One prominent example is the best-seller *"Downsize This!: Random Threats from an Unarmed American"* by author and filmmaker Michael Moore.

[2] Dramatic examples of lifeboats with too little water or food for all can be found in novels and are hopefully more often fictional than factual.

million euro.[3] In February 2008, the household products company Henkel and the automobile manufacturer BMW announced layoffs of 3,000 and 8,100 workers, respectively, although their profits had increased (to about 1 billion euro and more than 3.75 billion euro, respectively). These firms justified downsizing by future risks due to the Global Economy (Henkel) or simply by higher rentability aspirations (BMW).[4]

Rather than discussing such arguments at length, we experimentally examine the occurrence and the behavioral effects of downsizing in a principal-agent setting with minimum wages as a labor market rigidity. In the treatment where a principal's profit increase from downsizing his labor force is positive, but rather small, we expect layoffs to occur less often than in the other treatment where this gain is rather large.

Our research questions are: Do game theoretic benchmark solutions predict contract offers and agents' behavior? Will there be differences between anticipated and unanticipated downsizing and, if so, will they question the predicted profitability of downsizing? And how will the remaining agents react to witnessing layoffs?

Major results are that a) downsizing occurs less often than (game) theoretically predicted and b) is relatively independent of the theoretical gains from downsizing; c) firms which downsize are mostly those earning less; d) individual payoffs and welfare are lower than predicted due to surprisingly low piece rates; e) there are no significant differences between anticipated and unanticipated downsizing; and f) there is no evidence that remaining agents react to layoffs of their coworkers.

The specific firm model, which we analyze theoretically and have implemented experimentally, is introduced in section 2, including its solution. Section 3 describes the experimental protocol and our hypotheses. After analyzing the data in section 4, we conclude in section 5.

[3]Nokia executives stated that these internal numbers are due to accounting regularities and do not represent the factual profitability of the factory at all (see FAZ(2008b)).

[4]See, e.g., Handelsblatt (2008) or Frankfurter Rundschau (2008) for detailed information concerning the downsizing announcements, profits, and reactions. Official numbers for BMW have not yet been published, but were estimated by CEO Norbert Reithofer in 2007 (see FAZ (2008a)). The automobile manufacturer's rate of return is claimed to have dropped below 6 %. Union representatives complained about a loss of corporate culture when downsizing was announced.

2 The principal-agent model

■ **Model description:** Principal P currently employs both highly productive agents ($h = 1, \ldots, m$) with symmetric cost functions of effort e_h, $C_h(e_h)$ and less productive ones ($l = 1, \ldots, n$) with cost functions $C_l(e_l)$. Both groups produce the same kind of output. Highly productive agents[5] have lower costs of effort, i.e., $C_h(e_h)$ is smaller than $C_l(e_l)$ for all positive effort levels e. More specifically, we rely on quadratic effort cost functions $C_h(e_h) = \frac{k}{2}e_h^2$ and $C_l(e_l) = \frac{d}{2}e_l^2$ with $0 < k < d$. Each unit of effort corresponds to a (> 0) units of output. We neglect asymmetric information; the principal can perfectly observe the agents' types and the amount of output produced.[6]

To allow for profit-increasing downsizing, we assume the labor market rigidity of minimum wages and impose nondiscriminatory contract offers for all workers.[7] Workers' outside options like unemployment benefits are denoted by U and are assumed to be smaller than the minimum fixed wage, M. The principal is assumed to be a price taker on the sales market with product price p. A linear employment contract specifies the fixed wage F and the same piece rates for all workers. We do not explicitly model labor market competition.[8] The principal's profit, Π, is

$$\Pi = (ap - r)[\sum_{h=1}^{m} e_h + \sum_{l=1}^{n} e_l] - (m+n)F \; . \tag{1}$$

Employees' earnings, ω, are

$$\omega_h = F + re_h - \frac{k}{2}e_h^2 \qquad \text{and} \qquad \omega_l = F + re_l - \frac{d}{2}e_l^2 \tag{2}$$

[5] We use "agent", "worker" and "employer" synonymously in the remainder and alternate between male and female pronouns for the principal.

[6] One can justify co-employment of more and less productive workers by new production techniques which are more easily adopted by some, e.g., the younger workers, but this questions the productivity of others who before were equally skilled. Many other principal-agent models use the same or similar convex functions (see, e.g., Richter/Furubotn (2003)).

[7] Collective wage agreements or strict antidiscrimination laws justify these assumptions. Even when the principal is aware of different productivities of different agent types, this may not be verifiable or, if so, may even be disregarded in labor court cases.

[8] This is a standard assumption of principal-agent models. See Berninghaus et al. (2007a/2007b) who discuss models of duopolistic sales and hiring competition.

for highly productive agents h and less productive agents l, respectively.

The game is played finitely often. In the first x stages, the principal employs all agents; downsizing is impossible in these stages by assumption. Output is produced, learned by all parties, and sold. Profits, efforts, and earnings are assumed to be common knowledge. After the first x stages the principal can lay off part of his labor force, i.e., downsize. More specifically, the group of less productive agents may be dismissed while the more productive agents remain in the firm. In the last y stages, output is produced by the agents still employed. Although game theoretically, it does not matter whether the principal and/or the agents know ex ante that downsizing is possible after x stages, behaviorally this might matter.[9] We now solve the downsizing game by backward induction.

■ **Effort decisions:** From (2) we can immediately derive the agents' optimal effort choices when being employed. They are $e_h^* = r/k$ for highly productive agents h and $e_l^* = r/d$ for less productive agents l, respectively.

Agents obviously want to be employed since by assumption $M > U$.[10] By inserting optimal efforts into the principal's profit function, we get

$$\Pi = (a \cdot p - r) \cdot [m \cdot \frac{r}{k} + n \cdot \frac{r}{d}] - (m + n) \cdot F \ . \tag{3}$$

■ **Contracts without downsizing:** Denoting the contract offer without downsizing, i.e., the contract prior to the downsizing decision, by $(\tilde{F}, \tilde{r})$, the principal maximizes (3) subject to the minimum wage constraint (MWC1) $\tilde{F} \geq M$, which is obviously binding in optimum.

The optimal piece rate[11] and the resulting effort levels are $\tilde{r}^* = ap/2$, $e_h^* = ap/(2k)$, and $e_l^* = ap/(2d)$. Both workers earn more than the outside option while the less productive workers earn less than the highly productive ones:

$$U \ < \ \tilde{\omega_l} = M + \frac{a^2p^2}{8d} \ < \ \tilde{\omega_h} = M + \frac{a^2p^2}{8k} \qquad \text{due to } k < d \ .$$

[9] Anticipation of potential downsizing may serve as a "Sword of Damocles" (see experimental design and discussion).

[10] To avoid further complexity, we thus refrain from giving agents the option to quit in our theoretical analysis and in our experiment.

[11] It obviously fulfills the non-negativity constraint; Second-order conditions hold.

The principal's profit,

$$\tilde{\Pi} = \frac{dm + kn}{4dk} a^2 p^2 - (m+n)M \ ,$$

is positive for

$$M \leq B_1 := \frac{dm + kn}{4dk(m+n)} a^2 p^2 \ .$$

■ **Contracts after downsizing:** After laying off the group of less productive workers, the principal's contract offer, denoted by $(\hat{F}, \hat{r})$, maximizes

$$(ap - \hat{r}) \cdot m \cdot \frac{\hat{r}}{k} - m \cdot \hat{F}$$

subject to the minimum wage constraint (MWC2) $\hat{F} \geq M$. The optimal piece rate and effort level are $\tilde{r}^* = ap/2$ and $e_h^* = ap/(2k)$. But only the highly productive workers earn more than the outside option:

$$\hat{\omega_l} = U \ < \ \hat{\omega_h} = M + \frac{a^2 p^2}{8k} \ .$$

The principal's profit,

$$\hat{\Pi} = \frac{m}{4k} a^2 p^2 - m \cdot M \ ,$$

is positive for

$$M \leq B_2 := \frac{1}{4k} a^2 p^2 \ .$$

■ **Downsizing** is profitable for the principal if

$$\hat{\Pi} - \tilde{\Pi} > 0 \qquad \text{or} \qquad n(-\frac{a^2 p^2}{4d} + M) > 0 \ .$$

This holds true for $d \to \infty$ or, more generally, for

$$M > B_3 := \frac{a^2 p^2}{4d} \ .$$

It is easy to show that $B_3 < B_1 < B_2$ always holds.[12] Thus, to guarantee that profits are nonnegative and downsizing is profitable, it suffices to impose $B_3 < M < B_1 < B_2$ for the minimum wage, what is fulfilled in all our experimental treatments. By varying M between B_1 and B_3, we generated treatments with differing downsizing profitability.

3 Experimental design

■ **Basic Design:** The experimental design follows the principal-agent model. In the beginning, three participants interact: one principal (P–participant), one highly productive agent (H-participant), and one less productive agent[13] (L-participant). Eight such triplets of participants formed a session with 24 participants. The experiment consisted of $x = 2$ rounds without the possibility to downsize and $y = 2$ rounds afterwards. We refer to these four rounds as "first phase." During the first phase, participants did not know that a perfect stranger repetition of the same four rounds (second phase) would be played afterwards. They were told, however, that another experiment would follow and that they would definitively not interact with the same participants again.

We distinguished between the following treatments: In the **announced downsizing, high incentive–treatment (AH)** all participants knew from the beginning that after the first $x = 2$ rounds the principal could downsize and that two more rounds would be played thereafter. Furthermore, the principal's theoretical profit gain from downsizing was rather large.

The only difference in the **announced downsizing, low incentive–treatment (AL)** was that the profit increase from downsizing was rather small.

In the **unannounced downsizing, high incentive–treatment (UH)** participants were not told ex ante that downsizing would be possible after two rounds, i.e., they played the first two rounds unaware of the subsequent downsizing opportunity.[14] Again, the profit increase was relatively large.[15]

[12] The distances between the boundaries increase with increasing d and vanish for $d \to k$.

