

IZA DP No. 4795

Bonus Payments and Reference Point Violations

Axel Ockenfels
Dirk Sliwka
Peter Werner

February 2010

Bonus Payments and Reference Point Violations

Axel Ockenfels

University of Cologne

Dirk Sliwka

*University of Cologne
and IZA*

Peter Werner

University of Cologne

Discussion Paper No. 4795
February 2010

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Bonus Payments and Reference Point Violations^{*}

We investigate how bonus payments affect satisfaction and performance of managers in a large, multinational company. We find that falling behind a naturally occurring reference point for bonus comparisons reduces satisfaction and subsequent performance. The effects tend to be mitigated if information about one's relative standing towards the reference point is withheld.

JEL Classification: D03, M52

Keywords: reference points, incentives, bonus payments, job satisfaction, job performance

Corresponding author:

Dirk Sliwka
University of Cologne
Albertus-Magnus-Platz
D-50923 Köln
Germany
E-mail: dirk.sliwka@uni-koeln.de

^{*} We thank David Jaeger and seminar audiences at Bonn, Cologne, Dortmund, Edinburgh, Granada, Heidelberg, Karlsruhe, London, Magdeburg, Munich and Toulouse for helpful comments. We also thank the board and many managers of the studied company for sharing their data and views with us. Ockenfels thanks the German Science Foundation for financial support.

1. Introduction: The bonus system, a reference point and hypotheses

Laboratory experiments and survey studies suggest that reference points can significantly affect the evaluation of economic outcomes and the effectiveness of incentives. Bewley (1999), for instance, concludes from his surveys covering more than 300 US firms that “Few workers have a precise notion of market wages or wages that are fair in some absolute sense. Rather, they make comparisons with their own past pay and with the pay of co-workers.” Moreover, his studies indicate that “within a company, pay inequity offends (indeed, sometimes outrages) employees and destroys trust.” Behavioral economics responded to the evidence with modeling reference-dependent preferences. Among the most influential models are Kahneman and Tversky’s (1979) prospect theory with the reference point being the status quo, Rabin’s (1993) model of reciprocity with the reference point being the perceived (un)kindness of the opponent, Fehr and Schmidt’s (1999) as well as Bolton and Ockenfels’ (2000) models of social comparison with the reference point being some norm of outcome fairness, and, more recently, Köszegi and Rabin’s (2006, 2007, 2009) models with the reference point being the decision maker’s rational expectations held recently about outcomes. Much of the related research in labor market contexts has been influenced by Akerlof and Yellen’s (1990) fair wage-effort hypothesis, postulating that workers withdraw effort as their actual wage falls short of their fair wage. Yet, despite the growing role of reference points in the economics literature in general and the labor market literature in particular, there is very little evidence based on naturally occurring field data. Part of the reason is that, without experimental control, it is often difficult, if not impossible, to identify the ‘right’ frame of reference in labor relationships.²

² Akerlof and Yellen (1990), Frank (1997) and Frank and Sunstein (2001) summarize much of the earlier economic, sociological, psychological and biological literature regarding the importance of reference points and relative position for satisfaction and behavior. They make a similar point regarding the difficulties with field studies. Gächter and Fehr (2002) survey laboratory labor market research. For more recent evidence see Götte, Fehr and Zehnder (2009) and the references therein. See also Fließbach et al. (2009) for a neuroeconomic perspective, and Gneezy and List (2006) as well as Levitt and List (2007) for field experiments bridging laboratory and naturally occurring field data, as well as for the potential problems with experimental approaches. Two intriguing studies that were able to overcome the reference point problem in a natural field environment are Mas (2006), who shows that police performance is sensitive to pay rises compared to reference points set by final offer arbitrations, and Kwon and Meyerson Milgrom

In this paper, we make use of a unique data set to study the causes and effects of reference point formation in the context of a bonus plan for managers in a large multinational company.³ The important feature of the company's bonus system is that it produces a clear reference point, largely consistent with the behavioral economics approaches mentioned above. Moreover, our data allow us to match demographic characteristics, salaries, and bonus payments from personnel records with survey responses of individual managers in Germany (where the headquarters are located) as well as the US.

The company's bonus system stipulates that each year each manager is assigned a bonus target (also called the 'bonus budget'). The individual target depends on company performance, divisional performance and the manager's salary grade, among other variables. The bonus payment is set by the supervisor as a percentage of the individual manager's target (the so called 'bonus percentage').⁴ A manager's bonus percentage is restricted by his performance rating, which is determined and communicated several weeks before the individual target is calculated and his bonus percentage is assigned. The performance rating is chosen from a five point rating scale ('excellent', 'above average', 'fully meets expectations', 'below average', and 'inadequate'). For each rating, the bonus percentage must lie within a certain interval. By far most managers' rating is 'fully meets expectations', and the bonus percentage of these managers must be chosen between 80% and 110%. Better rated managers must be assigned 110% or more, and worse rated managers must be assigned less than 80%.

Within this context, the 100% bonus percentage is a natural reference point. For one, it seems reasonable that, if reference points play a role at all, managers care about not falling behind their individual *target*. Moreover, the performance rating reinforces the

(2009), who analyze exogenous changes in workers' relative wages during M&As and find that status in the workplace affect their turnover decisions.

³ The company has more than 100.000 employees worldwide. The largest subsidiaries are located in Germany and the US. Our study covers the incentive scheme for all managers in both countries. The term "manager" refers to executives and to all other employees not covered by a collective wage agreement. The company asked us to preserve its anonymity.

⁴ In this paper we refer to managers as those who are assigned bonuses, and to supervisors as those who assign bonuses. Managers in upper hierarchical levels can also be supervisors, something we will exploit in our analyses in Section 4.

100% reference point of the bonus percentage in the sense that a manager who was rated ‘fully meets expectations’ is likely to interpret this as hitting the target, and so is likely to *expect* a bonus not below his target (= 100%). Finally, the 100% reference point is strengthened by the fact that the bonuses paid to all managers in a department are limited such that the sum of bonus payments cannot exceed the sum of all individual bonus budgets. As a result, the department’s supervisor always has the option to choose an equal distribution in the sense of paying all managers in her department the respective budget (that is, a bonus percentage of 100%). Yet if she decides to give one manager more, others must get less. Therefore, falling behind the 100% bonus percentage implies not only that the manager is individually below target (and probably behind expectation), but also that others get a share of one’s own budget, leading to socially unfavorable inequality measured in bonus percentage terms. In this sense, and in line with the role of inequity aversion observed in various economic contexts, the 100% bonus percentage is also a natural *social* comparison standard.

One characteristic feature of reference-dependent preferences is loss aversion: falling below the respective reference point (a loss) has a larger effect than a same-sized gain above the reference point. For example, in one-person decision tasks many theoretical and empirical studies, starting with Kahneman and Tversky (1979) and further developed by Köszegi and Rabin (2006, 2007 and 2009), suggest that compared to the reference point, losses are much more powerful than gains. In social contexts, Bolton (1991) and Fehr and Schmidt (1999), among others, present evidence and theory suggesting that disadvantageous inequality hurts more than advantageous inequality, which might even often be irrelevant altogether. Taken together, we hypothesize that assigning a manager less than 100% bonus percentage (which we call a “reference point violation”) will strongly negatively affect a manager’s *satisfaction*, while assigning more yields – if at all – a smaller effect.

