Incentives, Sorting, and Prosocial Behavior in Organizations

DISSERTATION

zur Erlangung des Doktorgrades

der Wirtschafts- und Sozialwissenschaftliche Fakultät

der Eberhard Karls Universität Tübingen

vorgelegt von

MICHAEL ROBERT HAYLOCK

aus Carlton, Australien

Tübingen

2022

1. Betreuer:	Prof. Dr. Patrick Kampkötter
2. Betreuer:	Prof. Dr. Valeria Merlo

Tag der mündlichen Prüfung: 01.06.2022

Dekan:	Prof. Dr. Josef Schmid
1. Gutachter:	Prof. Dr. Patrick Kampkötter
2. Gutachter:	Prof. Dr. Valeria Merlo

Acknowledgements

First, I would like to thank my first supervisor Patrick Kampkötter for his unconditional support and discussion of thoughts and ideas, academic guidance, and always being a great source of motivation. I am also very grateful for feedback and guidance from my second supervisor, Valeria Merlo for feedback and giving me presentation opportunities. I also thank Martin Biewen for very helpful feedback on Chapters 2 and 5, and thank Marian Rümmele for helpful feedback on Chapter 2. I would like to thank Dagmar Hegedüs, and all other research assistants for outstanding research assistance throughout my time working on this dissertation.

I have enjoyed working with co-authors Mario Macis, Daniel Wiesen, Michael Kosfeld, Ferdinand von Siemens, and have especially learned a lot from the guidance and support of Robert Slonim. I am also very thankful to the DKMS, especially Susanne Seitz, Jürgen Sauter, Alexander H. Schmidt for being great co-authors to work with.

I gratefully acknowledge the hospitality of the Central European University Budapest, and the University of Melbourne for hosting me for extended research stays. Part of this research was conducted while I visited CEU Budapest and the University of Melbourne. The faculty members at both universities provided helpful support in discussing research ideas at the initial stages of projects. Special thanks to Nisvan Erkal for being a great host at the University of Melbourne.

I am also grateful for the feedback received at talks and conferences, at the Universities of Melbourne, CEU Budapest, Hohenheim, Wuppertal, Hannover, Reading, CINCH Duisbug-Essen, Frankfurt, at the EEA Manchester 2019, VfS Org. Econ. at HU Berlin 2019, and our Brown Bag at Tübingen.

I also dearly thank my parents, Simon Ross Haylock and Susanna Jane Reid, my partner Hanna Christ, and all my friends and dear colleagues at the University of Tübingen, without whose support this dissertation could not have been completed. Patrick Maier, Viktoria Müller, and Stefanie Ehmann were great colleagues and very inspirational. Thanks also to everyone else who provided me feedback and support along the way.

Contents

List of Figures

List of Tables

Chapter 1
Introduction

This dissertation is motivated by a ubiquitous economic problem, namely how to bring about desired actions through an *organization*. But what do we mean by an *organization*? As the Introduction of the Handbook of Organizational Economics quotes Kenneth Arrow, (1974, p. 33), “Organizations are a means of achieving the benefits of collective action in situations where the price system fails,” (Gibbons and Roberts, 2013). This not only includes firms, but also schools, social movements, and beyond (Gibbons and Roberts, 2013). Further, Arrow maintained that organizations are helpful, because of “the need for collective action and the allocation of resources through nonmarket methods.” Where market mechanisms are missing, an organization can be created to coordinate human behavior. An organization in this dissertation can be viewed as a collection of (relational) contracts (Alchian and Demsetz, 1972), or as a hub to reduce transaction costs by centralizing information collection (Coase, 1937). The dissertation contributes to answering important organizational economics questions, such as “How are people rewarded?” (Chapter 2), and the effect of information collection on collective action (Chapter 5), as well as how social norms in firms are determined (Chapter 4) (Gibbons and Roberts, 2013). Chapter 2 uncovers interesting heterogeneity in how executives are rewarded. The bottom line of Chapters 4 and 5 is that *who* (in terms of personal characteristics) is in the organization, and *how* these persons are asked to be involved with the organization, that matters for collective action to succeed.

The dissertation takes a new angle on three specific sub-questions an own chapter. While firms, governments, and the economics literature, have been interested in using *explicit* incentives to induce wanted behavior, these have a limited scope, first, if they reward based on short-term performance (Chapter 2), second, if the desired outcome is not contractible for ethical and legal reasons (Chapter 5), and third, in case no verifiable outcome measure is available (Chapter 4).

Thus, the dissertation first explores the distribution of wage incentives in a panel of executives in Chapter 2. Chapter 3 explores the contemporaneous sorting of employees, in private establishments, by their behaviors, traits and preferences. Here, helping, a proxy for cooperation, is a behavior upon which employees are sorted. Thus, Chapter 4 explores what employee characteristics determine such proxies of employee cooperation, in itself an important non-verifiable outcome, which is also related to firm productivity. Also looking at prosocial behavior in a volunteer context, I ask in Chapter 5 how to motivate a larger group of volunteer stem cell donors to donate in the long run.

Each study uses a different labor market—executives in Chapter 2, private establishments in Chapters 3 and 4, and a not-for-profit organization that has recruited a large number of volunteer donors in Chapter 5, to help partly answer these three important sub-questions. All three sub-questions of the dissertation are answered using unique observational data and modern microeconometric methods.

Chapter 2 is motivated by an empirical anecdote: The Global Financial Crisis of 2007-2008 was partly caused by short-term-oriented compensation practices. Ever since, there have been increased ethical concerns about highly-paid executives, and whether incentives set for executives lead them to behave in the firm’s long-term interest. What has not been answered until now, is whether pay is heterogeneously associated with short-term and long-term firm performance? Thus, the first paper of this dissertation explores whether there is significant heterogeneity in the data, holding unobserved time-constant individual factors constant. This accounts for unobserved baseline differences in pay, and unobserved time-constant differences in the distribution of pay.

Chapter 3 is a research note, and shows that the pattern of worker allocation across firms with respect to behavioral outcomes, such as helping, is consistent with sorting of workers on unobserved establishment characteristics. This sorting is strongest for helping behavior, relative to other observed worker traits. This finding adds to the importance of researching the origins of helping in firms, studied in Chapter 4.

Chapter 4 asks what determines cooperation, namely *helping* and *antisocial* behavior in organizations? These are two proxies for the success of collective action in firms, and thus important economic outcomes. While the first research question of the dissertation focuses on executives, this chapter focuses on all employee levels. Although cooperation has been researched in many settings, there has been to date little *representative* evidence on what drives helping behavior and antisocial behavior in the workplace, using a linked employer-employee panel survey with validated survey items. Thus, this chapter can also contribute to the external validity of the body of research on cooperation in firms. The results of this chapter add to the first chapter by showing that it is *which* workers are in the firm, not necessarily the incentives in place in the organization that determine cooperative behavior. This is an important contribution to the existing literature, showing empirically with representative data that the production of social goods can increase as a result of the *composition* of employees, in terms of their preferences and traits, and the leadership they are subject to.

Chapter 5 explores the behavior of a large volunteer organization of stem cell donors. Stem cells are an important social good that can in many cases only be obtained through an unpaid, and unrelated donor. How donors can best be motivated is a difficult economic question, since traditional monetary incentives are neither ethical nor legal. This chapter expands the definition of the organization to a large donor center *and* its recruited volunteers. I ask whether initiatives from the top of the organization, requesting extra information from donors, promote prosocial behavior that is potentially required at a distant point in time. Assessing which kind of initiatives motivate stem cell donors has not yet been explored, especially with our rich data on over 90,000 donation requests. The results can be generalized to other situations that share similar characteristics, such as volunteering situations with high stakes. The market mechanism has important novel characteristics, too. Importantly, there are very few matching persons from each side of the market, due to the sparsity of genetic characteristics. There is also a high dependence from the patient on the production of the good by few individuals they are matched to. Different to organ donations, stem cell donation can be done repeatedly, and from live donors. Thus, the organization, the donor center, aims to coordinate exchanges that have no monetary reward, and very high stakes consequences if they do not take place. The results show that initiatives run by the donor center with donors, can sort donors on their future availability, and also positively impact the availability of participants, while not reducing the aggregate availability.

Methodically, this dissertation applies modern microeconometric methods that are appropriate to the question at hand. Chapter 2 uses a recently developed quantile regression framework, Method of Moments–Quantile Regression, to account for time-constant unobserved heterogeneity of baseline pay, and of the volatility of pay (Machado and Santos Silva, 2019). Chapters 3 and 4 use modern panel estimation methods appropriate for survey data. Here, particular care is taken to assess the robustness of correlations, and to control for many individual and organizational confounding factors, such as management practices. Chapter 5 uses modern methods traditionally used in policy evaluation, such as intention-to-treat effects, and local average treatment effects, but also quantifies the amount of self-selection that contributes to changes in the outcome of interest. Further, Chapter 5 uses methods to quantify the amount of selection on unobservable characteristics that would lead to a null-effect Oster (2019).

The empirical work in this dissertation is based on a large body of theory and existing empirical results to motivate the hypotheses. This helps to understand potential mechanisms, assess the validity, and make new suggestions for managers. The theory is generally from a large field of behavioral economics, organizational economics, management accounting, and corporate finance research.

Regarding policy relevance, much care is also taken to make results useful for a manager in a firm, who wants to understand the mechanisms at hand. For example, when understanding which preferences correlate with helping behavior, I maintain that while changing the worker composition could potentially lead to a change in helping behavior, it is rather hard to change preferences of already employed workers. While not causal in nature, the correlations between helping (and antisocial behavior) and workforce observables (leadership, social preferences, trust and personality) show that measuring worker traits before hiring and sorting workers based on these can *potentially* improve the collective action in the firm. While personality and leadership skills are potentially assessed to some degree by hiring managers, economic preferences are usually ignored here.