[13] This means we set $m = 1$ and $n = 1$.

[14] By *unaware* we mean that we did not give any hint. Of course, we could not effectively avoid that – at least some – participants might have expected future layoffs anyhow.

[15] We collected fewer data for this treatment to get a first impression whether results fundamentally differed from treatment AH. In particular, the first phase of this experiment

After the first phase, participants in treatments AH and AL were told that one repetition of all four rounds of their corresponding treatment would be played (in a perfect stranger design). We denoted the first phase of treatment AH with AH(1stphase), the second phase with AH(2ndphase), and so on. Since participants in the UH–treatment would anyhow anticipate the downsizing opportunity in the second phase, we also announced the downsizing option to them in UH(2ndphase).[16]

■ **Treatment parameters:** We constantly set $d = 12$, $k = 2$, $a = 4$, $U = 15$, $p = 6$, in experimental currency units (ECU).

In treatments **AH and UH**, we furthermore set $M = 24$. For the sake of readability, we denote the timing (without (W) or after downsizing (A)) with an index on the lower right (after the index of the player) and abstain from always indicating the treatment.

The optimal contracts without downsizing then consist of $F_W = 24$ and $r_W = 12$ in both treatments. Optimal effort levels are $e_h^* = 6$ for the highly productive agent, h, and $e_l^* = 1$ for the less productive one, l. Payoffs are $\omega_{l,W} = 30$ for the less productive worker, $\omega_{h,W} = 60$ for the highly productive one, and $\Pi_W = 36$ for the principal.

The principal should downsize and offer the same contract to the remaining agent. Payoffs are then $\omega_{l,A} = 15$, $\omega_{h,A} = 60$, and $\Pi_A = 48$. Thus, principal P can increase the profit by about 33.3 %. Note that the principal's absolute gain is smaller than the less productive agent's loss, and welfare thus decreases even when assuming the unemployment benefit to be, miraculously, cost neutral.[17]

In treatments **AL** we set $M = 16$. The optimal contract without downsizing consists of $F_W = 16$ and, again, $r_W = 12$. Optimal efforts are unaffected. Payoffs are lower for the workers ($\omega_{l,W} = 22$ and $\omega_{h,W} = 52$) and higher for the principal ($\Pi_W = 52$) due to the lower minimum wage.

Nevertheless, the principal should downsize and offer the same contract to the highly productive agent, who invests the same effort as before. Payoffs

was used as a separate treatment while data of the repetition was pooled with treatment AH; see section 4.1 for details. Since our predictions can be tested by these three treatments, we neglected a fourth (one) with unannounced downsizing and low incentives.

[16] The consequences are discussed in section 4.1.

[17] This is another aspect where our study also fundamentally differs from Fischer et al. (2008), where social welfare is maximized with downsizing.

are then $\omega_{l,A} = 15$, $\omega_{h,A} = 52$, and $\Pi_A = 56$. This means that principal P can increase the profit by about 7.7 %. Again, the agent's loss more than outweighs the principal's profit increase so that conservatively estimated welfare decreases.

We refer to these results as *benchmark predictions* from now on.

■ **Software, framing, miscellaneous:** All participants received a fixed fee of $FF = 90$ ECU in each of the two phases of the experiment. We split this amount into two parts, FF_1 and FF_2 (45 ECU for rounds 1 and 2, 45 ECU for rounds 3 and 4), in the first phase of the UH–treatment. The experiment was programmed and conducted with z-Tree (Fischbacher (2007)).

Contract offers and effort choices were restricted to reasonable intervals.[18] Furthermore, employees were prohibited from choosing effort levels that would result in negative round earnings for the given contract. Principals were asked to make conjectures about the effort choices of the agents. Their contract offer in combination with those conjectures about effort choices was not allowed to imply negative expected payoffs. Other experimental studies also use techniques to avoid financial suicide of participants (see, e.g., Falk et al. (2008)). These restrictions could be checked by participants with a calculator integrated into the experimental software. Each participant could use this device to calculate all resulting payoffs from any combination of F, r, e_i, and e_j.[19]

Nevertheless earnings for principals could be negative due to overestimated efforts. The losses, if occurring, were subtracted from the other rounds' earnings and the fixed fee. Unlike other experiments that simply excluded participants whose aggregate payoff approached zero (see, e.g., Fehr et al. (2007)), we informed subjects that aggregate losses had to be paid out of pocket or by administrative work. We stressed that this was very unlikely to occur. In fact, moderate overall losses did occur in only one case.

The instructions were cautiously framed referring to *"employer"*, *"employee"*, *"fixed wage"*, *"piece rate"*, and *"layoffs"*. Although framed instructions could strengthen imported views, we explicitly wanted to analyze a

[18] Principals were restricted to fixed wage offers F with $24 \leq F \leq 40$ in treatments AH and UH, and $16 \leq F \leq 40$ in AL. For piece-rate offers we demanded $0 \leq r \leq 20$, for efforts $0 \leq e \leq 10$. One decimal point was allowed.

[19] We restricted the use to two minutes per round, after which participants had to decide *"at the latest."*

phenomenon connected to labor markets. Other labor market experiments use similar frames; compare, e.g., Fehr et al. (2007).[20] Representative instructions are given in Appendix A.

Two sessions of each treatment were played. This means we had 48 participants per treatment and 144 participants altogether. They were recruited, using the software ORSEE (Greiner(2004)). All sessions were conducted at the computer laboratory of the Max Planck Institute, Jena, in July 2008. Participants were students; the exchange ratio was 30 ECU = 1 euro. Average earnings were 14.55 euro (standard deviation: 3.99 euro), experimental sessions lasted about 100 minutes.

■ **Hypotheses:** Although the game theoretic benchmarks always predict downsizing, we expect layoffs to occur less often when the principal's gain from downsizing is rather small. Many experiments, ranging from ultimatum to dictator and public good games have shown that other-regarding preferences play a role in human behavior.[21] Similarly, some P–participants in our downsizing experiment might face a trade-off between own and other participants' payoffs. Since material gains from downsizing are larger in the AH–treatment, we expect downsizing in this case to occur more often, but not always. The finding of Charness/Rabin (2002) that efficiency concerns might also be important only strengthens this view.

Hypothesis 1: *There is less downsizing in the AL-treatment than in the AH-treatment. Downsizing occurs in AH as often as in UH. There will be some cases without downsizing in all treatments.*

For contract offers motives like inequity aversion, altruism, or fairness (see, e.g., Bolton/Ockenfels (2000) or Rabin (1993)) might induce at least some P–participants to offer better terms than predicted. P–participants may try to inspire higher, efficiency enhancing[22] effort levels by increasing fixed

[20] The authors also found almost no differences when comparing their results to an ostensibly neutral control treatment, using terms such as *"buyer," "seller,"* etc., instead.

[21] See Davis/Holt (1993) or Kagel/Roth (1995) for comprehensive discussions of dictator and public good experiments and, e.g., Andreoni/Miller (2002), Charness/Rabin (2002), or Fehr/Fischbacher (2003) on human altruism and social preferences. Charness/Rabin (2002, p. 817) support our perception of experimental results when summarizing: *"Participants in experiments frequently choose actions that do not maximize their own monetary payoffs when those actions affect others' payoffs."*

[22] It is easy to check that doubling the effort is welfare maximizing.

wages and/or the piece-rate offers.[23] Since low piece-rate offers harm both groups – principal and agents – we do not expect them to be substantially smaller than predicted.

Hypothesis 2: *Piece-rate offers are never below their benchmark level, but exceed the benchmark level for some P–participants. Fixed wage offers are higher than the minimum fixed wage.*

Welfare, fairness, or reciprocity concerns could increase effort levels. In particular, contract offers significantly above the optimum could be perceived as a kind action of the principal and trigger reciprocity, i.e., effort levels above the individual benchmark predictions (Brandt/Charness (2004), Dufwenberg/Kirchsteiger (2004)).

Hypothesis 3.1: *Effort levels of L– and H–participants without downsizing are either near to, or above, their benchmark level.*

Announcing the downsizing opportunity early could affect effort levels of less productive workers. In the HA-treatment, L–participants knowing that they can be fired might exert more effort before this downsizing decision than in HU-treatment and reduce their effort afterwards. This is called a probation period effect in the literature.[24]

Hypothesis 3.2: *Effort levels follow the pattern of a probation period effect for less productive workers anticipating downsizing. Round 1 and 2 effort levels of L–participants in the UH-treatment are lower than in the AH-treatment. There is no difference between effort levels of still employed L–participants after downsizing in all treatments.*

Finally, we want to explore how the remaining agents react to witnessing layoffs. They might perceive the firing of other agents as an unkind action of their principal and react reciprocally – here by lowering their effort level.

Hypothesis 3.3: *Effort levels of some H–participants decrease after witnessing the firing of L–participants.*

[23] In a broader sense, our experiment could be perceived as a trust game (see Cox (2004) or Bleich/Kirstein (2006)).

[24] The first two rounds of each phase of our experimental setting are similar to such a probation period. Ichino/Riphan (2005) in a field study (using absenteeism as a proxy for effort) and Falk et al. (2008) in an experiment find evidence that effort levels during probation, when dismissals are still possible, are higher than afterwards.

Overall, these considerations should, on average, lead to payoffs slightly above the benchmark predictions.

Hypothesis 4: *Payoffs and welfare slightly exceed their benchmark levels.*

4 Experimental results and discussion

After some methodological statements (4.1) we discuss the downsizing decisions (4.2). We then present the results of treatments AH and AL in two separate subsections (4.3 and 4.4) before comparing them (4.5). Since for treatment UH results are quite similar to treatment AH, we summarize all results and comparisons in a final subsection 4.6.