Regarding the effect of reference point violations on the managers’ *performance*, there are competing hypotheses. On the one hand, one might think that (the possibility of) falling below the reference point creates additional incentives, beyond the pure monetary implications, to improve one’s performance and to reach the bonus target. From an

economics point of view, the bonus system resembles a tournament, because a total department's bonus budget is fixed, and the bonus percentages rank workers according to performance. Tournament theory suggests that compensating workers according to a 'percentage rank' of the performance may increase effort (Lazear and Rosen, 1981).⁵ Furthermore, the psychological literature emphasizes the idea that goals create reference points, and that there may be not only an extrinsic but also an intrinsic interest in achieving targets set by others (see Heath et al., 1999, building on Locke, 1968, Locke and Latham, 1990). Camerer et al. (1997), Fehr and Götte (2007) and Abeler et al. (forthcoming) provided convincing field and laboratory evidence that income targets act as reference points and strongly affect labor supply decisions. Finally, field experiments in public good contexts show that information about one's own cooperation falling behind others' contributions can provoke increased efforts (Frey and Meyer, 2004, Chen et al., forthcoming).

On the other hand, there is also literature suggesting that reference point violations can decrease performance. Most prominently, Akerlof and Yellen (1990) motivate their fair wage-effort hypothesis, postulating that reference point violations reduce effort, with research in psychology and sociology and as it "accords with common sense" (see also the quote from Bewley in our first paragraph). More recently, behavioral economics research suggests that if reference point violations are perceived as an unkind act, managers may reciprocally respond with unkindness, and decide to reduce effort levels in the future (Rabin, 1993, Dufwenberg and Kirchsteiger, 2004, Falk and Fischbacher, 2006). Indeed, many laboratory and some field studies, starting with Fehr, Kirchsteiger and Riedl (1993), suggest that reciprocity can play an important role in labor relationships.⁶ Other experimental studies find that wage comparisons affect behavior in experimental labor relations (see, for example, Rivas and Sutter, 2008, Greiner et al., 2009, Clark et al., forthcoming).

⁵ Grund and Sliwka (2005) and Demougin and Fluet (2003) have shown that the incentive effect of tournaments may even become stronger when social reference points matter and individuals are inequity averse.

⁶ See Cooper and Kagel (forthcoming) for a recent survey on other-regarding preferences. Also, there is laboratory evidence that subjects, who fall below a reference point, be it the status quo (Kahneman and Tversky, 1979) or a social reference point (Bolton and Ockenfels, forthcoming), are willing to take more risks, which may additionally affect performance.

So, summing up, theory does not yield unambiguous predictions regarding the effect of reference point violations on employee performance. Moreover, to our knowledge, there is no earlier field evidence about the effect of reference point violations on employee performance in companies. Thus, we refrain from formulating a null hypothesis about how reference point violations affect performance. However, we expect that any satisfaction and performance effects due to reference-dependent preferences that we observe in Germany are mitigated in the US. The reason is that while the bonus schemes are otherwise identical, bonus percentages are explicitly communicated to managers in Germany but not to managers in the US. This is because labor regulations in Germany require a much higher degree of transparency in compensation systems. As bonus percentages cannot easily be inferred from absolute bonus payments and other data available to US managers (see Section 2), we hypothesize that the lack of transparency in the US system diminishes the role of the 100% reference point.

Overall, our data strongly support the key predictions of reference-dependent preferences and of Akerlof and Yellen's (1990) fair wage-effort hypothesis. We find that violations of the managers' 100% reference point significantly negatively affect both satisfaction and future performance in Germany, where reference point violations are transparent. This holds in particular if we restrict our analyses to managers who have been rated 'fully meets expectations' so that potentially confounding rating effects are excluded. However, no such reference point effects are found for the company's managers in the US, who operate under a less transparent but otherwise identical bonus system as their colleagues in Germany.

Section 2 describes our data. Section 3 analyzes the determinants of job satisfaction and establishes the strong relevance of the 100% reference point. Section 4 investigates the impact of reference point violations on performance. Section 5 discusses the findings and concludes.

2. A first look at the managers' compensation

Supported by the board of the company, we conducted a survey among the managerial staff in Germany (autumn 2007) and the US (summer 2008). As the survey was part of a larger study, managers were asked some 60 questions about workplace characteristics and other work-related aspects, including job satisfaction. We independently collected data about performance evaluations, compensation, demographic characteristics, and department affiliations of the managers over the years 2004-2006 (Germany) and 2004-2007 (US) from the personnel records of the firm. The technical environment allowed us to connect this background data with the survey answers of the individual managers in a way that guaranteed anonymity of the participants of the survey.⁷ Altogether 4,997 managers took part in the survey (3,122 in Germany and 1,875 in the US), which corresponds to a participation rate of 59% and 41%, respectively. This section reports compensation data of all managers in the company, and the next section investigates the determinants of reported job satisfaction.⁸

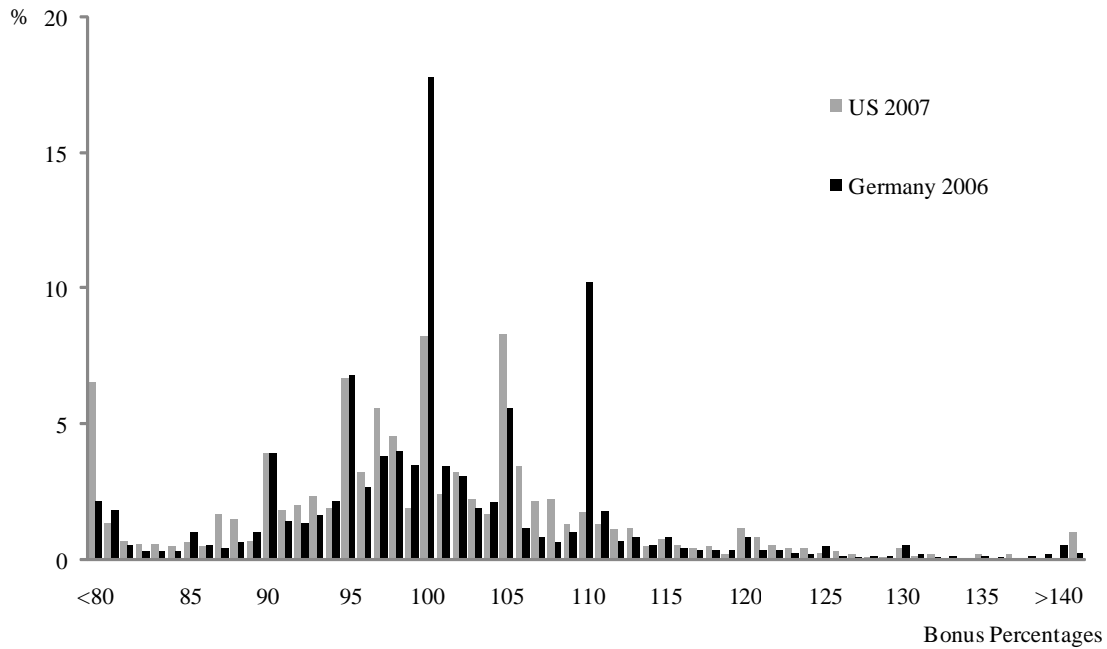
Salaries and absolute bonus payments of German and US managers are roughly comparable during the period of analysis and remained largely stable over the years that we analyze. In 2007, the average fixed salary of a manager in Germany (in the US) was 117,300 USD (113,500 USD). The bonus payments are substantial, as they account for some 20% of fixed salaries on average (25,600 USD in Germany and 22,500 USD in the US).