Second, Chapter 2 finds that compensation practices do not excessively benefit the top of the distribution using yearly pay as a proxy, since long-run incentives are stronger at the top of the pay distribution. However, this is reversed when looking at total wealth, indicating that higher pay is associated less strongly with long-run firm performance. Thus, firms could align their yearly pay practices more with what the executive receives in the long run. Examples of stock-related pay practices that could lead to these findings are given, and suggestions are made on what can change this, which are based on theory. However, one must be careful when making any policy recommendations for firms, since this is only an observed pattern, and any new policy could change other factors not accounted for in the analysis. One firm (or country) unilaterally changing (regulation of) incentives can of course not affect the entire market, and may be disadvantageous if it leads to turnover, for example.

Third, when assessing what initiatives run by a large stem cell donor center do, Chapter 5 shows that participation in an initiative is predictive of availability. This is useful for the donor center per se, if they aim to later predict which donors more likely to follow though with donation. Second, the chapter also assesses the average effect of the initiatives on average donor availability, and the effect on the treated. The latter mentioned effects are helpful for the donor center to understand whether the initiatives change the availability of the donor pool. These two contrasting interpretations are thus helpful for different reasons.

Chapter 2

Distributional differences in the time horizon of executive compensation *

2.1 Introduction

The short-term orientation of executive pay is a fundamental shortcoming of compensation practices. Former U.S. Treasury Secretary Geithner stated that

"This financial crisis had many significant causes, but executive compensation practices were a contributing factor. Incentives for short-term gains overwhelmed the checks and balances meant to mitigate against the risk of excess leverage...—Companies should seek to pay top executives in ways that are tightly aligned with the long-term value and soundness of the firm."[2]

A fundamental recommendation made by the Treasury in the same press release is that "compensation should be structured to account for the time horizon of risks." Past crises have shown that if firms do not account for short-term changes to firm performance in their compensation contracts, this can have severe consequences.

Despite recent policy changes, compensation practices are continuing to be heavily criticized in the media.[3] Thus, this paper addresses short-termism in executive compensation. I focus specifically on distributional heterogeneity in the time horizon of performance–pay elasticities using yearly total compensation, and accumulated wealth. The main question I address asks if pay at the top of the conditional distribution is more short-term oriented?

The research question comes from the criticism that executives benefit excessively from short-term changes in firm value (Bebchuk and Fried, 2004; Edmans et al., 2017a). Graham et al. (2005) find that 78% of executives are willing to sacrifce long-term firm value to outperform the market's expectations. I test if the relation of executive pay to short-term and long-term firm and industry performance is heterogeneous across the conditional yearly compensation and total wealth distribution, using the Method of Moments–Quantile Regression (MM–QR) (Machado and Santos Silva, 2019). I also allow for asymmetric response of pay to positive and negative short-term firm performance in a second specification, as in the asymmetric benchmarking literature (Garvey and Milbourn, 2006; Campbell and Thompson, 2015; Daniel et al., 2019).

I employ a panel quantile regression methodology developed by Machado and Santos Silva (2019), to account for endogeneity driven by risk preferences and other latent personality traits, assuming they are time-constant, and executive-firm specific. The strength of the estimator is that it accounts for unobserved average, and *distributional* heterogeneity with executive-firm fixed effects, which is not the case for most other panel quantile regression estimators (Machado and Santos Silva, 2019). It also gives direct inference on the significance of distributional effect heterogeneity, which I use to test the hypotheses.

The literature hitherto has identified short-termism as a problem (Narayanan, 1985; Bebchuk and Stole, 1993; Edmans et al., 2019; Marinovic and Varas, 2019), but not systematically assessed distributional differences in the time horizon of executive compensation. I find significant distributional heterogeneity of short-term and long-term performance–pay relations.[4] Total yearly compensation, a flow measure of pay, is more sensitive to short-term firm performance in the left tail of the conditional distribution, and more sensitive to long-term firm performance in the right tail. By contrast, total wealth, a stock measure of pay, is not always significantly, but quantitatively more sensitive to short-term firm performance in the right tail of the conditional distribution, and more sensitive to

* This chapter is based on Haylock (2022)

[2] https://www.treasury.gov/press-center/press-releases/Pages/tg163.aspx, also cited by, e.g. Peng and Röell (2014)

[3] See, for example, *The Economist*, July 11, 2020, "How CEO pay in America got out of whack".

[4] A cross-section test of distributional heterogeneity of performance–pay has been done before by Hallock et al. (2010)

long-term firm performance in the left tail. This suggests there are weaker incentives to invest in long-term projects for conditionally wealthier executives (Edmans et al., 2017b). Putting this into context, Gopalan et al. (2014) find that firms react to higher stock returns by increasing the duration of compensation, however this can also be to retain talent.

Previous literature has suggested asymmetric benchmarking as a possible driver of managerial skimming, which would be so if the asymmetry is stronger in the right tail of the conditional distribution (Garvey and Milbourn, 2006; Bizjak et al., 2008; Daniel et al., 2019). When allowing for asymmetric short-term performance–pay elasticities, the degree of asymmetry from negative short-term firm performance is very similar across the distribution. This makes asymmetric benchmarking an unlikely mechanism driving differences in conditional pay.

The results of this study show the importance of carefully implementing stock-based pay as an incentive, if its aim is to induce the executive to maximize long-term firm value.

This chapter proceeds as follows. Section 2.2 discusses related literature and develops hypotheses. Section 2.3 describes the data. Section 2.4 discusses the empirical methodology and its application. Section 2.5 presents the results. Section 2.6 discusses the potential mechanisms and potential policy relevance of the findings.

2.2 Related literature and hypotheses

I build an empirical model, including four main variables, to explain executive pay, the dependent variable. In a similar framework, Hallock et al. (2010) find distributional differences in performance–pay elasticities for CEOs, using conditional quantile regressions (Koenker and Bassett, 1978), ranging from 0.07 at the first decile, to 0.15 at the 9th decile. Following Hallock et al. (2010), my empirical model allows for heterogeneous performance–pay relations across the conditional distribution of compensation.

Firms optimally pay executives to maximize long-term firm value, which captures all relevant outcomes of executives' behavior, e.g. changes in growth or profit, and restructuring (Jensen, 2001; Edmans et al., 2012, 2017b). Managers may not act in the best interest of the firm, and aim to increase own wages or reputation with short-term oriented action (Narayanan, 1985). Bebchuk and Stole (1993) argue that asymmetric information between managers and shareholders can lead to sub-optimal investment. Examples of short-term behaviors are forgoing positive-NPV projects that sacrifice short-term performance but would raise long-term performance, undertaking negative-NPV projects that boost short-term performance but sacrifice long-term performance, M&A announcements, and stock repurchases with free cash (Edmans et al., 2017b, 2019). Bizjak et al. (1993) and Cadman et al. (2013) show that long-term equity is used more frequently in sectors in which short-term performance is a noisy predictor of long-term performance (Peng and Röell, 2014).

The distributional model also tests for differences in short-term and long-term firm performance–pay elasticities. I include short-term firm value to model managerial actions that have a short-term effect on firm value (Narayanan, 1985; Bebchuk and Stole, 1993), which also captures luck and productivity changes. I argue that industry and macroeconomic controls capture most other factors influencing firm value over the business cycle. While Gopalan et al. (2014) directly measure the duration of executive pay with a weighted average of vesting periods for pay components, I complement this by assessing how different pay measures correlate with long-term and short-term stock performance. It is possible that higher short-run performance also increases the value of long-term pay, which is an unintended consequence of this kind of pay, if executives can cash in on short-run changes to their equity holdings once they have vested. Supporting this conjecture, Edmans et al. (2019) provide empirical evidence that vesting equity provides an incentive for executives to engage in behavior that sacrifices long-term firm value.

I test whether incentives for short-term and long-term firm-performance differ across the distribution. If the null hypotheses below are not rejected, this would suggest that short-termism is a greater problem when pay is (conditionally) greater. This would support the arguments above.

Hypothesis 1 The short-term firm performance–pay relation is positive and increases with the conditional pay quantile.

Hypothesis 2 The long-term firm performance–pay relation is positive and decreases with the conditional pay quantile.

In an optimal contract, an executive's variable compensation is positively correlated with firm performance, and exogenous measures that are also correlated with firm performance are used to filter luck (Holmstrom, 1979; Edmans et al., 2012; Edmans and Gabaix, 2016). Other studies show mixed evidence on relative evaluation (Frydman and Jenter, 2010), and that managers are rewarded positively for external forces affecting firm performance (Bertrand and Mullainathan, 2001).

Here, I re-explore the hypothesis that managers in the lower-tail of the distribution are more likely to be benchmarked against the industry, and receive higher pay when economic conditions are good (Bizjak et al., 2008). If firms adjust pay upwards when the industry is performing better, holding own firm performance constant, this is often used to retain executives (Bizjak et al., 2008; Campbell and Thompson, 2015). This can be explained in equilibrium by the manager's outside option increasing if other firms link their manager's pay to their firm performance. I include long-term and short-term average industry shareholder wealth as the two other variables of interest. I also control for macroeconomic indicators, which are other exogenous factors potentially correlated with firm performance.

I test whether benchmarking against long and short-term industry performance is distributionally heterogeneous. If Hypotheses 3 and 4 are supported, they support the findings of Bizjak et al. (2008).

Hypothesis 3 The short-term industry performance–pay relation is positive and decreases in the conditional pay quantile.

Hypothesis 4 The long-term industry performance–pay relation is positive and decreases in the conditional pay quantile.