4.1 Methodology

All tests performed are two-sided with $\alpha = .05$ and using SPSS 16.[25] In each session of a treatment 8 triplets played the game twice, i.e., each session supplied us with 16 triplets of averaged data without and after downsizing. For treatment AH, e.g., we used the averages of the two rounds without downsizing as independent observations (observations: AH(1_W), AH(2_W), ..., AH(32_W)). The averages of the two rounds after downsizing were used as independent observations (observations: AH(1_A), AH(2_A), ..., AH(32_A)) that are dependent on observations AH(1_W) - AH(32_W). Since we used a perfect stranger design, we pooled the data of the first and second phases of treatments AH and AL when testing treatment effects. Specifically, we pooled AH(1stphase) with AH (2ndphase) and AL(1stphase) with AL2(2ndphase), respectively.[26] We denoted the pooled data sets with AH(32) and AL(32).

In treatment UH we could not assume participants to be unaware of the downsizing opportunity in the repetition. UH(1stphase) thus remained as a single treatment with 16 observations without and after downsizing

[25] This software adjusts for ties. Exact p-values were calculated whenever possible. We rounded cell entries after the tenth decimal point, if necessary.

[26] Checking this procedure for the most important variables, we found only one difference to be slightly significant. All performed tests were Wilcoxon independent two-sample tests with 16 observations each. Less productive workers' efforts after downsizing are omitted. AH(1stphase) vs. AH(2ndphase): F_W ($p = .288$), r_W ($p = .716$), $e_{h,W}$ ($p = .689$), $e_{l,W}$ ($p = .165$), F_A ($p = .320$), r_A ($p = .582$), $e_{h,A}$ ($p = .496$). AL(1stphase) vs. AL(2ndphase): F_W ($p = .139$), r_W ($p = .731$), $e_{h,W}$ ($p = .326$), $e_{l,W}$ ($p = .043$), F_A ($p = .503$), r_A ($p = .801$), $e_{h,A}$ ($p = .595$).

(named UH(16)) and served as a small-scale check of differences to treatment AH. The question is how to deal with UH(2ndphase). The perfect stranger design, the identical procedure, and essentially the same instructions[27] as in AH(2ndphase) suggest to pool the data of UH(2ndphase) with AH(32). We checked this and found it confirmed by the data without downsizing,[28] and only violated for the piece-rate offer after downsizing.[29] This small lack of congruency is probably due to the small number of observations. We decided to pool the data without and after downsizing (denoted with AH(48)) and just to mention the few minor experience effects.

The last issue concerns the statistical tests. Non-parametric testing, especially the Wilcoxon test, has become predominant in experimental economics for a variety of reasons (small samples, non-normal distributions, ordinal data). Nevertheless, using the Wilcoxon test for detecting differences in central tendency between two small samples is debatable when dealing with unequal variances or differently shaped distributions. In these cases robust rank-order tests seem more appropriate (see, e.g., Feltovich (2003, 2005), Fligner/Policello (1981), Ruxton (2006), or Siegel/Castellan (1988)).

We used standard non-parametric tests for small samples ($n < 10$), for all medium samples ($n \leq 30$) violating normality, and for ordinal or nominal data, but calculated robust rank-order tests in addition to Wilcoxon tests otherwise.[30] As expected, p-values of both tests were always qualitatively similar.

For large ($n > 30$) or normally distributed samples, we always used parametric tests. When comparing population means, we reverted to the Welch-Satterthwaite independent two-sample t test (hereafter "WS test") without prior variance checks.[31]

[27] They differed only in a few words in one line. For participants in AH they stated: *"again, you will receive your **participation fee**"* FF. For participants in UH they stated: *"again, you will receive both your **participation fees** (45 ECU each)"* $FF_1 + FF_2 = FF$.

[28] Wilcoxon independent two-sample tests, 32 and 16 observations, respectively. AH(32) vs. UH(2ndphase): F_W ($p = .414$), r_W ($p = .265$), $e_{h,W}$ ($p = .519$), $e_{l,W}$ ($p = .788$).

[29] F_A ($p = .516$), r_A ($p = .028$), $e_{h,W}$ ($p = .056$).

[30] We checked normality and equality of distributions by eyeballing histograms, stem-leaf plots, and QQ-plots and using Kolmogorov-Smirnov and Lilliefors tests as a double check. We used Levene's test to check for unequal variances; there is no study we are aware of that criticizes the use of an unequal variance test prior to a robust rank-order test as is the case for the t test (see next par.). We interpolated the p-values delivered by Feltovich (2005) and used the normal distribution as approximation otherwise.

[31] Several studies have shown this to be superior to a two-way procedure (see, e.g., Moser

4.2 Downsizing decisions

Table 1 summarizes the absolute and relative frequency of firms that chose to lay off the less productive worker after the first two rounds. We observe cases without downsizing in all three treatments: in the AH–treatment 35 out of 48 firms (72.9 %) fired their less productive workers after two rounds. This significantly differs from the theoretical prediction that all firms downsize (χ^2-test on abs. freq., $p < .001$).

	Downsizing		No downsizing	
Treatment	Abs. freq.	Rel. freq	Abs. freq.	Rel. freq
AH(48)	35	72.9 %	13	27.1 %
AL	25	78.1 %	7	21.9 %
UH	13	81.2 %	3	18.8 %

Table 1: Downsizing decisions

In the AL–treatment 25 of the 32 firms (78.1 %) downsized while 7 firms did not. Again, this significantly differs from the theoretical prediction of full downsizing (χ^2-test on abs. freq., $p < .001$).

Surprisingly, the percentage of firms that downsize was lower in the AH–treatment where theoretical gains from downsizing are rather large (33.3 %) compared to the AL–treatment where they are rather small (7.7 %). Factual gains are also higher in AH than in AL (see subsection 4.5 below). The differences in downsizing are insignificant, using a χ^2-test with Yates correction or a Fisher-Exact test for the resulting 2x2-table (p = .792 and p = .793, respectively). This refutes our hypothesis that the size of theoretical gains are relevant for downsizing.

Finally, the percentage of firing was highest in UH–treatment with 81.2 % (13 out of 16), which still significantly differs from full downsizing (χ^2-test on abs. freq, $p < .001$). The more frequent downsizing in UH when compared to treatment AH – with equal incentives, but anticipated downsizing - is insignificant (χ^2-test with Yates correction: p = .739; Fisher-Exact test: p = .740).[32] Thus, we summarize:

et al. (1989), Neuhäuser (2002), Ruxton (2006), or Zimmerman (2004)).

[32]All results stay qualitatively the same when using AH(32) instead of AH(48). In AH(32), e.g., 71.9% of the firms chose to downsize instead of 72.9 % in AH(48).

Result 1: *There is insignificantly more downsizing in AL than in AH and in UH than in AH. The number of cases without downsizing is significantly greater than zero in all three treatments.*

4.3 Treatment AH

Table 2 below describes the decisions and resulting payoffs for all participants in AH(48) without and after the downsizing decision.

The table distinguishes between a) all 48 observations (abbreviated and indexed "all," if inevitable), b) firms that did (firms D, 35 obs.) and c) firms that did not downsize their labor force (firms ND, 13 obs.). In cases without observations, e.g., for the effort e_l of a fired worker, table cells are left empty. We will not index the treatment when it is obvious.

	a) All			b) Only firms D			c) Only firms ND		
	Without	After		Without	After		Without	After	
F	25.22 (.000**)	25.64 (.000**)	.238	25.37 (.000**)	25.79 (.001**)	.274	24.81 (.250)	25.23 (.125)	1.000
r	8.42 (.000**)	8.06 (.000**)	.326	7.73 (.000**)	7.46 (.001**)	.516	10.27 (.008**)	9.67 (.010*)	.450
e_h	4.40	4.17	.296	4.06	3.79	.273	5.32	5.19	.790
e_l	.92			.81			1.23	1.03	.043*
Π	23.24 (.001**)	30.09 (.000**)	.011*	17.72 (.000**)	28.90 (.001**)	.000**	38.12 (.305)	33.31 (1.000)	.635
ω_h	46.04 (.000**)	45.45 (.000**)	.702	44.28 (.000**)	43.84 (.001**)	.810	50.79 (.005**)	49.77 (.003**)	.970
ω_l	27.81 (.000**)	18.65 (.000**)	.000**	28.04 (.002**)	15.00	.000**	27.19 (.083)	28.47 (.229)	.685

Table 2: Main results, treatment AH

In each cell, averages without downsizing are given first, followed by averages after downsizing and the p-value for the dependent two-sample test. Numbers in brackets below the averages are the p-values for one-sample tests against the benchmark predictions. For example, the average offered wage of all firms was 25.22 without and 25.64 after downsizing. Both averages differ significantly from the benchmark of $F = 24$ ($p = .000$ both times), but not from each other ($p = .238$). A complete list of the used tests shown in Table 2 can be found in Appendix B. Asterisks denote significance on the

5% (*) and 1%level (**). In case c) where firms did not downsize, we test for optimality from then on.

As expected, fixed wage offers are slightly, but significantly larger than the minimum wage of 24 without and after downsizing, at least for groups a) and b). Firms D always offered higher fixed wages than firms ND,[33] although this difference is insignificant (without: Wilcoxon independent two-sample test (hereafter "Wilcoxon test"): p =.183, after: p=.517, same test).

Piece-rate offers were much smaller than their benchmark level $r = 12$ (p-values between p = .000 and p = .010), although calculators were in heavy use.[34] The average piece rate offered over all firms without downsizing was only $r_W = 8.42$ and remained at about the same low level, $r_A = 8.06$, (p = .326). When distinguishing between firms D and ND, we see that the former, on average, offered lower piece rates without downsizing, $r_{W,D} = 7.73$, than the others, $r_{W,ND} = 10.27$. After downsizing, firms D offered $r_{A,D} = 7.46$, firms ND $r_{A,ND} = 9.67$. Due to the much larger variance among firms D, these differences are slightly insignificant (without downsizing: robust rank-order tests: p = .060 (Wilcoxon test: p = .087), after: p = .052 (.063).[35]

Result 2.AH: *Offered fixed wages are slightly above the benchmark predictions for most participants. Piece-rate offers are much below optimum throughout. Firms ND offer higher piece rates than firms D.*

Due to the low piece rates, benchmark tests for efforts are not illuminative at all. Instead, we compute measures of absolute and relative deviation given in Table 3. The absolute deviation of subject h's effort, $absdev_h$, is simply the difference between the chosen effort level and the effort level that would be optimal given the offered contract,[36] i.e., 0 is the benchmark we tested against. One might argue that absolute deviations are of little importance since the relative effects of an absolute deviation z can be quite different in payoff space. For an offered piece rate $r = 15$, e.g., a deviation of $z = 0.15$

[33] All comparisons between those groups are not given in tables, but in the text.