⁷ The procedure guaranteed that no party involved in data generation and processing could combine the information about contact details, background data and survey answers of the participants. First, we collected the administrative data from the personnel records of the company and encoded the data set. Then the company received the encoded data set, added email addresses of all managers and transferred the data to an independent consulting firm who administered the survey. This firm then invited the managers by email to take part in the survey, which was conducted in the intranet of the firm with an anonymous code, and deleted the email addresses afterwards. Finally, we received the encoded data set including survey answers for our analysis without information that allowed identification of individuals. Managers were informed on the first page of the survey that their answers would be matched anonymously with their individual compensation and background data and had to agree to this procedure to participate in the survey.

⁸ Comparing participating and non-participating managers, we do not find significant differences in background or compensation variables, so that we have no reason to suppose that the representativeness of the sample is limited in any way. The lower response rate in the US can be attributed to the fact that the US survey was run at a time when a larger fraction of the employees was on holiday.

However, there are considerable differences when we look at the distributions of bonuses in Germany and the US. Figure 1 shows the bonus percentage distributions in Germany for 2006 and in the US for 2007, the respective years of our survey (the distributions are very similar in the other years).

Figure 1. Distribution of Bonus Percentages in Germany and the US



Compared to the US data, bonus percentages in Germany are more compressed towards the 100% reference point. Most strikingly, in Germany, there is a significant peak at 100%, with 17.6% of the managers receiving exactly their target bonus, while in the US only 8.3 % of the managers receive a 100% bonus percentage ($p = 0.001$, two-sided χ^2 -test). The next largest peak in Germany is at 110%, which is the feasible bonus closest to 100% for the ‘above average’ managers. In the US, on the other hand, 110% exhibits no specific prominence – only 4.3% of all ‘above average’ managers in the US receive exactly 110%, which is about one eighth of the corresponding number in Germany (36.1%; $p = 0.001$). Also, the share of managers in Germany with bonus percentages below 80% (above 110%) is with 2.2% (11.6%) compared to 6.6% (13.4%) in the US

(weakly) significantly smaller with $p = 0.001$ (0.054). Looking at the whole distribution, the standard deviation of bonus percentages in Germany is with 12.2 also significantly smaller than the standard deviation of 15.1 observed in the US ($p = 0.001$, two-sided Levene test for differences in variances).

We attribute the differences in the distributions of bonus percentages to the different information policies of the company in Germany and the US. While the short term incentive system otherwise follows the same rules as in Germany, managers in the US are informed only about absolute bonus payments, but not about the bonus budgets or the assigned share of the budget.⁹ Thus, in the US, ‘fully meets expectations’ managers cannot observe their relative standing towards the reference point. Because pushing managers below 100% reduces satisfaction and performance if the reference point violations are transparent (see Sections 3 and 4), there is more reason for compression towards the reference point for ‘fully meets expectation’ managers in Germany than in the US. A similar argument holds for ‘above average’ managers. The more managers are pushed above 100% the larger the budgetary pressure to violate other managers’ reference point, which is harmful when the violations are transparent. A bonus of 110% for ‘above average’ managers, as often observed in Germany, minimizes this adverse effect.

Our evidence provides empirical support for other studies that associate compressions in payments with reference points and social concerns. In particular, it is in line with Frank’s (1984) seminal work showing that pay compression, relative to worker productivity, is the result of workers caring for their relative position (see also Frank 1985) as well as Akerlof and Yellen’s (1990) fair wage-effort hypothesis.¹⁰

⁹ Due to tighter transparency regulations, managers in Germany, when receiving their yearly bonus letters, are explicitly told their bonus percentage together with their bonus payments. In the US, on the other hand, managers are only told their bonus payment. Individual bonus budgets are computed based on salary grade and several performance indicators measuring the financial success of the company as a whole as well as of the relevant division, subject to exceptions and discretion. These complexities make it generally very difficult for a US manager to learn his individual budget. In fact, many open survey answers of US managers show that some managers wish to get better informed. Typical comments include: “The letter that comes with the award should clearly spell out your rating”, and “At a minimum, the employee should be told what % was applied”.

¹⁰ Our evidence is also in line with psychological and economic literature emphasizing that subjective performance evaluations tend to be compressed; see Murphy and Cleveland (1995) for an extensive survey

We summarize our findings in:

Observation 1. *The distribution of bonus percentages is more compressed towards the 100% reference point in Germany, where bonus percentages are made transparent, than in the US, where managers do not learn their bonus percentages. In particular, there is a significant fraction of managers receiving 100% in Germany, while the corresponding frequency is less than half as high in the US.*

3. *How does the bonus matter for job satisfaction?*

In this section, we investigate whether there is a kink in the managers' preferences at the 100% reference point as postulated by loss aversion and inequity aversion. While we cannot measure preferences directly, our study design allows us to investigate how deviations from the reference point affect perceived job satisfaction. A commonly used proxy for job satisfaction is the degree of approval to the statement: "I am very satisfied with my job."¹¹ In our survey, managers could respond to this question on a scale from 1 (totally disagree) to 7 (fully agree), so that increasing values of the score reflect higher satisfaction. The distributions of answers in the US and Germany are rather similar, with a mean of 5.28 (standard deviation 1.25) in Germany and 5.24 (1.25) in the US.

Following Freeman (1978), we use a unit normal transformation of the satisfaction score by subtracting the sample mean from a person's score and divide the result by the standard deviation of the sample. We estimate linear regressions with job satisfaction as the dependent variable for the years 2006 (Germany) and 2007 (US), in which the respective survey studies took place. To control for the background of the manager, we include dummies for age intervals, gender, firm tenure and promotion in the previous business year. As we can track the hierarchical relationships in our data, we use fixed effects per supervisor in order to also control for unobserved supervisor-specific effects (e.g., general leadership qualities) that can be expected to influence job satisfaction.

of the psychological literature and Murphy (1991) and section 2.2.3 in Prendergast (1999) for an economics perspective. Bolton et al. (2009) find rating compressions in electronic reputation systems and attribute their finding to reciprocal feedback giving.

¹¹ See for instance Freeman (1978), Clark and Oswald (1996), and Hamermesh (1977, 2001). Frey and Stutzer (2002) and Luttmer (2005) discuss the use of self-reported measures of well-being in economics.

Finally, we control for the direct effects of the performance rating on job satisfaction by calculating each model also with dummies for performance ratings. Table 1 shows our model specifications.

In our first specification, we focus on the impact of absolute bonus payments and bonus percentages. As a manager's bonus budget depends on his position in the company, his salary, and on the performance of both the company and his division, bonus payments and bonus percentages are not collinear, so that the effect of the bonus payment can be disentangled from that of the bonus percentage. In Model 1, the coefficient of the absolute bonuses is insignificant, while the coefficient of bonus percentages is highly significant and positive for both Germany (Model D1) and the US (Model US1).¹² Hence, despite their substantial economic relevance, absolute bonus payments do not have a robust influence on job satisfaction – while bonus percentages do.

When we include rating dummies (Model 2), the described effect remains robust for Germany, yet the US coefficient of the bonus percentage becomes statistically as well as economically insignificant.¹³ As managers in the US do not learn their bonus percentages, they cannot infer their position *within* the percentage interval corresponding to a given rating. Hence, controlling for the rating, one's relative position towards the reference point does not matter for satisfaction. In Germany however, one's relative standing within a given performance rating is known to managers and can therefore still substantially correlate with job satisfaction.