Garvey and Milbourn (2006) estimate that for a CEO at the mean, the performance–pay relation is between 25% and 45% lower when firm performance changes due to luck is negative, than when it is positive. Addressing a potential mechanism, Bizjak et al. (2008) find that asymmetric benchmarking of yearly compensation is used to retain CEOs, and is not strongly associated with poor corporate governance. If a firm engages in such behavior, a CEO can threaten to leave. In a comprehensive study testing robustness of this asymmetry, Daniel et al. (2019) find no significant interaction between bad luck and the pay benchmark in the majority of specifications for US firms.

I re-explore this mechanism and test for distributional heterogeneity of the short-term firm performance–pay asymmetry, from the perspective of when pay is granted, using total yearly compensation. This leads to the fifth and sixth hypotheses:

Hypothesis 5 Total compensation is more sensitive to positive than to negative short-term changes in firm value.

Hypothesis 6 The reduction to the short-term firm performance–pay relation, when performance is negative, increases with the conditional total compensation quantile.

If these hypotheses are both supported, especially Hypothesis 6, then asymmetric performance benchmarking is one possible reason for higher conditional pay. If only Hypothesis 5 is supported, and there is no evidence for distributional heterogeneity, then asymmetric benchmarking is not a driver of managerial skimming.

2.3 Data

I use compensation data from a 11-year unbalanced panel of executives in the C-suite of publicly listed firms for 34 countries over the years 2003-2013, provided by BoardEx. Most observations come from the USA, UK, Western Europe and Scandinavia. The unit of observation is the pay of an executive i, in a firm f, in an industry s, at year t. There are 143 executives who switch firms within the observation period in the final sample, which can be calculated by the difference in the 6,939 executive-firm matches, and the 6,796 of executives in the data altogether.[5] Matching firm financial data is from the ORBIS database provided by Bureau van Dijk. Executives are included in the final sample if there is at least one non-missing observation of total compensation.

The main dependent variables of the study are total compensation and total wealth. Total compensation consists of salary and cash bonus plus the grant-date value of newly emitted equity-linked compensation (such as stock options, restricted stock awards), and long-term incentive plans (restricted bonuses) awarded in each year, as used by Fernandes et al. (2013). Total compensation measures the grant-date opportunity cost to the shareholders of the executive's pay package (Fernandes et al., 2013). Executives accumulate stock holdings and other equity-linked pay. Firm performance can affect executive utility through wealth to a larger degree than yearly total compensation (Frydman and Jenter, 2010). Edmans and Gabaix (2016) show that incentives for executives are larger when using total wealth to proxy utility. I use total wealth as an outcome variable, which is the sum of estimated market value of and executive's cumulative holdings of stock-related pay, in-the-money options, and long-term incentive plans for an executive (Fernandes et al., 2013).[6]

Firm value is a widely used proxy of firm performance and managerial effort (Jensen, 2001; Edmans et al., 2017b). Firms' market value of equity at year end is used to generate the main independent variables (Jensen and Murphy, 1990; Bertrand and Mullainathan, 2001).[7] I control for macroeconomic indicators, GDP per capita, GDP-growth in percent, provided by the World Bank and inflation as the percentage change in average consumer prices, provided by the International Monetary Fund. These serve to control for time-variant country-level heterogeneity, although there may still be residual variance not captured here. Controls for age and age squared of an executive are included, in line with Bertrand and Mullainathan (2001). Since the estimator accounts for executive-firm fixed effects and age, additionally including tenure would be collinear. Executive-firm fixed effects, discussed below, control for unobserved average and distributional differences in wealth or total compensation that are time-constant.

Although other studies control for firm size (Murphy, 1999; Garvey and Milbourn, 2006), growth potential using the market-to-book ratio (DeVaro et al., 2017), and leverage (DeVaro et al., 2017), I have chosen explicitly not to include these controls, as they can likely cause biased estimates of coefficients. Including control variables that are simultaneously determined with the outcome variable of interest by the independent variables leads to this bias (Angrist and Pischke, 2008; Swanquist and Whited, Forthcoming).[8]

2.4 Measuring short-term and long-term performance

Short-term and long-term firm and industry performance is identified using the band-pass filter (Christiano and Fitzgerald, 2003), and is based on the theory of business cycles (Burns and Mitchell, 1947). This separates the performance, proxied by market value of a firm f or industry s at time t into a trend component $Trend$, and a cyclical component, $Shock$, so that

[5] 50% of executives are observed for at least 3 years in the data, 25% are observed at least 5 years, and 5% at least 8 years. The probability of an executive remaining in the sample for the following period is 0.78 on average. Most years have an attrition rate between 0.14 and 0.2. However, there is no general time pattern in the attrition rate.

[6] The calculation of total wealth for each executive, done by BoardEx, is explained in more detail in Appendix section 2.8 Note that options exercised and stock sold by an executive disappear from total wealth.

[7] This variable is astk_market_cap in ORBIS.

[8] For example, if performance affects the firm's market-to-book ratio, the relation between firm performance on compensation is not identified. The resulting bias can be shown in simultaneous equation system:

$$\begin{aligned} Pay &= \alpha + \rho \cdot Performance + \gamma \cdot GrowthPotential + e_i \\ GrowthPotential &= \lambda_0 + \lambda_1 \cdot Performance + u_i \\ Pay &= (\alpha + \gamma\lambda_0) + (\rho + \gamma\lambda_1) Performance + \gamma u_i + e_i . \end{aligned}$$

$$Market\ Value_{f\,t} = Shock_{f\,t} + Trend_{f\,t}. \tag{2.1}$$

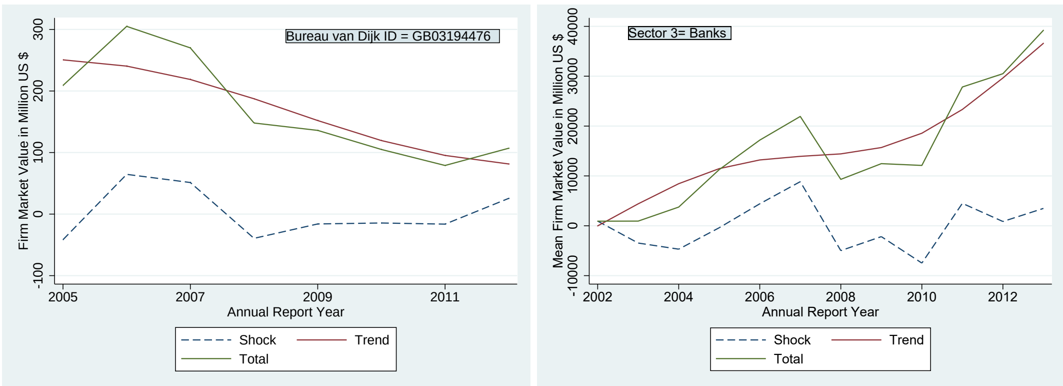
This filter is also used to identify the effects of firm shocks on executive compensation in a different setting done by DeVaro et al. (2017). I apply it to the time series of the market value of each firm to generate the shocks of year end market value for each firm, and to the time series of the mean year-end shareholder wealth of the industry. The band-pass filter is explained in more detail in Appendix section 2.9.

I separate stochastic cycles from the trend that range from two to eight years, as data are yearly. This filtering method is in accordance with Burns and Mitchell (1947), who define business cycles as stochastic cycles in business data between 1.5 and 8 years. The time period to measure short-term changes in firm value is also in line with compensation practices, where yearly bonuses are 'short-term', and pay withheld for longer than one year, normally 3 to 5 years, is 'long-term'. Long-term pay aims to remove such productivity cycles from compensation plans. Figure 4.1 shows the application of the filter to a firm, and an industry, from my sample. The figures show that it works to identify stochastic changes in performance, with mean zero. The shock and trend for each firm, $Shock_{f\,t}$ and $Trend_{f\,t}$, and each industry, $Shock_{st}$ and $Trend_{st}$, serve as the proxies for short-term, and long-term firm and industry performance.

As short-term performance is measured by *yearly* cycles of firm value, it is not driven entirely by exogenous factors to the firm, such as luck (DeVaro et al., 2017), but also captures factors endogenous to the firm, such as short-term oriented behavior (Edmans et al., 2019), and yearly productivity changes. Productivity is likely to sink if managers put more attention on public relations, and actions that focus more on short-term stock price manipulation, than on operations (Peng and Röell, 2014). The shock variable measures "short-term performance", and the trend component measures "long-term performance".

Table 2.1 describes the main variables of interest, using all observations from the final sample. An advantage of the data set is to be able to track executives over a long time frame. In the final sample, I have data from 34 countries, adding to the generality of the findings to countries outside the U.S., which has been the central focus in the literature hitherto.

Fig. 2.1: Short-term and long-term firm and industry performance



Business cycles of firm with BvD ID GB03194476, and of the banking sector, showing the time series of market value, the cyclical and trend components from the band pass filter, removing cycles from 2 to 8 years and accounting for drift.