[34] For example, principal participants in this treatment used the calculator for an average of 95 seconds in each of the 4+4=8 rounds they played.

[35] We will give the results of Wilcoxon test in analog situations in brackets. Standard deviations are 4.38 for firms D and 1.98 for firms ND. The only minor difference when using AH(32) is that the difference in piece rates after downsizing – $r_{A,D} = 6.34$, $r_{A,ND} = 9.31$ – is slightly significant (robust rank-order test: p = .038 (.041)).

[36] Due to the compulsory discrete choice set of employee participants, the optimal choice may not be selectable. In these cases, we redefined the optimum effort level as "the optimum among selectable effort levels."

decreases a less productive worker's payoff by about 1.4% and increases the principal's profit by about 12%. If $r = 6$, however, the same deviation is costlier for the worker (9%) and more beneficial for the employer (30%). It can be generally shown that the quotient between z and the optimum effort, hereafter called $reldev$, always behaves in this way.[37] Of course, the benchmark is 0 again.

	a) All		b) Only firms D		c) Only firms ND	
	Without	After	Without	After	Without	After
$absdev_h$	.19 (.170)	.14 (.136)	.19 (.193)	.06 (.410)	.18 (.591)	.35 (.229)
$absdev_l$	.22 (.000**)		.16 (.001**)		.38 (.000**)	.23 (.004**)
$reldev_h$	.11 (.127)	.04 (.135)	.13 (.198)	.02 (.446)	.08 (.383)	.08 (.194)
$reldev_l$	.62 (.009**)		.62 (.051)		.64 (.000**)	.45 (.004*)

Table 3: Effort deviations, treatment AH

Summarizing the results of Tables 2 and 3, we see that highly productive workers behave optimally (deviations insignificantly differ from 0, ranging from .06 to .35 for absolute deviations and .02 to .13 for relative deviations). Since piece-rate offers without downsizing are higher among firms ND, effort levels of highly productive workers, $e_{W,ND} = 5.32$, are above efforts in firms D, $e_{W,D} = 4.06$. This difference is significant (WS test: p = .007). The same is true after downsizing ($e_{A,ND} = 5.19$, $e_{A,D} = 3.79$; robust rank-order test: p = .030 (.023)). We do not observe a sharp decrease in effort levels of highly productive workers who witness layoffs. In fact, the moderate decrease of $e_{W,D} = 4.06$ to $e_{A,D} = 3.79$ is insignificant (p = .273) and accompanied by lower piece-rate offers. The decreases in effort levels for the whole data set and firms ND are also insignificant (p = .296 and p = .790, respectively).

[37]The quotient of the principal's gross profit$((a - r)z)$ and her optimal gross profit $((a-r)r/k)$ in fact equals $reldev_h$. This means the larger the piece-rate offer, the smaller the relative profit increase. The quotient of absolute worker's payoff decrease $(kz^2/2)$ and her optimum payoff $(r^2/(2d))$ equals $(reldev_h)^2$. We thereby accept the few undefined cases where the piece-rate offer was $r = 0$. When only one of the two values that were used to compute averages was missing, we used the other one only. Then only 2 or 3 of 48 cases were still missing (see Appendix B for details).

The picture changes for less productive agents who tend to spend more effort than is optimal. Absolute deviations are between .16 and .38, being significantly different from 0 (see Table 3). Relative deviations are even higher and lie between $reldev = .45$ and $reldev = .64$, all except one being significantly different from 0 (p between .000 and .009 and p = .051 for $reldev_{W,D}$).

The difference between the relative deviation of L–participants without downsizing, $reldev_{l,W,all} = .62$, and that of H–participants, $reldev_{h,W,all} = .11$, is significant (WS test: p = .039). Absolute deviations, $absdev_{l,W,all} = .22$ and $absdev_{h,W,all} = .19$, are not different (same test, p = .835).

We also find a hint of a probation period effect among less productive employees: Although contract offers do not change much after downsizing among firms ND, the average effort level of workers without downsizing, $e_{W,ND} = 1.23$, is significantly higher than the average after downsizing, $e_{A,ND} = 1.03$ (dependent sample Wilcoxon test, p = .043).[38] The difference between the efforts of workers employed by firms D, $e_{W,D} = .81$, and firms ND, $e_{W,ND} = 1.23$ (Wilcoxon test: p = .003) can partly be explained by the difference in piece-rate offers. But the difference in the absolute deviation between firms D and ND, $absdev_{W,D} = .16$ and $absdev_{W,ND} = .38$ (significant; robust rank-order test: p = .018 (.022)), and in the relative deviation, $reldev_{W,D} = .62$ and $reldev_{W,ND} = .64$ (insignificant; Wilcoxon test: p = .063), weakly indicates that workers that tend to be "more generous" are those working in firms that do not downsize.

Result 3.AH: *Highly productive workers behave opportunistically. There is no evidence that they react to witnessing layoffs. Less productive workers tend to spend more effort than is optimal, especially during probation periods.*

The P–participants' average payoffs differ from their benchmark (see Table 2 for p-values, all significant), what is not surprising considering the low piece rates. There is a significant increase of about 30% in the principals' average payoff from $\Pi_{W,all} = 23.24$ without downsizing to $\Pi_{A,all} = 30.09$ (p = .011). The same pattern holds when limiting the analysis to firms D where the increase from $\Pi_{W,D} = 17.72$ to $\Pi_{A,D} = 28.90$ is even larger with 63% (p

[38] This result remains when restricting to AH(32) (p = .047). Deviations for highly productive workers become moderately higher and slightly significant for the whole data set, but not for subgroups. Although less productive workers' relative deviation is again higher than that of highly productive ones (.83 and .20), this difference is insignificant in HA(32) (p = .085). Data are available on request.

= .000). Firms that do not downsize suffer a payoff loss from $\Pi_{W,ND} = 38.12$ to $\Pi_{A,ND} = 33.31$, partly because of the probation period effect. This difference is not statistically significant (p = .635), nor is the deviation from the benchmark payoff of 36.

When comparing payoffs between different groups of firms, it is striking that firms D earned only $\Pi_{W,D} = 17.72$ on average without downsizing while firms ND earned more than twice as much: $\Pi_{W,ND} = 38.12$. This difference is, of course, significant (robust rank-order test, p = .002 (.001)) and suggests that payoffs are the dominant indicator of the downsizing decision (see the logistic regression below). The difference in payoffs after downsizing between firms D, $\Pi_{A,D} = 28.90$, and firms ND , $\Pi_{A,ND} = 33.31$, is much smaller (Wilcoxon test: p = .684).

Due to the low piece rates, payoffs of H–participants are smaller than their benchmark of 60 for all groups and invariant to the firing decision (see Table 2 for p-values). Highly productive workers earn higher payoffs in firms ND (about 50 ECU) than in firms D (about 44 ECU), but the differences are insignificant (without: WS test: p = .101; after: robust rank-order test: p = .141 (.168)).

Less productive workers, in total, suffer an income loss after downsizing: their average earnings are $\omega_W = 27.81$ without and $\omega_A = 18.65$ after downsizing (p = .000). This decrease is caused by the firing decisions of firms D where average earnings decrease from $\omega_{W,D} = 28.04$ without downsizing to the unemployment benefit of 15 afterwards (p = .000). Among firms ND, profits do not change much ($\omega_{A,ND} = 28.47$ and $\omega_{W,ND} = 27.19$, p = .685). Payoffs are near to their benchmark levels because fixed wages play a dominant role for less productive workers (see Table 2 for p-values).

With regard to welfare, WF, naively set equal to the sum of payoffs, we restrict ourselves to a few observations about the whole group of firms. Here, welfare is slightly higher without downsizing, $WF_W = 97.10$, than afterwards, $WF_A = 94.18$ (insgnificant, dependent two-sample t test: p = .411), but both are much lower than the benchmark of 126 and 123, respectively (independent one-sample t tests: p = .000 in both cases). The payoff increase of firms amounts to 6.85 on average, the less productive workers' payoffs decrease with reversed sign to 9.16. The difference between them is not significant (WS test: p = .414).[39]

[39] Again, all results remain qualitatively the same for AH(32), only the principals' average profit increase for all firms (from 21.32 to 26.79) becomes insignificant (p = .128).

Result 4.AH: *For all firms average payoffs and welfare are lower than predicted. Downsizing firms earn more after downsizing than before. Firms ND earn much more without downsizing than firms D. Highly productive workers' payoffs are about equal in all groups. Less productive workers suffer an income loss when being fired.*

A few remarks on the – cheap talk – effort level conjectures demanded from principals. Denoting the quotients of conjectured effort and optimal effort for both employees as $relconj_h$ and $relconj_l$, respectively, a value of 1 indicates rational expectations while a value smaller 1 suggests that principals overestimate efforts. Restricting ourselves to all firms, we observe that all averages are significantly below 1: $relconj_{h,W} = .87$, $relconj_{h,A} = .83$, and $relconj_{l,W} = .57$ (one-sample t tests: p = .033, p = .001 and p = .000, respectively). The difference between $relconj_{h,W} = .87$ and $relconj_{l,W} = .57$ is statistically significant (WS test: p = .000).[40] We conclude that P–participants tend to overestimate both effort levels, especially those of less productive workers.

We finally want to test whether a principal's downsizing is dominantly influenced by her own profit. Therefore, we perform a simple logistic regression, using a forward algorithm to check which of all 17 explanatory variables we observed without downsizing[41] influence the binary dependent dummy variable *Down* that takes the value 1 if a firm downsizes and 0 otherwise. This algorithm uses score tests to decide which variable to include next as well as likelihood ratio as a criterion for exclusion. Formally, our model with j variables can be described by

$$Z_i = \ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \sum_j \beta_j x_j + \varepsilon_i$$

with Z_i as the latent variable and p_i determined by the logistic function.