As we have pointed out in Section 1, one characteristic of reference-dependent preferences is loss aversion: people dislike falling below a reference point more than they like exceeding the reference point by the same amount. Therefore, we hypothesize that reference point violations – getting a bonus percentage below 100% – significantly

¹² Absolute bonus payments and bonus percentages have both a positive and mostly significant impact when only one variable is included in the regression. Also, observe that the regressions consistently do not show a significant positive relation between the salary in the year of the survey and self-reported satisfaction. This evidence is in line with a large (but sometimes disputed) body of research suggesting that an increase in already high income levels does not substantially affect satisfaction levels (see, for instance, Easterlin, 1974, Clark and Oswald, 1996, and Clark, 1999).

¹³ Regarding performance ratings we observe that in both countries the dummy for a rating 'below average' exhibits a strong negative effect on satisfaction. This is consistent with the fact that the performance rating affects future career opportunities, as the probability of promotion increases with higher ratings.

decrease job satisfaction, whereas increasing the bonus beyond 100% has a smaller effect on satisfaction. Model 3 tests this hypothesis.¹⁴ In this specification we allow for the possibility that the effect of the bonus percentage has different slopes below and above the reference point. The model is defined as

$$\begin{aligned} \text{JobSatisfaction}_{is} = & \alpha + \beta \cdot X_{is} \\ & + \gamma \cdot (z_{is} - 100) \cdot I_{\{z_i > 100\% \}} \\ & + \delta \cdot (100 - z_{is}) \cdot I_{\{z_i < 100\% \}} + a_s + \varepsilon_{is} . \end{aligned}$$

z_{is} is the bonus percentage of individual i and $I_{\{z_i > 100\% \}}$ ($I_{\{z_i < 100\% \}}$) is a dummy variable taking the value 1 if the bonus percentage is above 100% (below 100%). Hence, γ captures the effect of a positive deviation and δ of a negative deviation from the 100% bonus percentage. If 100% constitutes a reference point relative to which employees evaluate their bonus, we expect γ and δ to be different from each other. In particular, loss aversion in our context implies $|\gamma| < |\delta|$. Finally, a_s denote the supervisor-specific effects, whereas ε_{is} captures the idiosyncratic error term per manager.

Consistent with loss aversion we find that reference point violations significantly decrease satisfaction levels, while increases above 100% do not increase self-reported satisfaction in Germany. The asymmetric perceptions of being below and above the reference point are exhibited in both countries – even though US managers do not know their bonus percentages. The reason is that bonus percentages are correlated with performance ratings (as explained in Section 1), so US managers can use their ratings as a proxy of their relative standing. E.g., a manager rated ‘above average’ (‘below average’) knows that he must have a bonus percentage above (below) 100%. However, when controlling for the effects of performance ratings (Model 4), the asymmetric effect of

¹⁴ As absolute bonus payments do not have a significant impact on reported satisfaction, we do not include them in the following models. The results do not change, however, if bonus payments are added as an explanatory variable.

deviations from 100% budget vanishes in the US sample while remaining robust in Germany.¹⁵

As a further robustness check, we estimate Model 3 with a sample including only managers who receive the ‘fully meets expectations’ rating. Managers rated ‘fully meets expectations’ in Germany can still exhibit a reference point effect, because they know their respective relative standing towards the reference point. Managers rated ‘fully meets expectations’ in the US, however, do not know whether they actually reach their target budget. The data analysis for Germany shows that the loss aversion effect remains stable in the German sample at about the same level as before, suggesting that the loss aversion effect observed in Models D3 and D4 is indeed driven by deviations from the 100% reference point. For the US data, the effects regarding bonus percentages again disappear.¹⁶

Following our discussion in Section 1, the statistical analyses above assumed that the reference point (i.e., the kink in the estimated function) is at 100%. In order to further justify our choice and to also allow for different shapes, we re-estimate our satisfaction models for Germany and the US, including dummies for percentage intervals instead of bonus percentages and controlling for performance ratings.¹⁷ The reference group consists of managers who receive exactly their budgets ($z_i = 100\%$). In the German sample, both dummies below 100% have statistically significant negative signs, indicating a lower satisfaction score compared to managers at the 100% threshold. Both dummies for intervals above 100% are positive, statistically insignificant and small in size. In the US, too, both dummies below (above) 100% have negative (positive) signs, but here all interval dummies are insignificant.

¹⁵ Our conclusion is identical if we use ordered-probit regressions (without supervisor fixed effects) to account for the ordinal character of the satisfaction score.

¹⁶ The coefficients γ and δ are significantly different in all relevant models for Germany (at $p = 0.001$ for D3 and D4 and $p = 0.003$ for D5; two-sided Wald tests), whereas in case of the US, γ and δ differ in specification US3 ($p = 0.001$), but not in US4 ($p = 0.861$) and US5 ($p = 0.326$).

¹⁷ See Table 3 in the Appendix. The results are similar if we exclude the performance ratings dummies from the models.

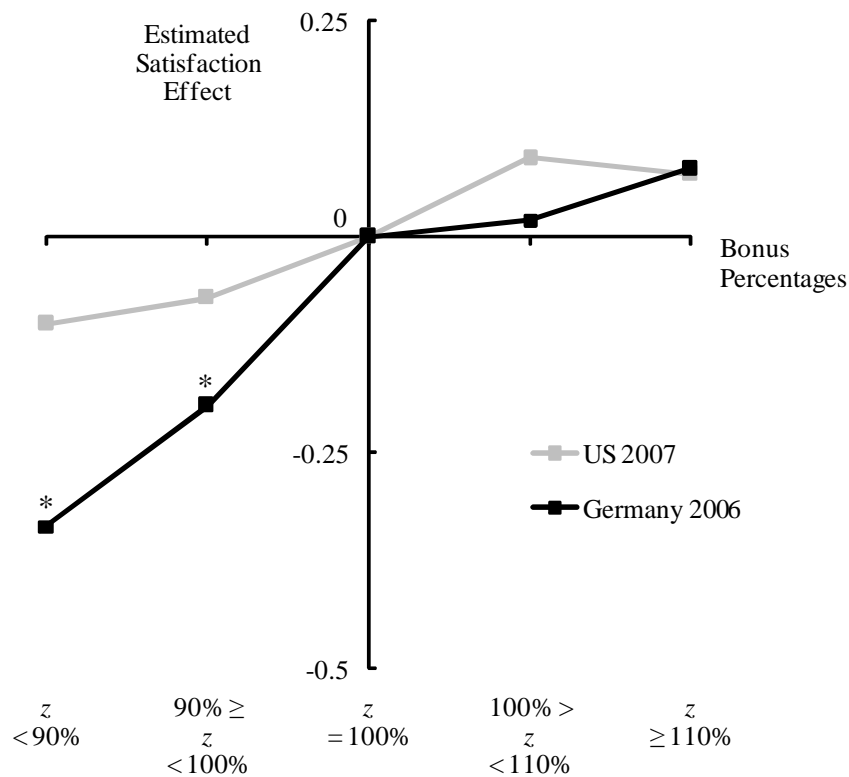
Table 1. Satisfaction Effects of Reference Point Violations
(Dependent Variable: Adjusted Satisfaction Score)