2.5 Application of Method of Moments-Quantile Regression

I aim to tackle numerous endogeneity concerns in my empirical application. Past research shows that the level of executive pay can be driven by selection of more talented managers into contracts with higher pay, through assortative

Table 2.1: Summary Statistics

Variable	N	Mean	S.D.	Min.	Max.
Age	25,636	51.7	8.1	25	92
Salary*	25,502	533.7	618.5	1	61,529.5
Bonus*	21,686	470.1	1,373.3	0	123,939.1
Equity-linked*	13,956	2,912.9	23,547.5	0	1,527,364.1
Variable compensation*	25,502	1,988.2	17,577.5	0	1,528,892.3
Total compensation*	25,636	2,514.3	17,614.3	1	1,530,183.7
Total wealth*	22,075	32,667.3	452,124.9	1	26,708,263
Market value of equity**	25,636	7,573	20,420.4	0	288,579
$Shock_f$***	25,636	0.1	3.8	-51	81
$Trend_f$***	25,636	7.5	19.8	0	266.9
$Shock_s$***	25,636	0.1	2.8	-29.3	24.9
$Trend_s$***	25,636	6.8	6.9	0.1	75.5
GDP growth	25,636	1	2.5	-8.3	15.2
GDP per-capita	25,636	42,785.1	7,348.3	1,724.7	115,109.3
Inflation	25,636	2.6	1.1	-1.7	14.1

*** scaled in Billions $US, ** scaled in millions $US, * scaled in thousands $US. The sum of salary, bonus, grant-date value of newly emitted equity-linked and long-term incentive plans, equals total compensation. Observations with zero salary or total compensation are dropped from the data. The sample includes executives for which there is at least one non-missing observation of total compensation. If an executive works in two firms at the same time, plausibility checks were done and some observations dropped according to the following criteria. If data is entirely missing, this observation is deleted. If there is a holding company or a subsidiary, and the executive had the same position at both firms, the observation belonging to the parent company is kept. If one position was only a representative or deputy position, this observation is deleted. If the firm is listed in two countries, the headquarter country is kept. If the executive switched positions and thus worked for two companies in one year, the first year of the new job is deleted, as this generally covers fewer months. Firm and industry shock and trend variables are generated using the band-pass filter, removing drift and cycles between 2 and 8 years from the raw data to generate the trend. GDP growth and GDP per-capita are from the World Bank, and Inflation is measured as the average percentage change in consumer prices in each country, which is from the IMF database.

matching (Gabaix and Landier, 2008; Tervio, 2008). More talented managers likely cause better firm performance as well. Estimating the conditional variance of pay is potentially confounded by managers' risk preferences. Confident managers are more likely to undertake in risky investments with free cash (Malmendier and Tate, 2005), close M&A deals (Malmendier and Tate, 2008), and hold more of own-company stocks (Malmendier and Tate, 2005). Thus, latent preferences and personality traits, such as risk tolerance and confidence, likely cause more volatile firm performance and pay, and even structurally different portfolios. I assume here, that these preferences and traits are more or less time-constant (Cobb-Clark and Schurer, 2012; Bernile et al., 2017; Schildberg-Hörisch, 2018). These endogeneity concerns are addressed by the MM-QR estimator, outlined in this chapter, which was developed by Machado and Santos Silva (2019).

The MM-QR uses estimates of conditional mean and the conditional scale function to estimate regression quantiles (Machado and Santos Silva, 2019). This makes it computationally easy to estimate a model with a large number of individual-specific fixed effects in a quantile regression framework. I estimate around 6,939 executive-firm fixed effects. I identify the response of the compensation Y_{ifst}, of an executive i, in firm f, in industry s, at time t, to performance variables that vary at the firm level, or the industry level to which the firm belongs. These measures are summarized for now by X of firm f, in industry s, at time t, as defined by X_{fst} for firm-level variables, and X_{st} for industry-level variables, written together as $X_{(f)st}$. Firm and industry performance measures, and all control variables are summarized for now under $X_{(f)st}$. The response of compensation to performance, summarized by the coefficient vector β, is allowed to depend on the position of pay in the conditional distribution, and is clustered at the executive-firm level, which is modeled by unobserved noise U_{ifst} distributed on the uniform interval $[0, 1]$. This is in order to estimate

$$\log Y_{ifst} = X'_{(f)st}\beta(U_{ifst}),\ i = 1, \ldots, n. \tag{2.2}$$

However, standard quantile regression methods do not deal with the panel nature of the data. This poses a problem for identification of β, if there is time-constant unobserved heterogeneity at the individual level, affecting both firm performance and compensation. Accounting for this by including individual intercepts, α_{ifs}, estimates the pay relation at the τ'th quantile as

$$Q_{\log Y_{ifst}}(\tau|X_{(f)st}) = \alpha_{ifs} + X'_{(f)st}\beta q(\tau), \tag{2.3}$$

where $q(\tau) = F_U^{-1}(\tau)$. However, including a large number of individual specific intercepts in the quantile regression is computationally burdensome. Further, variance estimates of other covariates may be increasingly large in proportion to the amount of fixed effects (Koenker, 2004). This is especially problematic if the panel is small, since standard errors for individual effects will be large.[9]

A second potential source of unobserved heterogeneity is in the conditional variance of pay. If the conditional variance of pay depends on time-constant unobserved factors, not accounting for these can bias estimates of the conditional variance, if they are correlated with independent and dependent variables. The method applied here accounts for unobserved differences in the conditional mean and the conditional variance of pay, using a location-scale model developed by Machado and Santos Silva (2019).

I assume in the analysis that the location and scale functions are known, to specify the empirical model of the relation between pay and performance with control variables as

$$\log Y_{ifst} = \alpha_{ifs} + X'_{(f)st}\beta + \sigma(\delta_{ifs} + X_{(f)st}\gamma)\varepsilon_{ifst} \tag{2.4}$$

where σ is the scale function. I assume the scale function to be linear in covariates. Here, regressors may only affect the distribution of the response variable through known location and scale functions (Koenker and Bassett, 1982). However, heteroskedasticity may not be linear, but can be multiplicative (Godfrey, 1978; Koenker and Bassett, 1982). In this case, the scale-shift at a quantile, q, is not linear in covariates but quadratic in covariates (Koenker, 2005).[10] Thus, results should be taken with some caution, as they do not account for second or higher order moments of performance. Most studies of performance–pay do not include polynomials of performance. I do not include polynomials to be in line with the literature, and keep results comparable. I estimate

$$\hat{Q}_{\log Y_{ifst}}(\tau|X_{(f)st}) = (\hat{\alpha}_{ifs} + \hat{\delta}_{ifs}\hat{q}(\tau)) + X'_{(f)st}(\hat{\beta} + \hat{\gamma}\hat{q}(\tau)). \tag{2.5}$$

The point estimate of the coefficient of interest l, at the τ'th quantile is

$$\hat{\beta}_l(\tau, X_{(f)st}) = \hat{\beta}_l + \hat{q}(\tau)\hat{\gamma}. \tag{2.6}$$

The scale parameter $\hat{\gamma}$ estimates the distributional heterogeneity.

In the estimation procedure, main variables are in logarithmic form in estimations below, but logarithmic notation is omitted here for brevity. The average estimated coefficients $\hat{\beta}$ in the MM-QR procedure are obtained by using OLS of time-demeaned independent and dependent variables, regressing $(Y_{ifst} - {}^{\sum_t Y_{ifst}}/_T)$ on $(X_{(f)st} - {}^{\sum_t X_{(f)st}}/_T)$.

[9] To account for a location shift, which is independent of the quantile estimated, Koenker (2004) uses l_1 shrinkage methods to control for the large number of fixed effects. Other methods to account for location shifts have been developed by Lamarche (2010) and Canay (2011).

[10] A model where a single explanatory variable has a quadratic effect on the scale of the conditional distribution is, for example, $y_i = \beta_0 + \beta_1 x_i + (1 + x_i)^2 u_i$ (Koenker, 2005).

Then, the location shift, which is the standard fixed-effect from a within regression, $\hat{\alpha}_{ifs}$, is predicted from the above estimation of $\hat{\beta}$, $\hat{\alpha}_{ifs} = \frac{1}{T}\sum_t (Y_{ifst} - X'_{(f)st}\hat{\beta})$. The residuals are $\hat{R}_{ifst} = Y_{ifst} - \hat{\alpha}_{ifs} - X'_{(f)st}\hat{\beta}$. The scale parameter, $\hat{\gamma}$, is estimated by regressing the time-demeaned absolute value of residuals $(|\hat{R}_{ifst}| - \sum_t |\hat{R}_{ifst}|/T)$ on $X_{(f)st}$.[11] The part of conditional variance that is time-constant and unobserved is estimated by $\hat{\delta}_{ifs} = \frac{1}{T}\sum_t (|\hat{R}_{ifst}| - X'_{(f)st}\hat{\gamma})$. The quantile $q(\tau)$ is then estimated by

$$\min_q \sum_i \sum_t \rho_\tau \left(\hat{R}_{ift} - \left(\hat{\delta}_{if} + X'_{f,st}\hat{\gamma}\right) q\right)$$

to obtain estimates of quantiles $\hat{q}(\tau)$ in the data, where ρ is the check-function (Machado and Santos Silva, 2019).

As the estimation procedure above shows, parameter estimates $\hat{\gamma}$ are amended of the executive-firm fixed effects $\hat{\alpha}_{ifs} + \hat{q}(\tau)\hat{\delta}_{ifs}$. Interpretations of point estimates at a quantile $\hat{q}(\tau)$ do not depend on time-constant individual characteristics, such as talent or risk preferences. This is an advantage of the estimation procedure not accounted for by most other quantile regression methodologies, and simple to implement. Further, standard errors are clustered via bootstrapping, to account for serial correlation of compensation.