A selection of relevant tables is given in Appendix D. The algorithm stops after step 1, including only the firm's payoff Π as explanatory variable.[42] Table 13 gives the results of the LR test for this first step compared to the null model without explanatory variables. The inclusion of Π contributes

[40] Results are qualitatively the same for AH(32).

[41] That is, the already mentioned 13 explanatory variables as well as both conjectured efforts, $conj_h$ and $conj_l$, and absolute deviations of conjectures from optimum effort, $absconj_h$ and $absconj_l$.

[42] We thus need not be concerned about multicollinearity.

significantly (p = .002). The parameter estimates are $\beta_0 = 3.245$ for the constant and $\beta_1 = -.075$ for the coefficient of Π. The effect of the principal's profit is significant (Wald test: p = .017, see Table 10).

This means, for example, that the probability of firing the less productive employee is about 92,4% when the principal's profit is $\Pi = 10$, about 63.1% at the theoretical profit without downsizing, $\Pi = 36$, and only about 40.9% when his profit reaches $\Pi = 48$. The model predicts 78.3 % of observations correctly, which is an improvement over the null model that classifies 71.7 % correctly (see Table 11). Nagelkerke's R^2 for our regression is $R^2 = .281$.[43]

Result 5.AH: *Firms tend to overestimate effort levels, especially those of less productive workers. Firing decisions are strongly influenced by principals' profits.*

4.4 Treatment AL

In treatment AL we have 32 observations in total, 25 for firms D and 7 for firms ND. Table 4 gives an overview of decisions and payoffs.[44]

	a) Aall			b) Only firms D			c) Only firms ND		
	Without	After		Without	After		Without	After	
F	18.47 (.000**)	17.45 (.002**)	.078	18.50 (.000**)	17.57 (.001**)	.150	18.36 (.500)	17.00 (.500)	.500
r	8.96 (.000**)	9.36 (.001**)	.285	8.83 (.001**)	9.61 (.014*)	.101	9.43 (.016*)	8.46 (.016*)	.250
e_h	4.26	4.34	.762	4.13	4.40	.220	4.74	4.14	.031*
e_l	.87			.88			.84	.80	.563
Π	32.82 (.000**)	36.56 (.000**)	.256	30.20 (.000**)	36.38 (.000**)	.039*	42.16 (.625)	37.21 (.313)	.156
ω_h	40.03 (.000**)	42.35 (.003**)	.211	39.62 (.001**)	43.83 (.037*)	.103	41.52 (.016*)	37.07 (.016*)	.313
ω_l	21.66 (.631)	16.05 (.008**)	.000**	21.52 (.563)	15.00	.000**	22.20 (.891)	19.81 (.016*)	.156

Table 4: Main results, treatment AL

Fixed wage offers are again higher than their benchmark $F = 16$, ranging

[43] Results are qualitatively the same for AH(32).

[44] It is constructed analogously to Table 2, performed tests are given in Appendix B.

from 17 to 18.5. These differences are significant for all groups except firms ND (see Table 4 for p-values). Fixed wage offers decline by about 1 ECU after the downsizing decisions in all groups, but these changes are insignificant. There are no noticeable differences between fixed wage offers without downsizing by firms D and (firms) ND (Wilcoxon test: p = .502) and the offers after downsizing by both groups (Wilcoxon test: p = .627).

Piece-rate offers are below the optimum of $r = 12$ without and after downsizing in all groups (p-values from .000 to .016). They are not significantly different in any group without and after downsizing (p = .285 for all, p=.101 for firms D, p=.250 for firms ND); but while firms D somewhat balance their declining fixed wage offers by an increase in the piece rate from $r_{W,D} = 8.83$ to $r_{A,D} = 9.61$, firms ND also lower their piece-rate offers from $r_{W,D} = 9.43$ to $r_{A,D} = 8.46$. The differences between piece-rate offers of firms D and ND without and after downsizing are insignificant (without: robust rank-order test: p = .894 (.922); after, same tests: p = .553 (.599)).

Result 2.AL: *Offered fixed wages are above their benchmark for most participants. Piece-rate offers are much below optimum throughout. There are no significant differences between piece-rate offers by firms D and firms ND.*

Efforts of highly productive workers decreased from $e_{W,ND} = 4.74$ to $e_{A.ND} = 4.14$ among firms ND, what is statistically significant (p = .031). Again, we find no hint of H–participants reacting to layoffs of less productive workers. On the contrary, their effort is higher after downsizing, $e_{A,D} = 4.40$, than without, $e_{W,D} = 4.13$ – probably due to increasing piece-rate offers. Absolute and relative deviations, presented in Table 5, suggest that H–participants spend a little less effort than is optimal. Only the deviations for the whole group after downsizing are significantly different from zero.

For less productive workers effort levels are almost identical in all groups (ranging from .80 to .88) and do not reveal a probation period effect (the difference between efforts among firms ND without downsizing, $e_{W,ND} = .84$, and after downsizing, $e_{W,ND} = .80$, is not statistically significant (p = .563)). L–participants seem to spend more effort than is optimal. These deviations are not significant except for the absolute deviations concerning the whole group without downsizing (p = .016). Absolute deviations of H–participants, $absdev_{h,W} = -.22$, and L–participants, $absdev_{l,W} = .13$, of the whole group without downsizing differ significantly from each other (WS test: p = .049). The same holds for relative deviations ($reldev_{h,W} = -.07$, $reldev_{l,W} = .26$; same test, p = .025).

	a) All		b) Only firms D		c) Only firms ND	
	Without	After	Without	After	Without	After
$absdev_h$	−.22 (.184)	−.34 (.029*)	−.29 (.081)	−.41 (.057)	.03 (.750)	−.09 (.250)
$absdev_l$	.13 (.016*)		.14 (.069)		.08 (.125)	.11 (.375)
$reldev_h$	−.07 (.154)	−.08 (.024*)	−.09 (.076)	−.10 (.164)	.04 (.625)	−.03 (.500)
$reldev_l$	.26 (.058)		.31 (.102)		.10 (.125)	.11 (.375)

Table 5: Effort deviations, treatment AL

Result 3.AL: *Highly productive workers behave rather opportunistically, spending only a little less effort than is optimal. There is no evidence that they react to witnessing layoffs. Less productive workers tend to spend more effort than is optimal; there is no probation period effect.*

The average payoffs in treatment AL follow a pattern similar to the AH–treatment with few exceptions: for all firms and firms D, profits increase after downsizing, but only the change for group D is statistically significant (see Table 4 for p-values). Again, the payoff difference between firms D ($\Pi_{W,D} = 30.20$) and firms ND ($\Pi_{W,D} = 42.16$) without downsizing is larger than the difference after downsizing ($\Pi_{A,D} = 36.38$ vs. $\Pi_{A,D} = 37.21$), but both are not significant this time (Wilcoxon tests, without: p = .085; after: p = .746). When performing a logistic regression analogous to that of the preceding subsection, the algorithm stops at the null model, not including any explanatory variable at all (see Appendix D for details). Highly productive workers earn less than their benchmark and about the same without and after downsizing. L–participants suffer from being fired. Their earnings without downsizing hardly differ from the benchmark of 22 and between firms D and ND (Wilcoxon test: p = .956).

Welfare in all firms, defined as above, is almost equal without, $WF_W = 94.51$, and after downsizing, $WF_A = 94.97$ (dependent two-sample t test: p = .921), but much smaller than the unaltered benchmarks (p = .000, in both cases). Again, less productive workers' average losses (5.61 ECU) insignificantly differ from firms' gains (3.75 ECU) (WS test: p = .578).

Result 4.AL: *Average payoffs and welfare for all firms are lower than their benchmark levels. Downsizing firms earn more after layoffs. Firms ND earn*

more without downsizing than firms D. Highly productive workers' payoffs are about equal in all groups. Less productive workers suffer an income loss when being fired.

Relative effort conjectures are larger than 1 for highly productive agents this time: $relconj_{h,W,all} = 1.16$ (p-value for a one-sample t test against 1: p = .066) and $relconj_{l,W,all} = 1.27$ (p = .021)), respectively, indicating that firms underestimate H–participants' effort choices. But they overestimate the efforts of less productive workers: the relative effort conjecture is $relconj_{l,W,all} = .75$ (p = .000), what differs from their estimation accuracy for highly productive workers (WS test: p = .000).

Result 5.AL: *Firms tend to underestimate effort levels of highly productive workers and overestimate those of less productive workers.*

4.5 Treatments AH and AL

When comparing the results of treatments AH(48) and AL, we mainly concentrate on aggregate firm level and on findings that are strikingly different.[45]

Naturally, fixed wage offers are lower in treatment AL than in treatment AH without downsizing ($F_{AH,W} = 25.22$ and $F_{AL,W} = 18.47$, WS test: p = .000) and after downsizing ($F_{AH,A} = 25.64$ and $F_{AL,A} = 17.45$, WS test: p = .000). Piece-rate offers, by contrast, are quite similar without downsizing, $r_{AH,W} = 8.42$ and $r_{AH,W} = 8.96$ (WS test: p = .519), and differ more, though insignificantly, after downsizing where firms offer $r_{AH,A} = 8.06$ in treatment AH, but $r_{AL,A} = 9.36$ in treatment AL (WS test: p = .147). In both treatments, piece-rate offers are much lower than optimum.

As for efforts and deviations, it is interesting to distinguish between subgroups again: While there are no significant differences for firms D,[46] the effort levels of less productive workers in firms ND are $e_{AH,W,ND} = 1.23$ in treatment AH, but only $e_{AL,W,ND} = .84$ in AL (robust rank-order test: p = .010 (.015)), although piece-rate offers do not differ significantly. Before the downsizing decision, L–participants in treatment AH seem to be concerned about future layoffs and thus spend relatively more effort than those in treatment AL. This view is strengthened by the findings after downsizing. Here, effort levels and absolute and relative deviations are also larger in AH than

[45] We performed tests for all subgroups. Data are available on request.