Country No.	Germany					United States				
	D1	D2	D3	D4	D5	US1	US2	US3	US4	US5
Salary (000s)	-0.002 [0.004]	-0.002 [0.004]	0.002 [0.002]	0.001 [0.002]	0.001 [0.002]	-0.005 [0.004]	-0.006 [0.004]	-0.004 [0.002]	-0.004 [0.002]	-0.003 [0.004]
Bonus Payment (000s)	0.011 [0.010]	0.010 [0.010]				0.003 [0.006]	0.003 [0.006]			
Bonus Percentage	0.012*** [0.003]	0.012*** [0.004]				0.016*** [0.005]	0.001 [0.011]			
Positive Deviation from Bonus Budget (=100%)			0.002 [0.004]	-0.001 [0.006]	0.008 [0.014]			0.013 [0.010]	-0.002 [0.016]	0.033 [0.058]
Negative Deviation from Bonus Budget (=100%)			-0.031*** [0.005]	-0.030*** [0.006]	-0.032*** [0.008]			-0.021** [0.010]	-0.006 [0.016]	-0.018 [0.024]
Sample	All	All	All	All	'fully meets'	All	All	All	All	'fully meets'
Rating Dummies	No	Yes	No	Yes		No	Yes	No	Yes	
No. of observations	2091	2091	2091	2091	1544	957	957	957	957	599
R-squared (within)	0.03	0.04	0.04	0.05	0.04	0.03	0.04	0.03	0.04	0.04

Standard errors are given in brackets. *, ** and *** denote significance on the 10%, 5% and 1%-level, respectively. Control variables include supervisor fixed effects, age interval dummies, gender, total years of affiliation to the company, and a dummy for a promotion in the previous year.

Figure 2 illustrates the results. For Germany, estimated interval dummies exhibit the kinked shape at the reference point, which is implied by loss aversion. The described relation remains robust if the analysis is restricted to managers rated ‘fully meets expectations’. No such patterns are observed for the US data.

Figure 2. Relation between Bonus Percentage and Job Satisfaction relative to Managers with a 100% Bonus Percentage



The figure shows the estimated values for bonus percentage interval dummies on job satisfaction, controlling for compensation, performance rating, and demographic background of the managers and including fixed effects per supervisor (see Table 3 in the Appendix). Significant interval dummies are marked with an asterisk (*).

An important question is, of course, the direction of causality. One may for instance raise the argument that the managers’ performance is affected by their job satisfaction, and that therefore more satisfied managers receive higher bonus percentages. However, the pronounced kink in Germany seems inconsistent with such a reversed causality argument:

Without reference point effects, there is no reason to expect that a manager, who receives a bonus percentage of 100%, exhibits a much larger performance effect when being marginally more satisfied than when being marginally more dissatisfied. Moreover, the US data provide a complementary argument: If satisfaction drove performance and in turn affected bonus percentages, one should also observe such effects within performance ratings, regardless of whether bonus percentages are known or not. However, there is no such relation between percentages and satisfaction in the US, indicating that it is the information about bonus percentages that causally affect satisfaction. Observation 2 summarizes our findings.

Observation 2. *Transparent reference point violations lead to decreased job satisfaction. Negative deviations from the reference point have a much stronger effect on satisfaction than positive deviations of the same size.*

4. How do reference point violations affect job performance?

It may well be that while reference point violations negatively affect employee perceptions they are a useful instrument to generate performance incentives. In this section, we investigate the competing hypotheses regarding the consequences of reference point violations on performance as outlined in Section 1. Our analysis exploits the fact that we have information about the hierarchical relations between the managers. Hence, we can use a supervisor's rating as a measure for the performance in her department: If reference point violations negatively affect performance, one would expect that having a higher fraction of managers with a bonus percentage below 100% in one year should lead to a reduced department performance and thus also negatively affect the department supervisor's performance rating in the subsequent year.¹⁸ Furthermore, exploiting the panel structure of our data, we can control for unobserved heterogeneity by including supervisor fixed effects. Specifically, we estimate the following specification:

$$\text{Rating}_{st+1} = \alpha + \beta \cdot \text{DevRef}_{st} + \gamma \cdot X_{st} + a_s + \varepsilon_{st}.$$

¹⁸ When studying the effect of the bonus on performance, we cannot just conduct our analysis on the level of the managers' individual performance. Reference point violations might be the result of unobserved factors such as low ability or low willingness to perform, which in turn yields low ratings and low bonuses.

X_{st} is a vector of individual background variables of a supervisor s and a_s are supervisor fixed effects. The background variables include a manager's age and hierarchy level in the year $t + 1$. Our hypothesis is that the share of managers rated below the 100% reference point (measured by $DevRef_{st}$) among all 'fully meets expectations' managers in year t by supervisor s has a negative effect on her performance rating in year $t + 1$.¹⁹ Table 2 reports the results for Germany and the US.

Model 1 uses a unit normal transformation of the supervisor's performance rating in $t + 1$ as the dependent variable. We find that reference point violations negatively affect the performance of a given supervisor in Germany (Model D1). That is, the more managers in a department are pushed below 100% bonus percentage in a given year, the worse is the supervisor's subsequent rating.

In Model 2, we additionally control for the general evaluation behavior of the supervisor by including the relative shares of the performance ratings different from 'fully meets expectations' assigned by a supervisor in t .²⁰ The benefit of this specification is twofold: For one, while the supervisor fixed effects capture time constant unobserved factors, the heterogeneity of the team regarding individual performance may still vary over time. Controlling for the distribution of ratings captures the effect of this heterogeneity. In addition, the rating distribution determines the restrictions under which a supervisor can (re)distribute the bonus budgets. A supervisor with a team in which many managers have received 'excellent' ratings, for instance, may have no choice but to violate reference points of 'fully meets expectations' managers, because the system requires that a disproportionately large share of the total budget needs to be spend on the excellent managers.

We find that the empirical results are robust: A supervisor in Germany with a given rating distribution will have a lower performance in the subsequent year when violating the

¹⁹ Performance ratings are coded with numbers ranging from 5 for 'excellent' to 1 for 'inadequate'. The mean rating for all supervisors in the sample is 3.31 (standard deviation = 0.51) for German supervisors and 3.46 (standard deviation = 0.57) for US supervisors.

²⁰ The share of 'inadequate' ratings is not included in the regressions, though, as the lowest rating is hardly ever assigned in the years of our sample.

100% reference point of a higher share of her subordinate managers. The effects are also economically significant. A rough interpretation of the coefficient is the following: If a supervisor who assigns a bonus percentage of less than 100% to all of her ‘fully meets expectations’ managers instead gave all of them a bonus of 100%, her own performance evaluation in the next year would increase by 0.68 standard deviations. This implies that a supervisor with a performance evaluation at the median would additionally outperform 24.6% of all other managers when changing her evaluation behavior in this way. To further quantify the performance effect of reference point violations in a simple way, we estimate Model 2 with supervisor fixed effects and a dichotomous dependent variable equal to one, if the supervisor is assigned an ‘above average’ rating in the subsequent year, and zero otherwise.²¹ A coefficient for $DevRef_{st}$ of -0.253 suggests that a supervisor who violates the reference points of all ‘fully meets expectations’ managers in her department can increase the likelihood of receiving an ‘above average’ rating by more than 25% if she assigns all ‘fully meets expectations’ managers 100% of their budgets.

It is important to note that there is indeed much variation in terms of the frequency of reference point violations in Germany. When supervisors are ranked according to the share of reference point violations in their team, the supervisor at the 33% percentile already gives 40% of her managers less than the 100% bonus. The supervisor at the 66% percentile violates reference points of even 80% of her subordinates. But there are also supervisors who do without any reference point violations.²² As our results indicate, this is indeed a good strategy to maintain team performance when percentages are transparent.