One potential problem of the empirical application is fixed-T asymptotic biases of the estimated scale parameter $\hat{\gamma}$ and quantile $\hat{q}(\tau)$ when the number of individuals n relative to the panel length T, n/T, is large (Machado and Santos Silva, 2019). This is because the MM-QR estimator assumes that asymptotically, the number of individuals is small compared to the panel length, or $(n, T) \to \infty$, $n = o(T)$. Parameters of average effects, however, remain consistent in short panels with large n. Machado and Santos Silva (2019) show in Theorem 4 that it is possible to remove the bias by using a jackknife. I therefore bias-correct point estimates in main results using the split-panel jackknife method (Dhaene and Jochmans, 2015).[12]

This method estimates two scale parameters of half panels, by splitting total executive-year observations N into odd and even years, N_{odd} and N_{even}. The half-panels have the same (or very similar) number of individuals n as the full panel, but only half as many time periods, allowing us to identify the bias from having a small T by using the sample size weighted differences of full-panel and half-panel estimates. MM-QR estimations are run on both half-panels separately, and, assuming the amount of bias is proportionate to the number of observations N, the scale parameter from MM-QR, $\hat{\gamma}_{MM-QR}$, is corrected accordingly. The corrected scale parameter is $\hat{\gamma}_{JK} = 2\hat{\gamma}_{MM-QR} - \hat{\gamma}_{odd}\frac{N_{odd}}{N} - \hat{\gamma}_{even}\frac{N_{even}}{N}$. The estimated quantiles, $\hat{q}(\tau)$, are also bias-corrected analogously to $\hat{q}_{JK}(\tau) = 2\hat{q}_{MM-QR}(\tau) - \hat{q}_{odd}(\tau)\frac{N_{odd}}{N} - \hat{q}_{even}(\tau)\frac{N_{even}}{N}$. Thus, if the scale parameter or quantile is over-estimated (under-estimated) when the panel becomes shorter, it is corrected downward (upward). The bias-corrected point estimates of the coefficient of interest l, at the τ'th quantile are

$$\hat{\beta}_l^{JK}(\tau, X_{(f)st}) = \hat{\beta}_l + \hat{q}_{JK}(\tau)\hat{\gamma}_{JK}. \tag{2.7}$$

2.6 Results

2.6.1 Total compensation

Testing the Hypotheses 1 to 4, I estimate regression quantiles using the MM-QR. The dependent variable is the logarithm of 1+total compensation, $\log Y_{ifst}$, for an executive i, in firm f, in industry s, at time t. Performance measures

[11] One may use an alternative transformation of residuals that has mean 0 conditional on X.

[12] Machado and Santos Silva (2019) show that the bias-corrected estimator performs far better than the uncorrected estimator for different panel lengths, and works reasonably well for a panel length of 10.

are the logarithm of 1+transformed short-term firm and industry performance, log $Shock_{ft}$ and log $Shock_{st}$,[13] the logarithm of 1+long-term firm and industry performance, log $Trend_{ft}$ and log $Trend_{st}$, macroeconomic controls Z_{ft} outlined in the Data section, and year indicators.[14]

I estimate the performance–pay elasticity with a log-log model in the main specification. Here, one assumes that managerial actions affect firm value proportionately to firm size, and that resulting bonuses relate proportionately to firm value. This is especially realistic if pay is also equity-linked, which is the case here (Edmans et al., 2017b). For example, a corporate restructure will likely increase the %-performance of the firm. On the other hand, perquisites, such as buying a private jet, may only reduce $-performance. In the estimation, the fixed effect that is identified is an executive-firm pair fixed effect, as outlined above in the empirical methodology. If an executive switches firms, another fixed effect is estimated. This accounts for unobserved, time-constant heterogeneity of executive-firm matches in the average level, and conditional variance of pay. I estimate

$$\hat{Q}_{\log Y_{ifst}}(\tau|\cdot) = (\hat{\alpha}_{ifs} + \hat{\delta}_{ifs}\hat{q}(\tau)) + (\log Shock_{ft} + \log Shock_{st} + \log Trend_{ft} + \log Trend_{st} + Z'_{ft} + \psi_t)(\hat{\beta} + \hat{\gamma}\hat{q}(\tau)) \quad (2.8)$$

using an unbalanced panel, after winsorizing the sample at the 1st and 99th percentiles.[15] Summary statistics of the winsorized sample are in Table 2.5 in section 2.10 of the Appendix. Results of the same estimation for the unwinsorized sample are shown in Table 2.6 in section 2.10 of the Appendix. I deal with serial correlation of pay by clustering standard errors via bootstrap (Parente and Silva, 2016). Even if there is intra-cluster correlation, estimates of quantile regression are also consistent under certain conditions (Parente and Silva, 2016). I resample from the regression sample by firm-executive cluster, with 200 replications. Location, scale and point estimates of coefficients of interest at the 10th, 30th, 50th, 70th, and 90th conditional quantiles are reported in Panel A of Table 2.2. In panel B, I show results from the split-panel jackknife bias correction (Dhaene and Jochmans, 2015).

Hypothesis 1 asks if the short-term firm performance–pay relation is positive and increases with the conditional pay quantile. Testing Hypothesis 1 in Table 2.2, the point estimates of interest belong to the variable Log $Shock_f$. Evidence from both the winsorized and unwinsorized sample (shown in Table 2.6 in section 2.10 of the Appendix) points in the same direction. Both location and scale parameters are estimated precisely in columns one and two. The location parameters reported in column 1 are from a standard fixed-effects estimator. The location function shows that one average, a 1% increase in short-term firm value is associated with 0.04% more total compensation. This elasticity is also quantitatively similar at the conditional median.

Now turning to the heterogeneity of the short-term firm performance–pay relation, the scale parameter is negative, showing that conditionally higher paid managers have a lower short-term firm performance–pay sensitivity, which decreases from 0.07 to 0.02, from the 10th to the 90th conditional quantiles. In panel B, the bias-corrected results find larger decreases in the elasticity, ranging from 0.08 at the 10th percentile to 0 at the 90th percentile. This rejects Hypothesis 1 using total compensation as a measure of pay.

In the unwinsorized sample, in Table 2.6 in section 2.10 of the Appendix, short-term firm performance–pay sensitivities are larger overall, which is to be expected, since compensation data are very right-skewed. The distributional heterogeneity is also significant, is qualitatively the same, and quantitatively larger, which is in line with the results from the winsorized regressions.

[13] Log $(1 + (Shock_{ft} + x))$, where x is the smallest number such that all values are non-negative), so Log $Shock_{ft}$ and Log $Shock_{st}$ are the logarithm of transformed variables in the following.

[14] Alternately, one could use time-by-country fixed effects instead of macroeconomic controls, but this made the estimation procedure burdensome with about 400 dummy variables and bootstrapping.

[15] It is doubtful that winsorizing is a sensible practice, especially when using quantile regression. Winsorizing is done for main results to be comparable with those of the previous literature, and all main results are additionally replicated with the unwinsorized sample.

Table 2.2: Winsorized MM-QR of log total compensation on log of firm and industry performance measures

	(1) Location	(2) Scale	(3) Q10	(4) Q30	(5) Q50	(6) Q70	(7) Q90
A: MM-QR							
Log $Shock_f$	0.0441***	-0.0189***	0.0733***	0.0593***	0.0431***	0.0290***	0.0162**
	(0.0117)	(0.0061)	(0.0191)	(0.0155)	(0.0121)	(0.0083)	(0.0078)
Log $Trend_f$	0.2073***	0.0142*	0.1853***	0.1958***	0.2081***	0.2187***	0.2283***
	(0.0161)	(0.0081)	(0.0208)	(0.0164)	(0.0167)	(0.0162)	(0.0193)
Log $Shock_s$	0.0128	-0.0079	0.0250	0.0192	0.0124	0.0065	0.0012
	(0.0157)	(0.0073)	(0.0230)	(0.0190)	(0.0159)	(0.0143)	(0.0150)
Log $Trend_s$	-0.0220	-0.0167	0.0038	-0.0085	-0.0229	-0.0353	-0.0467
	(0.0246)	(0.0124)	(0.0334)	(0.0261)	(0.0256)	(0.0251)	(0.0308)
B: Jackknife							
Log $Shock_f$	0.0441***	-0.0258***	0.0872***	0.0610***	0.0422***	0.0274***	0.0035
	(0.0117)	(0.0061)	(0.0191)	(0.0155)	(0.0121)	(0.0083)	(0.0078)
Log $Trend_f$	0.2073***	0.0185**	0.1764***	0.1952***	0.2087***	0.2193***	0.2365***
	(0.0161)	(0.0081)	(0.0208)	(0.0164)	(0.0167)	(0.0162)	(0.0193)
Log $Shock_s$	0.0128	-0.0137*	0.0357	0.0218	0.0118	0.0039	-0.0088
	(0.0157)	(0.0073)	(0.0230)	(0.0190)	(0.0159)	(0.0143)	(0.0150)
Log $Trend_s$	-0.0220	-0.0127	-0.0007	-0.0137	-0.0229	-0.0303	-0.0421
	(0.0246)	(0.0124)	(0.0334)	(0.0261)	(0.0256)	(0.0251)	(0.0308)
N	25,636	25,636	25,636	25,636	25,636	25,636	25,636

Results show Method of Moments-Quantile Regressions (Panel A) of log(1+total compensation) on transformed firm and industry short-term performance, $\log(1 + (Shock + x))$, where x is the smallest number such that all values are non-negative, and $\log(Trend + 1)$ for firm and industry long-term performance, with executive-firm fixed effects. Panel B shows bias-corrected estimates of coefficients estimated in Panel A. This is done using the split-panel jackknife bias correction by Dhaene and Jochmans (2015), where the panel is split into even and odd years. Standard errors clustered at executive-firm level via bootstrap (200 reps) are in parentheses. *** $p < .01$, ** $p < .05$, * $p < .1$. Full set of controls: Age, age squared, GDP, GDP growth, inflation, and year dummies used. Controls not reported. Data set used is the full unbalanced panel from BoardEx executive compensation database, ORBIS for market value variables, and span years 2003-2013 and 34 countries. GDP growth and GDP per capita are from the World Bank, and inflation is from the IMF database. Observations with zero total compensation or base salary are removed. Compensation, shock and trend variables are winsorized at the 1st and 99th percentiles.

Results in Tables 2.9 and 2.12 in section 2.10 of the Appendix show that distributional heterogeneity of the short-term performance pay elasticity is robust to a log-level model, for both winsorized and unwinsorized data. Shock values in the log-level model are not scaled by firm size, so the pay-performance sensitivity is a %-$ relation. The results imply that short-term performance–pay elasticity is smaller in the right tail of the conditional distribution. They are against Hypothesis 1, from the perspective of granting compensation, as total compensation is a flow measure. The findings are naturally dependent on this specific measure of pay.