[46] Robust rank-order test for e (p = .816 (.838)), Wilcoxon tests for *absdev* (p = .345) and *reldev* (p = .242).

in AL, but the differences become smaller and insignificant. For efforts, e.g., we then observe $e_{AH,A,ND} = 1.03$ and $e_{AL,A,ND} = .80$ (Wilcoxon test: p = .128).

The difference in the average gain of firms without and after downsizing in treatment AH, 6.85, is larger than in treatment AL, 3.75, though this difference is not significant (WS test: p = .457).[47] But even in combination, these insights about efforts and payoffs only partly explain why relatively more firms keep their less productive employee in AH than in AL.

Payoffs of workers and welfare are about the same in both treatments except for the rent reallocation due to the lower fixed wages.[48]

Result 6: *Treatments AH and AL differ in fixed wages and thereby in payoffs: while principals earn more in AL, agents earn less. Piece-rate offers in both treatments do not differ greatly, but the less productive workers who are not fired in treatment AH spend relatively more effort than those in treatment AL.*

4.6 Treatment(s) UH (and AH)

In UH, participants were not aware of the downsizing opportunity. Since the trends in UH are similar to those in AH, we only list the main results of UH and state important differences at the end. The following Tables, 6 and 7, are constructed analogously to the preceding tables. Note that we had only three firms ND in this treatment (13 firms D, 16 firms in total).

Fixed wage offers are above the optimum but do not really differ without and after downsizing. Piece-rate offers are always far below the optimum (see the Tables for concrete p-values). Effort levels are relatively the same across groups of firms and do not differ without and after downsizing. There is no indication of a probation period effect (less productive workers' effort is almost identical without, $e_{W,ND} = .67$, and after downsizing, $e_{A,ND} = .68$, among firms ND). Deviations are comparable to treatment AH, except those of less productive workers among firms ND (see Table 7).

Just as in treatment AH, payoffs of all groups are lower than predicted, although the difference is insignificant for less productive employees. The principals' average payoff increase over all firms from without to after down-

[47] This also holds for firms D only (Wilcoxon test: p = .298).

[48] The only noteworthy difference when using AH(32) instead of AH(48) is that after downsizing piece rates are now significantly larger in AL than in AH (WS test: p = .032).

	a) All			b) Only firms D			c) Only firms ND		
	Without	After		Without	After		Without	After	
F	26.52 (.002**)	25.71 (.004**)	.328	26.79 (.008**)	25.95 (.016*)	.328	25.33 (.500)	24.67 (.500)	1.000
r	7.98 (.001**)	8.93 (.001**)	.452	7.98 (.001**)	9.11 (.003**)	.503	8.00 (.500)	8.17 (.250*)	1.000
e_h	4.46	4.48	.551	4.37	4.47	.750	4.83	4.50	.750
e_l	.79			.82			.67	.68	1.000
Π	22.65 (.042*)	33.31 (.000**)	.056	23.14 (.064*)	35.02 (.000**)	.127	20.53 (1.000)	25.88 (1.000)	.750
ω_h	44.74 (.001**)	48.13 (.002**)	.847	44.28 (.001**)	49.24 (.012*)	.622	46.75 (.500)	43.33 (.250)	.750
ω_l	28.97 (.385)	17.34 (.250)	.000**	28.95 (.438)	15.00	.000**	29.05 (.750)	27.47 (.250)	.750

Table 6: Main results, treatment UH

sizing is large in magnitude and only slightly insignificant (p = .056); the same holds for firms D only (p = .127). H–participants earn almost the same in all groups. Less productive employees suffer a significant income loss when being fired and earn about the same without and after downsizing otherwise. The payoff differences between firms D and ND are insignificant.[49]

Welfare without downsizing, $WF_W = 96.36$, is about the same as afterwards, $WF_A = 98.78$ (dependent sample Wilcoxon test: p = .860). Both are much smaller than the benchmark (without: p = .016, after: p = .000, Wilcoxon one-sample rank sum tests).

Relative conjecture accuracy is similar to treatment AH: principals overestimate highly productive workers' efforts ($relconj_{h,W} = .76$ and $relconj_{h,A} = .86$; Wilcoxon one-sample tests against the benchmark: p = .003 and p =. 025) as well as the effort of L–participants ($relconj_{l,W} = .54$, p = .000, one-sample t test). P–participants' conjectures are better for highly productive workers since the difference between $relconj_{h,W} = .76$ and $relconj_{l,W} = .54$ is significant (Wilcoxon test: p = .047).

When comparing treatments AH und UH, we checked all variables with-

[49]The p-values without downsizing are p = .875 (.900) for Π, using a robust rank-order test, and p = .900 for ω_h, and p = .986 for ω_l with Wilcoxon tests. After downsizing we observe p =.687 for Π and p = .596 for ω_h, using Wilcoxon tests. We did not calculate a logistic regression because of the small sample size.

	a) All		b) Only firms D		c) Only firms ND	
	Without	After	Without	After	Without	After
$absdev_h$	.46 (.023*)	.01 (.617)	.38 (.188)	−.08 (.920)	.83 (.250)	.42 (.500)
$absdev_l$	.12 (.216)		.14 (.203)		.00 (1.000)	.02 (1.000)
$reldev_h$	.15 (.039*)	.00 (.556)	.08 (.313)	−.02 (.977)	.46 (.250)	.08 (.500)
$reldev_l$	.55 (.328)		.75 (.189)		−.29 (1.000)	−.19 (1.000)

Table 7: Effort deviations, treatment UH

out and after downsizing in all subgroups (a complete list of tests is given in Appendix C). The major difference is found where it could be expected, namely among firms ND without downsizing. There the rather large, but insignificant difference between effort levels of less productive workers in treatment AH, $e_{AH,W} = 1.23$, and in treatment UH, $e_{AH,W} = .67$ (robust rank-order test: p = .203 (.179)), is partly due to the lower piece-rate offers. This, however, cannot account for the large differences between absolute deviations ($absdev_{AH,W} = .38$ vs. $absdev_{UH,W} = .00$) and especially relative deviations ($reldev_{AH,W} = .64$ vs. $absdev_{UH,W} = -.29$), which are both significant, despite the small simple size of firms ND (Wilcoxon tests, p = .032, and robust rank-order test, p = .041 (.041), respectively). For comparison, among firms D the relative deviations do not differ much (Wilcoxon test: p = .790). Apparently, L–participants who anticipated the downsizing decision (treatment AH) were more generous than those that could not foresee being fired – at least among the firms that eventually did not fire them.[50]

Result 7: *The main trends in treatment UH are comparable to those of treatment AH: Low piece-rate offers determine the payoff structure; principals' payoffs increase after downsizing while highly productive employees earn as much as before, and less productive workers suffer an income loss. Treatments AH and UH are almost identical. Among firms ND, L–participants seem to be more generous in AH than in UH.*

[50]For the sake of completeness: without downsizing, relative deviations of highly productive workers were higher in treatment UH than in AH (Wilcoxon test: p = .039). Again, all these results stay qualitatively the same for AH(32).

5 Conclusion

It may appear rather bold to explore the economic and emotional effects of downsizing in the typical tradition of laboratory experiments, i.e., by employing student participants and offering rather small incentives, compared to the sometimes lifelong dramatic effects of downsizing. We do, of course, acknowledge that our lab study can at best provide a clue to what has to be expected in the field. However, relying on student participants seems to be less debatable, partly because the behavior of student participants is rather typical of their age bracket (see Güth et al. (2007)), partly because there has been some evidence that lab behavior may be a good indicator of actual behavior (see, e.g., Falk/Fehr (2003), Karlan (2005)).

Rather than repeating all the specific findings reported above, let us summarize our main conclusions:

- More often than not, game theory fails to predict behavioral choices; this also applies to our downsizing game: fixed wages are higher (what, in view of its border prediction, could be partly attributed to noise) and piece rates are remarkably lower than expected.

- Effort choices sometimes deviate from best replies to piece-rate offers. In particular, as in probation periods detected in empirical and experimental studies (Ichino/Riphan (2005), Falk et al. (2008)), less productive workers in treatment AH invest more effort when anticipating layoff decisions by their principal.

- Contrary to our intuition, the major driving force of downsizing is neither the theoretically predicted nor the factual profitability of such a measure (as captured by treatments AH and AL), but rather the actually earned profit before the downsizing decision, suggesting that poorly performing firms are less scrupulous.[51]

- If downsizing opportunities are not anticipated (as in treatment UH), there is no probation period effect in the sense of higher efforts by less productive workers. This might explain why sometimes firms threaten to lay off (larger shares of) their workforce and withdraw from this measure later on.

[51] More generally, the message could be that entrepreneurial ethics are a luxury to be enjoyed only by highly succesful firms.

- There is hardly any indication that those not threatened by layoffs care about the fate of their less fortunate fellow workers. We admit that this may be partly due to our specific scenario where it is exogenously determined who can (cannot) be fired. In the field, it may not always be that obvious who is going to lose his or her job. But it remains an important insight that there is hardly any intrinsic solidarity among fellow workers.

Altogether, we believe that our results show that experimentally exploring downsizing decisions and workers' reactions to such measures can be instructive and suggest qualitative effects which, we hope, can be confirmed by field research.

References

Andreoni, J., Harbaugh, W. and L. Vesterlund (2003) : The Carrot or the Stick: Rewards, Punishments, and Cooperation, in: *American Economic Review*, Vol. 93, No. 3, 893-902, 2003.

Andreoni, J. and J. Miller (2002) : Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism, in: *Econometrica*, Vol. 70, No. 2, 737 - 753, 2002.

Berninghaus, S. K., W. Güth, C. Hoppe, and C. Paul (2007a) : International Competition in Hiring Labour and Selling Output: A Theoretical and Experimental Analysis, in: Dynamik internationaler Märkte. Wirtschaftswissenschaftliches Seminar Ottobeuren (eds. W. Franz, H.-J. Ramser, and M. Stadler), Mohr Siebeck, 2007.