²¹ We restrict the model to supervisors who receive either ‘above average’ or ‘fully meets expectations’ ratings in the year of the bonus assignments.

²² For instance, supervisors who have rated at least one subordinate manager as ‘below average’ typically face no necessity at all to pay any of their ‘fully meets expectations’ subordinates less than 100% of their budget even if they also have ‘above average’ managers in their team.

Table 2. Performance Effects of Reference Point Violations

(Dependent Variable: Adjusted Supervisor Performance Rating in Subsequent Year)

No.	Germany					United States				
	D1	D2	D3	D4	D5	US1	US2	US3	US4	US5
Share of Managers with Bonus Percentage <100% among 'fully meets' Managers	-0.572*** [0.217]	-0.676*** [0.220]		-0.689*** [0.217]	-0.671*** [0.221]	0.009 [0.159]	-0.087 [0.162]		-0.087 [0.162]	-0.090 [0.161]
Share of Managers rated 'fully meets' with Bonus Percentage <100% in department			-0.786*** [0.297]					0.006 [0.194]		
Bonus Budget per Manager				-0.014 [0.020]					0.000 [0.022]	
Team Size					-0.020 [0.031]					-0.025 [0.036]
Controls for Rating Distribution	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Observations	468	468	468	468	468	504	504	504	504	504
R-squared (within)	0.05	0.09	0.07	0.09	0.09	0.04	0.08	0.08	0.08	0.08

The regression includes supervisor fixed effects. Standard errors are given in brackets. *, ** and *** denote significance on the 10%, 5% and 1%-level, respectively. All regressions include control variables for the age and hierarchy level of the supervisor. A supervisor's rating distribution is measured by the share of performance ratings 'excellent', 'above average' and 'below average'.

The results of our statistical analysis are strongly supported by the free text comments on the incentive system given by managers at the end of our survey. Many respondents express their concerns that bonus percentages below 100% ‘demotivate’ managers. For instance one respondent wrote “A good (meets expectations) employee should always be evaluated with 100%, not with 80% or 90%. This always leads to discouragement,” and another one expressed that “If 6 people are evaluated in a department and one gets 105% while the others receive 99%, you have discouraged nearly a complete department”.²³

Again, the data on the US subsidiary provide the possibility for a falsification exercise. As bonus percentages are not transparent in the US, bonus percentages below 100% should not be as harmful for future performance. Indeed, we do not find any significant effect of reference point violations in the various US models.

The regression results are robust in alternative specifications. In Model 3, we include the share of ‘fully meets expectations’ managers below 100% of all managers in the department as an alternative measure for the degree of reference point violations. We use this specification to control for the relative importance that the group of potentially discouraged managers has for the whole department or division.²⁴ However, this does not change our results, as the respective coefficient stays negative and highly significant for Germany (Model D3). Also, in order to exclude potential explanations driven by changes in the hierarchical composition of the department, we control for the average bonus budget. The promotion of a team member to a higher salary grade might possibly lower budgetary pressure, because higher salary grades imply higher bonus budgets and may offer the supervisor more freedom to reallocate bonus payments. However, including the budget per manager has no significant impact, and the adverse effect of reference point

²³ Other examples are “Good managers do not care about the difference of some hundreds of Euros associated with an evaluation with ‘fully meets expectations’ below 100%, but are strongly offended by the evaluation and feel like an underperformer.”, “The absolute bonus amount is not an instrument for motivation. [...] A manager who receives a large bonus payment can calculate that his personal target achievement was, for example, 99%. He will immediately realize that he was rated not even as average, and will be discouraged by thousands of Euros.”; “The expectation of every manager who has done a decent job is 100% - every %-point below that is a disappointment.”; “If I assign somebody a bonus below the average, this will not have a positive effect on motivation – no matter how high the payment is.” Comments have been translated from German.

²⁴ The two alternative measures for reference point violation are highly correlated, as the majority of managers receives a ‘fully meets expectations’ rating in both countries.

violations remains robust (Model D4). Finally, variations in the size of the team could also change the restrictions to bonus allocations as well as performance, but controlling for team size also does not affect the results (Model D5). In none of the specifications for the US, we find a significant effect of the variables capturing reference point violations.

One implication of our analysis is that supervisors can achieve a stronger variation in bonus percentages among their managers in the US (as observed in Figure 1), without having to fear that reference point violations reduce satisfaction and performance. For instance, supervisors can assign a higher share of ‘above average’ ratings to their managers without making the associated reference point violations among the ‘fully meets expectations’ managers apparent. In fact, in the US, the share of ‘above average’ is some 11%-points higher, and average bonus percentages of ‘fully meets expectations’ managers are about 3.6%-points lower than in Germany. In departments with supervisors who belong to the quartile with the strongest dispersion in ratings, this difference increases to 4.6 %-points (both differences between countries are significant at $p = 0.001$, two-sided t -tests).

We can summarize our main result in:

Observation 3. *Transparent reference point violations subsequently lead to decreased performance.*

5. Discussion and conclusion

We investigate how bonus payments affect satisfaction and performance of managers in a large, multinational company. The company’s short term incentive system produces a clear, naturally occurring reference point for bonus comparisons. Utilizing a unique data set that combines survey and compensation data, we find that if a manager’s bonus transparently falls behind the reference point, self-reported job satisfaction is significantly reduced. In particular, in line with several behavioral economics approaches to reference-dependent preferences, falling below the reference point has a much stronger effect on satisfaction than being ahead. Moreover, transparent reference point violations reduce

subsequent performance: A supervisor who pushes more managers in her department below the reference point subsequently gets a lower performance rating herself. One implication is that bonus payments are strongly compressed towards the reference point. Both effects are implied by Akerlof and Yellen's fair wage-effort hypothesis.

One potential way to deal with the negative effects of reference point violations is to make it non-transparent. Frank and Sunstein (2001), for instance, note that at the University of Chicago Law School "there is an exceptionally strong norm against public discussion of salaries, even among good friends. The evident basis of the norm is to prevent dissatisfaction about relative position in the face of satisfaction with absolute position." We can confirm the effectiveness of the approach; a lack of transparency makes it more difficult to identify reference point violations. As a likely result, the evidence for reduced job satisfaction, reduced job performance and for the compression of bonuses and evaluations, as we see it in Germany, is much weaker in the US. At the same time, we acknowledge that we cannot be sure that the differences regarding reference-dependent satisfaction and performance are not partly due to cultural differences in preferences between Germany and the US, as opposed to the differences in transparency. However, there are at least two reasons why culture is unlikely to be a complete explanation. As we have seen in Section 3, managers in the US, too, exhibit loss aversion (Model US3); yet, if we control for the information about bonus percentages contained in the ratings, no evidence for reference-dependent effects is left (Models US4 and US5). This indicates that the difference stems from the (non-) transparency of bonus percentages and not from absence of loss aversion in the US. Moreover, laboratory experiments typically find little differences in reference-dependent behavior between Northern Europe and the US, suggesting that reference-dependent preferences are similar (see, for example, Brandts et al., 2004, for a cross-cultural study of cooperation in dilemma games and Oosterbeek et al., 2004, for a meta-study of bargaining experiments).