Hypothesis 2 explores if the long-term firm-performance pay sensitivity is positive and decreases with the conditional pay quantile. The coefficient of interest is Log $Trend_f$ in Table 2.2. The elasticity of total compensation to long-term firm performance is 0.21 at the median (column 5). Both location and scale parameters are estimated moderately precisely, and the positive scale parameter shows that predicted earnings respond more to long-term changes in performance at the right tail. The elasticity is about 23% higher at the 90th percentile than the 10th percentile, and the difference is statistically significant at the 10% level. This evidence rejects the premise of

Hypothesis 2, in which the sensitivity of pay to performance is decreasing in the conditional total compensation quantile.

Regarding the robustness of results, the bias-corrected results in panel B are quantitatively similar. The unwinsorized regressions in Table 2.6 in section 2.10 of the Appendix reveal, however, no significant heterogeneity across the conditional distribution. The log-level specification in Tables 2.9 and 2.12 reveals significant heterogeneity in the same direction as the winsorized log-log specification. The long-term firm performance-total compensation relation is increasing in the conditional quantile in three out of the four tested specifications, and in no case does it go in the opposite direction. The evidence is in line with the notion that higher conditional total compensation is more strongly benchmarked against long-term firm value. It is also worth noting that short-term performance–pay elasticities are smaller than long-term elasticities across the distribution in the winsorized sample, but this reverses in the raw data.

These results speak somewhat against the interpretation that greater conditional total compensation results from short-term managerial actions or managerial skimming, when compensation is granted (Edmans et al., 2019). A potential explanation of the short-term performance–pay relation is that pay is associated with firms' liquidity. For example, in the financial crisis, firms also cut bonuses of non-managerial employees, even though these employees do not affect overall firm performance to a large degree (Efing et al., 2018). Results are also in line with the story that boards take the long-term firm performance into account when granting compensation.

Hypothesis 3 questions whether short-term industry performance–pay relation is positive and decreases in the quantile. I find no significant relation between industry short-term performance and total compensation in winsorized results in Table 2.2. The estimate of the location parameter is close to zero. The results from unwinsorized regressions in Table 2.6 suggest that there is a negative relation between short-term industry performance and pay, but this could be driven by outliers in the data. The results from the log-level specifications in Tables 2.9 and 2.12 in section 2.10 of the Appendix show similar results. Firms do not appear to use short-term industry benchmarking, but we cannot rule it out. This could be due to measurement of industry shocks, which may not capture special groups of peers used for relative evaluation (Bizjak et al., 2008).

Hypothesis 4 asks if the long-term industry performance–pay relation is positive, and decreases in the conditional pay quantile. The coefficient of interest belongs to the industry trend in Table 2.2. The location and scale parameters are imprecisely estimated. This is in line with the findings of the literature, in which there is mixed evidence that firms use industry performance benchmarking and also select peers in special groups (Edmans et al., 2017b). I can not entirely rule out long-term industry benchmarking, as estimates are noisy, and again the unwinsorized regressions suggest a negative correlation between long-term industry performance and compensation.

I next test Hypothesis 5, which asks whether short-term firm performance–pay sensitivity is the same for positive and negative short-term firm performance. I also test Hypothesis 6, which predicts reduction to the short-term firm performance–pay relation, when performance is negative, increases with the conditional total compensation quantile. This specification allows for different performance–pay sensitivities for positive and negative short-term performance, proxied by positive and negative short-term firm performance. I estimate this using an analogous regression to above, with an added interaction term with $\mathbf{I}\{S_f < 0\}_{ft}$, indicating negative short-term firm performance in firm f, at year t, and run

$$\begin{aligned}\hat{Q}_{\log Y_{ifst}}(\tau|\cdot) = (\hat{\alpha}_{ifs} + \hat{\delta}_{ifs}\hat{q}(\tau)) + (\log Shock_{ft} + \mathbf{I}\{S_f < 0\}_{ft} \times \log Shock_{ft} + \log Shock_{st} \\ + \log Trend_{ft} + \log Trend_{st} + Z'_{ft} + \psi_t)(\hat{\beta} + \hat{\gamma}\hat{q}(\tau)). \quad (2.9)\end{aligned}$$

Turning to estimation results in Table 2.3, the location parameters for the coefficients of interest, Log $Shock_f$, and the interaction with the indicator for negative short-term performance, are precisely estimated. The elasticity between total compensation and short-term firm performance, at the mean and median, is 0.05, in case performance is positive, but reduces to zero, in case firm performance is negative. The estimate of the scale function for the interaction term is positive for the bias-corrected point estimate, but very close to zero. Thus, executives are, quantitatively, equally well insured for bad performance across the conditional distribution, and the lower tail gains more from positive short-term performance.

This is significant evidence for an asymmetry in pay for positive and negative short-term firm performance, which supports Hypothesis 5. However, the asymmetry is quantitatively similar across the conditional distribution, rejecting Hypothesis 6. If anything, the asymmetry from negative shocks becomes smaller.

Regarding robustness of results, the unwinsorized results in Table 2.7 show similar results quantitatively and qualitatively, and here the degree of asymmetry is much smaller. The elasticity at the median in case of positive short-term firm performance is 0.22 (not significant), which reduces by 0.028 to about 0.19, in case short-term firm performance is negative. This supports the interpretation that actions and events that affect firm performance proportionately, correlate asymmetrically with compensation. The asymmetry is not driven by actions and events that only affect the dollar change in firm value to a small degree, as the interaction is not significant in any log-level models in the Appendix.

A potential confound of the short-term performance–pay relation is executives leaving the firm, or being fired if they perform poorly. Campbell and Thompson (2015) find that asymmetry in performance–pay for good and bad firm performance is likely used as a retention device, as the asymmetry is stronger when labor market conditions are favorable for executives. This means the executive's outside option is higher. Even if some managers are fired, we should still expect to observe asymmetry in short-term performance–pay sensitivity.

2.6.2 Executive wealth

The long-term development of accumulated pay is harder for the board to predict, since once pay is granted, it is harder to be renegotiated and can be strategically influenced by the executive. Firms grant executives stock-related pay over a successive number of periods that eventually vest after at least three years for stock at the median or four years for stock options (Edmans et al., 2019). The stock-related pay holdings co-move mechanically with the stock price, which is not the case for salary and cash bonus. This justifies a second test of the hypotheses using total wealth as the dependent variable as a proxy for executive utility. I test Hypotheses 1 to 4 again, using total wealth, the value of all accumulated equity-linked, and deferred compensation, as the dependent variable and measure of pay. Results are shown in Table 2.4, and in Table 2.8 for unwinsorized regressions.

Hypothesis 1 asks again, if the short-term firm performance-wealth relation is positive, and increases with the conditional wealth quantile. Testing Hypothesis 1, the coefficient of interest belongs to Log $Shock_f$. Column 1 shows an average increase of executive wealth by 0.05% for an increase of short-term firm value of 1%. The scale parameter on column 2 is not significantly estimated, especially the bias-corrected estimate in Panel B. This shows the short-term performance-wealth sensitivity is increasing in the distribution, but that this is imprecisely estimated. The point estimates for short-term firm performance pay substantially increase from the 10th percentile (0.03) to the 90th percentile (0.07). This is an economically significant increase of over 100% change in the short-term firm-performance pay elasticity. This result is in contrast to using total compensation as the measure of pay, and is in line with Hypothesis 1.

The results from log-level models in Tables 2.11 and 2.14 also support this finding. There is a significant and positive scale parameter estimated of Log $Shock_f$ in the winsorized regression. In Table 2.14, the scale parameter has the same sign with positive, and increasing point estimates of Log $Shock_f$ with the quantile.

An explanation for these findings is that stock-related pay in total wealth is very sensitive to the firm value. Further implications of this finding are also important and can be inferred from the literature. If stock-related pay is about to vest in a given month, executives have an incentive to announce M&A transactions or share buy-backs,

Table 2.3: Winsorized MM-QR of log total compensation on firm and industry performance measures with asymmetry

	(1) Location	(2) Scale	(3) Q10	(4) Q30	(5) Q50	(6) Q70	(7) Q90
A: MM-QR							
Log $Shock_f$	0.0544***	-0.0184***	0.0827***	0.0693***	0.0534***	0.0398***	0.0273***
	(0.0117)	(0.0062)	(0.0207)	(0.0158)	(0.0113)	(0.0099)	(0.0093)
$\mathbf{I}\{S_f < 0\}\times$	-0.0515***	0.0020	-0.0545***	-0.0531***	-0.0514***	-0.0499***	-0.0486***
Log S_f	(0.0036)	(0.0017)	(0.0048)	(0.0042)	(0.0037)	(0.0039)	(0.0040)
Log $Trend_f$	0.2083***	0.0143*	0.1863***	0.1968***	0.2091***	0.2197***	0.2294***
	(0.0159)	(0.0080)	(0.0235)	(0.0175)	(0.0160)	(0.0163)	(0.0183)
Log $Shock_s$	0.0041	-0.0095	0.0188	0.0118	0.0036	-0.0035	-0.0099
	(0.0157)	(0.0072)	(0.0207)	(0.0172)	(0.0167)	(0.0140)	(0.0158)
Log $Trend_s$	-0.0214	-0.0209*	0.0108	-0.0045	-0.0225	-0.0380	-0.0521*
	(0.0244)	(0.0123)	(0.0301)	(0.0304)	(0.0237)	(0.0257)	(0.0273)
B: Jackknife							
Log $Shock_f$	0.0544***	-0.0251***	0.0961***	0.0711***	0.0526***	0.0383***	0.0152
	(0.0117)	(0.0062)	(0.0207)	(0.0158)	(0.0113)	(0.0099)	(0.0093)
$\mathbf{I}\{S_f < 0\}\times$	-0.0515***	0.0034**	-0.0572***	-0.0538***	-0.0512***	-0.0493***	-0.0461***
Log S_f	(0.0036)	(0.0017)	(0.0048)	(0.0042)	(0.0037)	(0.0039)	(0.0040)
Log $Trend_f$	0.2083***	0.0201**	0.1749***	0.1950***	0.2098***	0.2212***	0.2397***
	(0.0159)	(0.0080)	(0.0235)	(0.0175)	(0.0160)	(0.0163)	(0.0183)
Log $Shock_s$	0.0041	-0.0159**	0.0305	0.0147	0.0030	-0.0061	-0.0207
	(0.0157)	(0.0072)	(0.0207)	(0.0172)	(0.0167)	(0.0140)	(0.0158)
Log $Trend_s$	-0.0214	-0.0188	0.0099	-0.0089	-0.0227	-0.0335	-0.0508*
	(0.0244)	(0.0123)	(0.0301)	(0.0304)	(0.0237)	(0.0257)	(0.0273)
N	25,636	25,636	25,636	25,636	25,636	25,636	25,636