Berninghaus, S. K., W. Güth, C. Hoppe, and C. Paul (2007b) : Duopolistic Hiring and Sales Competition : A Theoretical and Experimental Analysis, Working Paper.

Bleich, S. and A. Kirstein (2006) : Screening of Endogenous Types: An Experiment, Working Paper, 2006.

Bolton, G.E. and A. Ockenfels (2000) : ERC: A Theory of Equity, Reciprocity, and Competition, in: *PNAS Proceedings of The National Academy of Sciences*, Vol. 100, No. 6, 3531-3535, 2000.

Brandt, J. and G. Charness (2004) : Do Labour Market Conditions Affect Gift Exchange? Some Experimental Evidence, in: *The Economic Journal*, July, 684 - 708, 2004.

Charness, G. and M. Rabin (2002) : Understanding Social Preferences with Simple Tests, in: *Quarterly Journal of Economics*, Vol. 117, No. 3, 817 - 869, 2002.

Cox, J.C. (2004) : How to identify trust and reciprocity, in: *Games and Economic Behavior*, Vol. 46, 260 - 281, 2004.

Davis, D.D. and C. A. Holt (1993) : Experimental Economics, Princeton University Press, 1993.

Dufwenberg, M. and G. Kirchsteiger (2004) : A theory of sequential reciprocity, in: *Games and Economic Behavior*, Vol. 47, 268 - 298, 2004.

Falk, A. and U. Fehr : Why labour market experiments?, in: *Labour Economics*, Vol. 10, 399 - 406, 2003.

Falk, A. and U. Fischbacher (2006) : A theory of reciprocity, in: *Games and Economic Behavior*, Vol. 54, 293 - 315, 2006.

Falk, A., Huffmann, D., and W. B. Macleod (2008) : Institutions and Contract Enforcement, IZA Discussion Paper, No. 3435.

FAZ (2008a) : BMW fährt Rekord ein - Aktie stottert hinterher, in: *Frankfurter Allgemeine Zeitung (German nationwide newspaper)*, January 30, 2008, p. 16.

FAZ (2008b) : Nokia: Bochums Ertragskraft wird falsch dargestellt, in: *Frankfurter Allgemeine Zeitung (German nationwide newspaper)*, January 31, 2008, p. 18.

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

Fehr, E., Klein, A. and K. M. Schmidt (2007) : Fairness and Contract Design, in: *Econometrica*, Vol. 75, No. 1, 121 - 154, 2007.

Feltovich, N. (2003) : Nonparametric Tests of Differences in Medians: Comparison of the Wilcoxon-Mann-Whitney and Robust Rank-Order Tests, in: *Experimental Economics*, 6, 273 - 297, 2003.

Feltovich, N. (2005) : Critical Values for the Robust Rank-Order Test, in: *Communciations in Statistics – Simulation and Computation*, 34, 525 - 547, 2005.

Fischbacher, U. (2007) : z-Tree: Zurich Toolbox for Ready-made Economic Experiments, in: *Experimental Economics*, 10(2), 171-178.

Fischer, S., Güth, W., and C. Köhler (2008) : Effects of Profitable Downsizing on Collective Bargaining. Discussion Paper Series of the Max Planck Institute of Economics ♯2008–011.

Fligner, M. A., and G. E. Policello II (1981) : Robust Rank Procedures for the Behrens-Fisher Problem, in: *Journal of the American Statistical Association*, Vol. 76, No. 373, 162 - 168, 1981.

Frankfurter Rundschau (2008) : BMW spart beim Stammpersonal, in: *Frankfurter Rundschau (German nationwide newspaper)*, February 28, 2008, p. 18.

Greiner, B. (2004) : An Online Recruitment System for Economic Experiments, in: Kremer, K. and V. Macho (eds.): Forschung und wissenschaftliches Rechnen 2003. GWDG Bericht 63,Göttingen : Ges. für Wiss. Datenverarbeitung, 79-93, 2004.

Güth, W., Schmidt, C. and M. Sutter (2007) : Bargaining outside the Lab – A Newspaper Experiment of a Three-Person Ultimatum Game: in: *The Economic Journal*, Vol. 117 (March), 449-469, 2007.

Handelsblatt (2008) : Henkel baut 3000 Stellen ab and BMW greift hart durch, in: *Handelsblatt (German nationwide business newspaper)*, February 28, 2008, pp. 14 and 17.

Ichino, A. and R. T. Riphan (2005) : The Effect of Employment Protection on Worker Effort – A Comparison of Absenteeism During and After Probation, in: *Journal of the European Economic Association*, 3 (1), 120 - 143, 2005.

Kagel, J. K. and A. E. Roth (1995) : Handbook of Experimental Economics, Princeton University Press, 1995.

Karlan, D. S. (2005) : Using Experimental Economics to Measure Social Capital and Predict Financial Decisions, in: *American Economic Review*, Vol. 95, No. 5, 1688-1699, 2005.

Moser, B. K., Stevens, G. R., and C. L. Watts (1989) : The Two-Sample t test versus Satterthwaite's approximate F test, in: *Communciations in Statistics – Theory and Methods*, 18(11), 3963 - 3975, 1989.

Neuhäuser, M. (2002) : Two-sample tests when variances are unequal, in: *Animal Behaviour*, 63, 823 - 825, 2002.

NYTimes (2008) : After a Downsizing, How to Motivate?, in: *The New York Times*, August 24, 2008, p. 8.

Rabin, M. (1993) : Incorporating Fairness into Game Theory and Economics, in: *American Economic Review*, Vol. 83, No. 5, 1281 - 1302, 1993.

Ruxton, G. D. (2006) : The unequal variance t-test is an underused alternative to Student's t-test and the Mann-Whitney U test, in: *Behavioral Ecology*, 17, 688 - 690, 2006.

Richter, R. and E. G. Furubotn (2003) : Neue Institutionenökonomik. Tübingen: Mohr Siebeck. 3rd Edition.

Siegel, S. and N. J. Castellan, jr. (1988) : Nonparametric Statistics for the Behavioral Sciences, McGrawHill, Inc., 1988.

Zimmerman, D. W. (2004) : A note on preliminary tests of equality of variances, in: *British Journal of Mathematical and Statistical Psychology*, 57, 173 - 181, 2004.

A Instructions

Given below are the instructions for the L-employee, treatment AH, partly reformatted to save space. All other instructions are available from the authors on request.

■ **Experiment 1:**

1. General instructions:

Please stop communicating with other participants from now on and turn off your mobile phone. Read the following instructions carefully. If you have a question, please raise your hand, and the supervisors will answer your question at your computer box. We will have to exclude you from the experiment and all payments if you violate these rules. The instructions are identical for all participants except for the subsequent role assignment. Your anonymity will be guaranteed. This means that no other participant is going to learn your identity during or after the experiment. To begin with, you are taking part in an experiment consisting of **4 periods**. After this you will be given new instructions for another experiment!

In the first experiment, three participants will interact. Two of them will take the roles of employees, one will take the role of an employer. One of the employees is of type H (**H-employee**), the other of type L (**L-employee**). There is only one type of **employer**. The role assignment is carried out randomly in the beginning of the experiment. Each participant keeps his role during the whole experiment. **You are an L-employee.**

The earnings of every participant depend on his or her own decisions and those of the other participants. Earnings are calculated in ECU (Experimental Currency Unit) during the experiment. At the end of both experiments, they will be converted into Euro at a fixed exchange rate. **This exchange rate is 30 ECU = 1 Euro.**

Additionally, participants receive a fixed participation fee that does not depend on decisions, but will be offset against payoffs if necessary. **This participation fee is: 90 ECU.**

2. Periods 1 and 2:

2.1 General rules:

Each triplet of participants, consisting of an employer, an H-employee, and an L-employee, interacts for 4 periods. Each of the two periods 1 and 2 is basically constructed as follows:

1. The employer offers one contract that applies to both employees. It consists of two components: a **fixed wage** W with $24 \leq W \leq 40$ and a **piece rate** r (with $0 \leq r \leq 20$) that must be paid for each unit of output. Furthermore, the employer has to make conjectures about the employees' effort levels (see 2.). Up to 1 decimal place is allowed for each of the inputs named above.

2. Knowing the offered contract, each employee independently chooses an effort level, i.e., the H-employee chooses e_H, the L-employee chooses e_L. Restrictions are: $0 \leq e_H \leq 10$ and $0 \leq e_L \leq 10$. Again, up to 1 decimal place is allowed. One unit of effort leads to exactly one unit of output the employer is selling. The **gross output** Q thus equals the sum of chosen efforts.

This ends the interactions of a period. Payoffs result as follows:

- Employer: $(24 - r) \cdot (e_H + e_L) - 2 \cdot W$.
- H-employee: $W + r \cdot e_H - 1 \cdot (e_H)^2$.
- L-employee: $W + r \cdot e_L - 6 \cdot (e_L)^2$.

After each period every participant gets to know effort levels, gross output, and payoffs of all participants.

2.2 Calculator:
Additionally, the software provides a **calculator** to each participant. You can use this calculator for two minutes in every period, after which you have to make your decision at the latest. The calculator allows every participant

- to calculate the employer's payoff for various levels of W and r and various effort levels e_H und e_L, and
- to calculate the employees' payoffs for various levels of W and r and various effort levels e_H und e_L.

Note that as an employer you are only able to make conjectures about effort levels since you do not know the employees' decisions yet. As an employee, however, you know the decisions of the employer. They are preset in the calculator. You will not learn the employer's conjectures.

2.3 Additional restrictions:

As an **additional restriction for the employees**, you are limited to choose effort levels – e_H or e_L – that guarantee that payoffs are larger than or equal to zero in each period. You can check this restriction with the help of the calculator.

As an **additional restriction for the employer**, you are limited to offers W and r that, in addition to the conjectures about effort levels e_H and e_L, also given by yourself, guarantee that expected payoffs are larger than or equal to zero in each period. You can check this restriction with the help of the calculator.