That said, we think that the role of transparency is worth of further studying. While our field study shows that transparency regarding one's standing relative to the reference point may have detrimental effects, (behavioral) theory may not adequately capture the role of transparency in our setting. Reference-dependent preference models either

(implicitly) assume that the relative position is known or that people rationally form beliefs about their positions. Our results together with other evidence from experimental economics seem to suggest, however, that if one's relative position is not known it is probably less relevant: "What you don't know won't hurt you." (see, e.g., Gehrig et al. 2007 and the references therein). Moreover, this effect is consistent with the social psychology literature, where a series of classic findings demonstrate that social judgments critically depend on which comparison standards are made accessible in the judgmental situation. That is, comparison standards that are not (made) accessible are not used (see Mussweiler and Damisch, 2008, for a review of the more recent relevant cognitive psychology research, and Chen et al., forthcoming, and Benjamin et al., forthcoming, for an application in the economics literature).

On the other hand, a lack of transparency may also have negative consequences, related to procedural fairness perceptions (see, in psychology, Schaubroek et al., 2000, and Werner and Ones, 2000, and in economics, Bolton et al., 2005, and Frey and Stutzer, 2005). One reason is that a lack of transparency bears the risk of inequitable pay rates that would be objected if transparent. Our study for instance shows that a less transparent system may exhibit significantly more reference point violations than a more transparent system.²⁵ Because the role of transparency is not yet well investigated in (behavioral) economics, we do not feel sufficiently confident to conclude from our study that transparency about relative positions is generally not recommendable. However, our evidence clearly suggests that transparency comes with a cost.²⁶

Another avenue for further research is to separate the contribution of different reference-dependent concerns; the managers in our study may compare their bonus percentage relative to their individual bonus target, to their expectations given their ratings, or relative to their peer group, or all three. In fact, we believe that one reason for why we

²⁵ Transparency of pay rates and bonuses is also on the political agenda in the US and in Germany. E.g., in August 2009, Colorado joined California, Michigan and Illinois in protecting 'employees' right to discuss their wages'. The so-called "Wage Transparency Act" prohibits employers from taking adverse actions against employees who discuss their wages with others.

²⁶ Transparency about the relative position seems particularly problematic if the performance measurement process cannot provide objective information justifying the relative ranking. It may be less problematic when objective performance measures are available that credibly explain performance differences between employees.

find so strong reference point effects is that different reference points coincide: falling behind one's target implies falling behind the peer group. However, it would be interesting to know more about the contribution of each of those effects for our results. Understanding the precise behavioral mechanisms that drive behavior is useful for testing and further refining theory.

Finally, to us, one of the main goals of this line of research is not only to get better in understanding the actual effectiveness of bonus systems, but also to be increasingly able to make recommendations for better systems in practice. Surveys of companies indicate that the demand for better systems is large: The consultancy firm Towers Perrin states in its report of a large company survey conducted in 2007 that “more than three-quarters of responding organizations have changed their variable pay programs in the past three years and nearly half expect to implement more changes in the future”.²⁷ The bonus system studied in this paper has also been recently changed, as a reaction to our study, attempting to reduce unwanted comparisons. However, the interaction and effectiveness of commonly used instruments is not well understood yet. A more applied scientific literature on behavioral incentive design would be useful both to shape and enrich economic theory and to give advice to decision makers in practice (see also Roth 2002).

²⁷ The report refers to the 2007 ‘Towers Perrin rewards challenges and changes survey’ including 637 HR and compensation executives from 21 countries.

References

- Abeler, J., Falk, A., Götte, L. and Huffman, D. (forthcoming), “Reference Points and Effort Provision”, *American Economic Review*.
- Akerlof, G. and Yellen, J. (1990), “The fair wage-effort hypothesis and unemployment”, *Quarterly Journal of Economics* 105 (2), 255-283.
- Ariely, D., Gneezy, U., Loewenstein, G. and Mazar, N. (2008), “Large Stakes and Big Mistakes”, *The Review of Economic Studies*, 76, 451-469.
- Benjamin, D. J., Choi, J. J., and Strickland, A. J. (forthcoming), “Social Identity and Preferences”, *American Economic Review*.
- Bewley, T. F. (1999), *Why Wages Don't Fall During a Recession*, Harvard University Press, Cambridge.
- Bolton, G. E. (1991), “A Comparative Model of Bargaining: Theory and Evidence”, *American Economic Review*, 81(5), 1096-1136.
- Bolton, G. E., Brandts, J. and Ockenfels, A. (2005), “Fair Procedures: Evidence from Games Involving Lotteries”, *The Economic Journal*, 115, 1054-1076.
- Bolton, G. E., Greiner, B., and Ockenfels, A. (2009), “Engineering Trust - Reciprocity in the Production of Reputation Information”, Working paper.
- Bolton, G. E., and Ockenfels, A. (2000), “ERC: A Theory of Equity, Reciprocity and Competition”, *American Economic Review*, 90(1), 166-193.
- Bolton, G. E., and Ockenfels, A. (forthcoming), “Risk Taking and Social Comparison. A Comment on ‘Betrayal Aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States’ ”, *American Economic Review*.
- Brandts, J., Saijo, T. and Schram, A. (2004), “How Universal is Behavior? A Four Country Comparison of Spite, Cooperation and Errors in Voluntary Contribution Mechanisms”, *Public Choice*, 119, 381-424.
- Camerer, C., Babcock, L., Loewenstein, G. and Thaler, R. (1997), “Labor supply of New York City cabdrivers: One day at a time”, *Quarterly Journal of Economics* 112 (2), 407–441.
- Chen, Y., Harper, M., Konstan, J. and Li, S. X. (forthcoming), “Social Comparisons and Contributions to Online Communities: A Field Experiment on MovieLens”, *American Economic Review*.
- Clark, A. E. (1999), “Are wages habit-forming? Evidence from micro data”, *Journal of Economic Behavior & Organization*, 39(2), 179-200.
- Clark, A. E., Masclot, D. and Villeval, M.-C. (forthcoming), “Effort and Comparison Income: Experimental and Survey Evidence”, *Industrial and Labor Relations Review*.

- Clark, A. E. and Oswald, A. J. (1996), "Satisfaction and comparison income", *Journal of Public Economics*, 61(3), 359-381.
- Cooper, D. and Kagel, J. (forthcoming), "Other Regarding Preferences: A Selective Survey of Experimental Results" in: J. Kagel and A. Roth, eds., *The Handbook of Experimental Economics vol. 2*, in preparation.
- Demougin, D. and Fluet, C. (2003), "Inequity Aversion in Tournaments", CIRPEE Cahiers de recherche 0322.
- Dufwenberg, M. and Kirchsteiger, G. (2004), "A Theory of Sequential Reciprocity", *Games and Economic Behavior*, 47, 268-298.
- Easterlin, R. (1974), "Does Economic Growth Improve the Human Lot?" in: P. A. David and M. W. Reder, eds., *Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz*, Academic Press, New York.
- Falk, A. and Fischbacher, U. (2006), "A Theory of Reciprocity", *Games and Economic Behavior*, 54, 293-315.
- Fehr, E. and Götte, L. (2007), "Do Workers Work More if Wages are High? Evidence from a Randomized Field Experiment", *American Economic Review*, 97(1), 298-317.
- Fehr, E., Götte, L. and Zehnder, C. (2009), "A Behavioral Account of the Labor Market: The Role of Fairness Concerns", *Annual Review of Economics*, 1, 355-384.
- Fehr, E., Kirchsteiger, G. and Riedl, A. (1993), "Does fairness prevent market clearing? An experimental investigation", *Quarterly Journal of Economics*, 108, 437-460.
- Fehr, E. and Schmidt, K. (1999), "A Theory of Fairness, Competition, and Cooperation", *Quarterly Journal of Economics*, 114, 817-868.
- Fließbach, K., Weber B., Trautner, P., Dohmen, T., Sunde, U., Elger, C. and Falk, A. (2007), "Social comparison affects reward-related brain activity in the human ventral striatum", *Science*, 318 (5854), 1305 – 1308.
- Frank, R. H. (1984), "Are Workers Paid Their Marginal Products?", *American Economic Review*, 74(4), 549-571.
- Frank, R. H. (1985), "The Demand for Unobservable and Other Nonpositional Goods", *American Economic Review*, 75(1), 101-116.
- Frank, R. H. (1997), "The Frame of Reference as a Public Good", *The Economic Journal*, 107, 1832-1847.
- Frank, R. H. and Sunstein, S. (2001), "Cost-Benefit Analysis and Relative Position", *The University of Chicago Law Review*, 68(2), 323-374.
- Frey, B. S. and Meier, S. (2004), "Social Comparisons and Pro-social Behavior - Testing 'Conditional Cooperation' in a Field Experiment", *American Economic Review*, 94, 1717-1722.
- Frey, B. S., and Stutzer, A. (2002), "What Can Economists Learn from Happiness Research?" *Journal of Economic Literature*, XL, 402-435.