Results show Method of Moments-Quantile Regressions (Panel A) of log(1+total compensation) on transformed firm and industry short-term performance, $\log(1 + (Shock + x))$, where x is the smallest number such that all values are non-negative, and $\log(Trend + 1)$ for firm and industry long-term performance, with executive-firm fixed effects. Panel B shows bias-corrected estimates of coefficients estimated in Panel A. This is done using the split-panel jackknife bias correction by Dhaene and Jochmans (2015), where the panel is split into even and odd years. Standard errors clustered at executive-firm level via bootstrap (200 reps) are in parentheses. *** $p < .01$, ** $p < .05$, * $p < .1$. Full set of controls: Age, age squared, GDP, GDP growth, inflation, and year dummies used. Controls not reported. Data set used is the full unbalanced panel from BoardEx executive compensation database, ORBIS for market value variables, and span years 2003-2013 and 34 countries. GDP growth and GDP per capita are from the World Bank, and inflation is from the IMF database. Observations with zero total compensation or base salary are removed. $\mathbf{I}\{S_f < 0\}$ takes value one if the shock in that year was negative, and zero otherwise. Compensation, shock and trend variables are winsorized at the 1st and 99th percentiles.

which can send the stock price soaring (Edmans et al., 2019). A prime example cited by Edmans et al. (2019) was the Bazaarvoice acquisition of PowerReviews in June 2012, which saw executives cashing in $90 Million US in stock after the stock went above $20. Executives were cited to know that the aim of the M&A transaction was to eliminate the primary competitor from the market. An antitrust law suit followed and the stock price declined to $7. This anecdote shows that the value of long-term incentives reacts to short-term behavior, in line with my findings. Higher conditional wealth is more strongly associated with short-term performance. This is in line with the concern that short-term behavior and stock price manipulation is a valid mechanism managers use to increase their own pay.

Since equity also vests, and the sale of equity is endogenous to the short-term stock price, executive wealth may be negatively correlated with the short-term firm performance for some observations. If the short-term firm value is down, executives may prefer to keep equity that is vested. If there is a jump in short-term firm performance due to a stock buy-back or an announced M&A transaction, they will sell equity that has vested (Edmans et al., 2019). This can reduce the coefficient size. Another reason is that winsorizing the data also introduces measurement error, causing attenuation bias.

Testing Hypothesis 2 again regarding long-term firm performance, the coefficient of interest belongs to Log $Trend_f$. In Table 2.4, the elasticity at the median (and mean, shown by the location parameter) is 0.75. Total wealth increases by 0.75% for a 1% increase in long-term firm performance. The mean and median long-term firm performance–pay elasticity is much larger using executive wealth (0.75), than using total compensation (0.21) as the dependent variable. This finding is in line with the empirical and theoretical literature on executive compensation, which shows that elasticities are generally larger when using total wealth as the dependent variable or utility measure (Edmans et al., 2017b). Further, results show a decrease in the correlation between long-term firm performance and executive wealth in upper quantiles for log-log models. The elasticity decreases significantly using the bias-corrected estimates in Panel B by 0.14 from the 10th to the 90th percentile. Short-term pay drives the wealth of managers comparatively more in the right tail of the conditional distribution.

This result is not supported by log-level models in Tables 2.11 and 2.14, which test incentives for a dollar change in firm-value. However, M&A transactions and share buybacks are more likely to affect firm value in a proportionate manner, rendering the log-log specification be a more appropriate model. The results show that long-term incentives are present and sizable for executives across the distribution, despite the heterogeneity. The concern about short-term behavior is supported when looking at the results in Table 2.8, and considering that some executives sell vesting equity when firm value increases.

I test Hypothesis 3 in Tables 2.4 and 2.8 using total wealth. The location parameter for short-term industry performance is noisily estimated, but point estimates show a positive correlation between executive pay and industry shocks in the lower tail. There is also significant heterogeneity, with the elasticity decreasing from 0.09 at the 10th percentile to 0.035 at the median. This is in line with Bizjak et al. (2008), who show that benchmarking of executives with respect to industry performance is more likely for executives receiving below median pay, as they have a competitive outside option if the industry is performing well, holding firm performance constant. However one must be careful comparing results directly, as estimates are of the conditional distribution.

Positive industry benchmarking in the short term is also supported by log-level models in Tables 2.11 and 2.14 in section 2.10 of the Appendix. One potential mechanism is that executives may receive more pay in the form of stock-related pay if the industry is performing better, with firm performance held constant. When testing Hypothesis 4, parameters are noisily estimated and if anything negative, but far from significant.

2.7 Discussion

This paper is concerned with distributional differences in the time horizon of executive compensation. The literature hitherto has identified short-termism as a problem, but not systematically shown whether higher (conditonal) pay is associated with more short-term, or long-term oriented incentives. The quantile regression framework with fixed effects used, MM-QR, allows a systematic analysis of this increasingly important question.

The findings show that performance pay is more short-term oriented in the lower tail of the distribution when using total compensation. This is potentially driven by risk-sharing, as firms must reduce bonuses in bad times to maintain liquidity (Efing et al., 2018). By contrast, performance pay is more short-term oriented in the upper tail of the distribution when total wealth is used to measure pay. This could be because yearly compensation is easier for the board to measure and control. It can be easily adjusted by the firm to account for short-run and long-run performance in each period. Compensation committees can adjust bonuses and amount of options and performance

Table 2.4: Winsorized MM-QR of log total wealth on log of firm and industry performance measures

	(1) Loc.	(2) Scale	(3) Q10	(4) Q30	(5) Q50	(6) Q70	(7) Q90
A:MM-QR							
Log $Shock_f$	0.0511***	0.0126*	0.0313	0.0411***	0.0525***	0.0617***	0.0692***
	(0.0131)	(0.0070)	(0.0198)	(0.0159)	(0.0135)	(0.0121)	(0.0126)
Log $Trend_f$	0.7516***	-0.0280**	0.7954***	0.7739***	0.7487***	0.7281***	0.7117***
	(0.0341)	(0.0121)	(0.0386)	(0.0358)	(0.0331)	(0.0340)	(0.0345)
Log $Shock_s$	0.0390*	-0.0247**	0.0777***	0.0587***	0.0364*	0.0182	0.0037
	(0.0203)	(0.0102)	(0.0290)	(0.0216)	(0.0191)	(0.0193)	(0.0197)
Log $Trend_s$	-0.0299	-0.0193	0.0003	-0.0146	-0.0320	-0.0461	-0.0575
	(0.0477)	(0.0199)	(0.0596)	(0.0547)	(0.0503)	(0.0471)	(0.0548)
B: Jackknife							
Log $Shock_f$	0.0511***	0.0088	0.0364*	0.0455***	0.0521***	0.0577***	0.0645***
	(0.0131)	(0.0070)	(0.0198)	(0.0159)	(0.0135)	(0.0121)	(0.0126)
Log $Trend_f$	0.7516***	-0.0445***	0.8261***	0.7800***	0.7465***	0.7186***	0.6843***
	(0.0341)	(0.0121)	(0.0386)	(0.0358)	(0.0331)	(0.0340)	(0.0345)
Log $Shock_s$	0.0390*	-0.0313***	0.0914***	0.0590***	0.0354*	0.0158	-0.0084
	(0.0203)	(0.0102)	(0.0290)	(0.0216)	(0.0191)	(0.0193)	(0.0197)
Log $Trend_s$	-0.0299	-0.0050	-0.0215	-0.0267	-0.0305	-0.0337	-0.0375
	(0.0477)	(0.0199)	(0.0596)	(0.0547)	(0.0503)	(0.0471)	(0.0548)
N	22,075	22,075	22,075	22,075	22,075	22,075	22,075

Results show Method of Moments-Quantile Regressions (Panel A) of log(1+total wealth) on transformed firm and industry short-term performance, $\log(1 + (Shock + x))$, where x is the smallest number such that all values are non-negative, and $\log(Trend + 1)$ for firm and industry long-term performance, with executive-firm fixed effects. Panel B shows bias-corrected estimates of coefficients estimated in Panel A. This is done using the split-panel jackknife bias correction by Dhaene and Jochmans (2015), where the panel is split into even and odd years. Standard errors clustered at executive-firm level via bootstrap (200 reps) are in parentheses. *** $p < .01$, ** $p < .05$, * $p < .1$. Full set of controls: Age, age squared, GDP, GDP growth, inflation, and year dummies used. Controls not reported. Data set used is the full unbalanced panel from BoardEx executive compensation database, ORBIS for market value variables, and span years 2003-2013 and 34 countries. GDP growth and GDP per capita are from the World Bank, and inflation is from the IMF database. Observations with zero total compensation or base salary are removed. Compensation, shock and trend variables are winsorized at the 1st and 99th percentiles.

stock granted, but accumulated long-term pay, mostly consisting of stock-related pay, cannot be easily adjusted once granted. Once awarded, there is more potential for executives to strategically increase their payout, which may not coincide with maximizing long-term business performance.