These restrictions imply that period payoffs smaller than zero are only possible for employers, e.g. if effort levels are below conjectured efforts. However, the employer is able to restrict this risk by choosing W and r appropriately; payoffs larger zero should be the norm. Payoffs of employers and employees are summed up over the first two periods that are played as described above.

3. Periods 3 and 4:

Before the third period, each employee may choose between two alternatives:

I. to **keep the L-employee** or

II. to **lay off the L-employee**.

The H-employee will always be kept.

If the employer hires the L-employee again (case I), periods 3 and 4 are played analogously to periods 1 and 2.

If the employer lays off the L-employee (case II), he or she will receive a payment of 15 ECU from the experimenters (not from the employer) in each of the periods 3 and 4. Consequently, the L–employee does not make any decisions and does not learn the other participants' payoffs in periods 3 and 4. In each of the periods of this case (II), the employer offers a new contract to the H-employee only. The same bounds for contracts and effort levels apply. The payoff of the H-employee is calculated as before. Of course, the employer now earns $(24-r)\cdot e_H - W$. Restrictions are unchanged; calculators are provided again.

Payoffs of periods 3 and 4 are added to those of periods 1 and 2 and to the participation fee, are converted into euro, and are paid out anonymously and

in cash at the end of both experiments. If the employer's payoff from periods 1 to 4 is smaller than zero, it will be subtracted from the participation fee. If the rest is smaller than zero, it will be offset against the payoffs from the other experiment. If there is still a debt, this has to be paid for at the end of both experiments – either in cash or by administrative work. Please note again as an employer that this situation can be avoided almost completely by choosing W and r appropriately; payoffs larger zero should be the norm. The employees' payoffs are always larger than or equal to zero.

In the following, last experiment, you will not interact with the same participants as in this experiment again. Before we start the experiment, you have to answer some control questions.

■ **Experiment 2:**
We will now repeat the same experiment one more time, i.e., **all 4 periods** are played again. This means that, again, you will receive your **participation fee** and additional payments, depending on your decisions. Payoffs of all periods are added, converted, and paid out as before. Furthermore, you keep the same role as in the previous experiment, but it is guaranteed that no one will be matched with the same participants again.

Please stay silently at your seat at the end of the experiment until we call you individually and anonymously with the help of your box number and pay you off.

B List of performed tests

The p-values given in section 4 were obtained, using the tests listed in Table 8 below. All tests were performed two-sided. The first column depicts the treatment(s), the second the data used (all observations (1), only observations where firms did downsize (2), or only observations where firms did not downsize (3)), the third the number of observations, and the fourth the category of tests (tests of benchmark levels (BL), or of differences between observations without and after downsizing (W/A)). The performed tests are listed in the last column. In two important cases with rather small sample sizes, we explicitly mention that results stayed the same when using non-parametric instead of parametric statistics.

Tr.	Data	#	Cat.	Tests; comments
AH	1	48	BL	One-sample t test: all; $reldev_{h,W}, reldev_{l,W}$: only 46 obs., $reldev_{h,A}$: 45 obs.
AH	1	48	W/A	Dependent sample t test: all
AH	2	35	BL	One-sample t test: all; $reldev_{h,W}, reldev_{l,W}$: only 33 obs., $reldev_{h,A}$: 32 obs.
AH	2	35	W/A	Dependent sample t test: all
AH	3	13	BL	One-sample t test: r_W ,r_A, $w_{h,W}$, $w_{h,A}$, $w_{l,A}$, $dev_{h,W}$, $dev_{l,W}$, $reldev_{h,A}$; one-sample Wilcoxon rank sum test: else
AH	3	13	W/A	Dependent sample t test: r; dependent sample Wilcoxon test: else
AL	1	32	BL	One-sample t test: all
AL	1	32	W/A	Dependent sample t test: all
AL	2	25	BL	One-sample Wilcoxon rank sum test: all
AL	2	25	W/A	Dependent sample t test: ω_l; dependent sample Wilcoxon test: else
AL	3	7	BL	One-sample Wilcoxon rank sum test: all
AL	3	7	W/A	Dependent sample Wilcoxon test: all
UH	1	16	BL	One-sample t test: r_W, Π_W, $\omega_{h,W}$, $\omega_{h,A}$, $\omega_{l,W}$; one-sample Wilcoxon rank sum test: else
UH	1	16	W/A	Dependent sample t test: Π (the result, p = .056) does not greatly change for a Wilcoxon test (p = .093)); dependent sample Wilcoxon test: else
UH	2	13	BL	One-sample t test: r_W, $\omega_{h,W}$, $\omega_{h,A}$, $\omega_{l,W}$; one-sample Wilcoxon rank sum test: else
UH	2	13	W/A	Dependent sample t test: ω_l (results identical when using Wilcoxon instead); dependent sample Wilcoxon test: else
UH	3	3	BL	One-sample Wilcoxon rank sum test: all
UH	3	3	W/A	Dependent sample Wilcoxon test: all

Table 8: Performed tests

C List of tests for differences between AH and UH

All tests were performed as independent two-sample test. Sample sizes were 48, 35, and 13 for all firms, only firms D, and only firms ND in treatment AH and 16, 13, and 3 for treatment UH. In Table 9 below, we abbreviate the performed tests as follows: Welch Satterthwaite t tests (WS), Wilcoxon tests (WI), robust rank-order tests together with Wilcoxon tests (rro). P-values are also given.

	a) All		b) Only firms D		c) Only firms ND	
	Without	After	Without	After	Without	After
F	.115 (.102) (rro)	.486 (WI)	.255 (.223) (rro)	.681 (WI)	.489 (WI)	.500 (WI)
r	.565 (.577) (rro)	.510 (WI)	.950 (WI)	.192 (.209) (rro)	.709 (.680) (rro)	.536 (WI)
e_h	.715 (WI)	.491 (WI)	.539 (WI)	.236 (.256) (rro)	.666 (.596) (rro)	.779 (WI)
e_l	.409 (WI)		.780 (.803) (rro)		.203 (.179) (rro)	.575 (WI)
Π	.997 (WI)	.444 (WI)	.567 (.555) (rro)	.292 (.320) (rro)	.814 (.818) (rro)	1.000 (WI)
ω_h	.764 (WS)	.526 (WI)	1.000 (WS)	.219 (.238) (rro)	.710 (.704) (rro)	.439 (WI)
ω_l	.307 (WI)	.489 (WI)	.532 (WS)		.637 (WI)	.800 (WI)
$absdev_h$	.218 (WI)	.978 (WI)	.574 (WI)	.720 (WI)	.143 (WI)	.511 (WI)
$absdev_l$	.207 (.179) (rro)		.738 (WI)		.032* (WI)	.409 (WI)
$reldev_h$	.221 (WI)	.963 (WI)	.583 (WI)	.830 (WI)	.039* (WI)	.621 (WI)
$reldev_l$	.237 (WI)		.790 (WI)		.041* (.041) (rro)	.312 (WI)
$relconj_h$	.459 (.456) (WI)	.820 (WI)	.457 (WI)	.683 (WI)	.763 (.718) (rro)	.282 (WI)
$relconj_l$	.582 (WI)		.941 (.950) (rro)		.421 (WI)	.692 (.686) (rro)

Table 9: Main results, treatment UH

D Logistic regression results, AH and AL

Given below are translated, relevant SPSS 16 output tables for the logistic regression of the dependent dummy variable *Down* defined in subsection 4.3 for treatment AH (Tables 10 - 13) and treatment AL (Table 14).

Due to the few undefined values of $reldev_h$, the final regression for treatment AH was performed for only of 46 of 48 values. The results differ negligibly when using all 48 observations or only Π.

For step 0 of treatment AH, where predictions suggest that downsizing took place, 33 out of 46 observations are predicted correctly, a ratio of 71.7%. This classification table is omitted due to space restrictions.

Variables in the equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1	Π	-.075	.032	5.666	1	.017	.928
	Constant	3.245	1.150	7.957	1	.005	25.663

Table 10: Logistic regression, AH(48): Relevant variables

Classification Table

			Predicted Down (1=yes)		Percentage of correct predictions
		Observed	.00	1.00	
Step 1	Down (1=yes)	.00	5	8	38.5
		1.00	2	31	93.9
		Overall			78.3
The cut value is .500					

Table 11: Logistic regression, AH(48): Classification Table

Variables not in the equation

			Value	df	Sig.
Step 1	Variables	F	.277	1	.599
		r	.076	1	.783
		$conj_h$	.000	1	.992
		$conj_l$	.071	1	.790
		e_h	.476	1	.490
		e_h	.396	1	.529
		ω_h	.027	1	.870
		ω_l	.435	1	.510
		$absdev_h$	2.769	1	.096
		$absdev_l$	.355	1	.551
		$absconj_h$	.672	1	.412
		$absconj_l$	.006	1	.938
		$reldev_h$	.549	1	.459
		$reldev_l$	.326	1	.568
		$relconj_h$	.019	1	.889
		$relconj_h$	.003	1	.955

Table 12: Logistic regression, AH(48): Other variables

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	10.021	1	.002
	Block	10.021	1	.002
	Model	10.021	1	.002

Table 13: Logistic regression, AH(48): Omnibus Tests

Variables not in the equation

			Value	df	Sig.
Step 0	Variables	F	.010	1	.919
		r	.167	1	.683
		$conj_h$	2.760	1	.097
		$conj_l$	.974	1	.324
		e_h	.615	1	.433
		e_l	.060	1	.806
		Π	1.706	1	.192
		ω_h	.101	1	.751
		ω_l	.170	1	.680
		$absdev_h$	.681	1	.409
		$absdev_l$	.257	1	.612
		$absconj_h$	1.408	1	.235
		$absconj_l$	.761	1	.383
		$reldev_h$	1.449	1	.229
		$reldev_l$	.441	1	.507
		$relconj_h$	2.906	1	.088
		$relconj_l$	.171	1	.680

Table 14: Logistic regression, AL