- Frey, B. S. and Stutzer, A. (2005), "Beyond outcomes: measuring procedural utility", *Oxford Economic Papers*, 57(1), 90-111.
- Freeman, R. B. (1978), "Job satisfaction as an economic variable", *American Economic Review*, 68(2), 135-141.
- Gächter, S. and Fehr, E. (2002), "Fairness in the Labour Market: a Survey of Experimental Results", in: B. Friedel and M. Lehmann-Waffenschmidt, eds., *Surveys in Experimental Economics. Bargaining, Cooperation and Election Stock Markets*, Physica, Heidelberg and New York.
- Gehrig, T., Güth, W., Levati, V., Levinsky, R., Ockenfels, A., Uske, T. and Weiland, T. (2007), "Buying a pig in a poke: An experimental study of unconditional veto power", *Journal of Economic Psychology*, 28(6), 692-703.
- Gneezy, U. and List, J. A. (2006), "Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets using Field Experiments", *Econometrica*, 74(5), 1365-1384.
- Gneezy, U. and Rustichini, A. (2000), "Pay Enough or Don't Pay At All", *Quarterly Journal of Economics*, 791-810.
- Greiner, B., Ockenfels, A. and Werner, P. (2009), "Social Wage Comparisons and Performance: A Real-Effort Experiment", Working Paper.
- Grund, C. and Sliwka, D. (2005), "Envy and Compassion in Tournaments", *Journal of Economics and Management Strategy*, 14, 187-207.
- Hamermesh, D. (1977), "Economic Aspects of Job Satisfactor", in: W. E. Ashenfelter and A. Oates, eds., *Essays in Labor Market and Population Analysis*, Wiley, New York.
- Hamermesh, D. (2001), "The changing distribution of job satisfaction", *Journal of Human Resources*, 36, 1-30.
- Heath, C., Larrick, R.P. and Wu, G. (1999), "Goals as reference points", *Cognitive Psychology*, 38, 79-109.
- Kahneman, D. and Tversky, A. (1979), "Prospect theory: An analysis of decision under risk", *Econometrica*, 47(2), 263-91.
- Köszegi, B. and Rabin, M. (2006), "A model of reference-dependent preferences", *Quarterly Journal of Economics*, 121(4), 1133-1165.
- Köszegi, B. and Rabin, M. (2007), "Reference-Dependent Risk Attitudes", *American Economic Review*, 97(4), 1047-1073.
- Köszegi, B. and Rabin, M. (2009), "Reference-Dependent Consumption Plans", *American Economic Review*, 99(3), 909-936.
- Kwon, I. and Meyerson Milgrom E. M. (2009), "Status in the Workplace: Evidence from M&A", Working Paper.
- Lazear, P. and Rosen, S. (1981), "Rank-order Tournaments as Optimum Labor Contracts", *Journal of Political Economy*, 89, 841-864.

- Levitt, S. D. and List, J. A. (2007), “What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World?”, *Journal of Economic Perspectives*, 21(2), 153-174.
- Locke, E. A. (1968), “Toward a theory of task performance and incentives”, *Organizational Behavior and Human Performance*, 3, 157–189.
- Locke, E. A., and Latham, G. P. (1990), *A theory of goal setting and task performance*, Englewood Cliffs, NJ: Prentice Hall International.
- Luttmer, E. (2005), “Neighbors as Negatives: Relative Earnings and Well-Being”, *Quarterly Journal of Economics* 120(3), 963–1002.
- Mas, A. (2006), “Pay, reference points, and police performance”, *Quarterly Journal of Economics*, 121(3), 783-821.
- Murphy, K. J. (1991), “Performance measurement and appraisal: Motivating managers to identify and reward performance”, in: W. Bruns, ed., *Performance measurement, evaluation and incentives*, Harvard Business School, Cambridge, MA.
- Murphy, K. R. and Cleveland, J. (1995), *Understanding Performance Appraisal*, Newbury Park, CA: Sage.
- Mussweiler, T. and Damisch, L. (2008), “Going back to Donald - How comparisons shape judgmental priming effects”, *Journal of Personality and Social Psychology*, 95, 1295-1315.
- Oosterbeek, H, Sloof, R. and van de Kuilen, G. (2004), “Cultural Differences in Ultimatum Game Experiments: Evidence from a Meta-Analysis”, *Experimental Economics*, 7(2), 171-188.
- Prendergast, C. (1999), “The provision of incentives in firms”, *Journal of Economic Literature*, 37(1), 7-63.
- Rabin, M. (1993), “Incorporating Fairness into Game Theory and Economics”, *American Economic Review*, 83, 1281-1302.
- Rivas, M. F. and Sutter, M. (2008), “Wage Dispersion and Workers’ Effort”, University of Innsbruck, Faculty of Economics and Statistics Working Paper 15-2008.
- Roth, A. E. (2002), “The Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics”, *Econometrica*, 70(4), 1341-1378.
- Schaubroeck, J., May, D. R. and Brown, F. W. (2000), “Procedural justice explanations and employee reactions to economic hardship: A field experiment”, *Journal of Applied Psychology*, 79(3), 455-460.
- Werner, S. and Ones, D. S. (2000), “Determinants of perceived pay inequities: The effects of comparison other characteristics and pay-system communication”, *Journal of Applied Social Psychology*, 30(6), 1281-1309.

Appendix: Additional Results

Table 3. Determinants of Job Satisfaction – Robustness Checks
(Dependent Variable: Adjusted Satisfaction Score)

Country No.	Germany D6	United States US6
Salary (000s)	0.001 [0.002]	-0.004 [0.002]
$z < 90\%$	-0.335** [0.135]	-0.101 [0.343]
$90\% \leq z < 100\%$	-0.196** [0.089]	-0.071 [0.268]
$100\% < z < 110\%$	0.019 [0.092]	0.091 [0.275]
$z \geq 110\%$	0.079 [0.150]	0.072 [0.325]
Sample	All	All
Rating Dummies	Yes	Yes
No. of observations	2091	957
R-squared (within)	0.04	0.04

Standard errors are given in brackets. *, ** and *** denote significance on the 10%, 5% and 1%-level, respectively. Control variables include supervisor fixed effects, age interval dummies, gender, total years of affiliation to the company and a dummy for promotion in the last year.