Stronger wealth sensitivity to short-term shocks is potentially driven by a ratchet effect of long-term incentive plans, found by (Bebchuk et al., 2010). Some firms pay a fixed value of stock options or fixed value of performance stock, which means the number of stocks granted is adjusted according to the last month's stock price. The calculation of the number of shares or options is usually based on averages of around 30 trading days, which is not enough to eliminate fluctuations that occur on a yearly basis. A low short-term firm value at the grant date would result in a larger number of shares being granted with fixed value, e.g. in the case of performance shares, and possibly also restricted stock. This gives greater leverage to the executive when firm value rises again. In the case of stock options, the same effect occurs if the share price is relatively low before allocation and therefore a lower strike price is set, or the executive receives a larger number of options. Supporting this mechanism, Yermack (1997) finds that stocks show negative abnormal returns before granting stock options, and positive abnormal returns after granting.

A potential confound of the findings in this paper is reverse causation. It is possible for higher wages to change performance, however there have been scant attempts by studies to show a causal direction from pay to firm performance (Edmans et al., 2017b). Another problem is that contracts are not observed in my study, and are endogenous (Edmans and Gabaix, 2016). The study cannot infer without knowing contracts how much of results are driven by strategic manipulation of stock prices and other short-term behaviors, and how much of the results are from an efficient market. Thus, the policy implications must be taken with caution. It is also possible under certain conditions for pay to be more strongly related to the short-term stock performance, when there is sufficient ambiguity about manipulation (Peng and Röell, 2014).

Long-term compensation can only be amended ex-post if there are contractual means to restrict the payout to certain conditions, or regain already granted stock-related pay in certain cases. In the United States, claw-back clauses that can recover pay in case of fraud are written in law, but these can not prevent short-term behavior that is *legal*, such as investment choice.[16]

Theory suggests multiple mechanisms to deal with the concerns about short-termism. An optimal contract with concerns about stock price manipulation after vesting relates current pay to all past performance dates of firm performance, conditions vesting on performance, and not on a fixed date, and can also shift some vesting into retirement (Edmans et al., 2012; Marinovic and Varas, 2019). Lengthening the vesting period into retirement is costly for the executive near career end, as this exposes the executive to more risk.

A practical implementation of this, would be to link the payout value of stock-related pay to the average price during the entire holding period. Since stock-related pay is often emitted in rolling windows of three or four years, this would not cause a large disadvantage to executives who do increase long-term firm value, and reduce the opportunity to game the incentive pay system. Since firms have limited liability, the stock would not be paid out in case of bankruptcy.

There is always discretion involved in choosing which system to use to incentivize executives. In reality, this is determined by a market equilibrium, making it potentially difficult for a single firm to implement policy recommendations. The results together cast some doubt with the growing body of evidence on effectiveness of long-term incentive plans to prevent managerial skimming.

[16] The Dodd-Frank Act of 2010 and the Sarbanes-Oxley Act of 2002 in the US enable firms to regain payments after fraud, even if the executive is not charged. This is implemented by the SEC (Edmans et al., 2017b).

2.8 Total Wealth

Definition of total wealth. The total wealth variable is calculated by BoardEx, who provide the following definition in the Data dictionary. The closing stock price of the annual report date is used. They implement a Black-Scholes option pricing model. Volatility is measured using 100 days of historic stock prices. The risk free rate is measured using the following: UK = 6 months Libor rate, Europe = EURIBOR, US = 10 year T-Bill, otherwise = 6.5% It is assumed that exercise is on expiry date whether known or assumed.

2.9 Band-Pass Filter

I now describe the random walk approximation from Christiano and Fitzgerald (2003) used to generate shock and trend variables, largely using description in the Stata 15 Manual. the proxies for short and long-term firm performance. Assuming the ideal band-pass filter generates the data y_t when applied to raw data x_t, and the data generated from the approximation yield $\hat{y}_t$, then the problem of getting the best approximation minimizes the mean squared error between the ideal filter and the approximation:

$$E\left[(s_t - \hat{s}_t)^2 \mid y\right], \quad y \equiv [y_1, \cdots, y_t]. \tag{2.10}$$

The ideal filter generates the cyclical component using

$$s_t = \sum_{j=-\infty}^{\infty} b_j y_{t-j}. \tag{2.11}$$

So $\hat{s}_t$ is a linear projection of s on y in each t. The aim is to identify a part of y_t that oscillates with a period between p_l and p_h, where $2 \leq p_l < p_h < \infty$, using the random walk filter in the non-stationary asymmetric case (Christiano and Fitzgerald, 2003). The filter works for non-stationary data in that it is robust up to one unit root, but it is not invariant to drift. One must remove drift so that the filter induces stationarity. Assuming a random walk plus drift process transforms the original series using

$$z_t = y_t - \frac{(t-1)(y_t - y_1)}{(T-1)}. \tag{2.12}$$

2.10 Additional tables

Table 2.5: Summary statistics of winsorized variables

Variable	N	Mean	S.D.	Min.	Max.
Age	25,636	51.7	8.1	25	92
Salary*	25,502	522.1	425.7	17.4	2,317.1
Bonus*	21,686	437.1	725.9	0	4,221.8
Equity-linked*	13,956	2,101.2	4,392.7	2.9	29,881
Variable compensation*	25,502	1,462.3	3,151.2	0	21,359
Total compensation*	25,636	1,988.1	3,393.1	18.8	22,615
Total wealth*	22,075	15,672.3	52,480.3	5.3	418,912.6
Market value of equity**	25,636	7,110	17,147.6	3	99,091
$Shock_f$***	25,636	0	2.6	-11.6	12.7
$Trend_f$***	25,636	7.1	16.8	0	95.1
$Shock_s$***	25,636	0.1	2.2	-7.4	8.9
$Trend_s$***	25,636	6.7	6.2	0.5	30.8
GDP growth	25,636	1	2.5	-8.3	15.2
GDP per capita	25,636	42,785.1	7,348.3	1,724.7	115,109.3
Inflation	25,636	2.6	1.1	-1.7	14.1

*** scaled in Billions $US, ** scaled in millions $US, * scaled in thousands $US. The sum of salary, bonus, grant-date value of newly emitted equity-linked and long-term incentive plans, equals total compensation. Observations with zero salary or total compensation are dropped from the data. The sample includes executives for which there is at least one non-missing observation of total compensation. If an executive works in two firms at the same time, plausibility checks were done and some observations dropped according to the following criteria. If data is entirely missing, this observation is deleted. If there is a holding company or a subsidiary, and the executive had the same position at both firms, the observation belonging to the parent company is kept. If one position was only a representative or deputy position, this observation is deleted. If the firm is listed in two countries, the headquarter country is kept. If the executive switched positions and thus worked for two companies in one year, the first year of the new job is deleted, as this generally covers fewer months. Firm and industry shock and trend variables are generated using the band-pass filter, removing drift and cycles between 2 and 8 years from the raw data to generate the trend. GDP growth and GDP per-capita are from the World Bank, and Inflation is measured as the average percentage change in consumer prices in each country, which is from the IMF database. All compensation, shock and trend variables are winsorized at the 1st and 99th percentiles.

Table 2.6: MM-QR of log total compensation on log of firm and industry performance measures

	(1) Location	(2) Scale	(3) Q10	(4) Q30	(5) Q50	(6) Q70	(7) Q90
A: MM-QR							
Log $Shock_f$	0.6676***	-0.1745***	0.9359***	0.8079***	0.6598***	0.5286***	0.4097***
	(0.1503)	(0.0664)	(0.2270)	(0.1647)	(0.1574)	(0.1255)	(0.1329)
Log $Trend_f$	0.1947***	0.0032	0.1899***	0.1922***	0.1949***	0.1972***	0.1994***
	(0.0173)	(0.0087)	(0.0218)	(0.0172)	(0.0177)	(0.0178)	(0.0206)
Log $Shock_s$	-0.0624**	-0.0068	-0.0519*	-0.0569**	-0.0627**	-0.0679**	-0.0725**
	(0.0277)	(0.0124)	(0.0307)	(0.0273)	(0.0270)	(0.0317)	(0.0343)
Log $Trend_s$	-0.0384*	-0.0117	-0.0205	-0.0290	-0.0389*	-0.0477**	-0.0556**
	(0.0225)	(0.0114)	(0.0305)	(0.0252)	(0.0224)	(0.0237)	(0.0279)
B: Jackknife							
Log $Shock_f$	0.6676***	-0.2237***	1.0374***	0.8130***	0.6541***	0.5237***	0.3161**
	(0.1503)	(0.0664)	(0.2270)	(0.1647)	(0.1574)	(0.1255)	(0.1329)
Log $Trend_f$	0.1947***	0.0015	0.1922***	0.1937***	0.1948***	0.1957***	0.1971***
	(0.0173)	(0.0087)	(0.0218)	(0.0172)	(0.0177)	(0.0178)	(0.0206)
Log $Shock_s$	-0.0624**	-0.0220*	-0.0261	-0.0481*	-0.0638**	-0.0766**	-0.0969***
	(0.0277)	(0.0124)	(0.0307)	(0.0273)	(0.0270)	(0.0317)	(0.0343)
Log $Trend_s$	-0.0384*	-0.0080	-0.0251	-0.0332	-0.0389*	-0.0436*	-0.0510*
	(0.0225)	(0.0114)	(0.0305)	(0.0252)	(0.0224)	(0.0237)	(0.0279)
N	25,636	25,636	25,636	25,636	25,636	25636	25,636

Results show Method of Moments-Quantile Regressions (Panel A) of log(1+total compensation) on transformed firm and industry short-term performance, $\log(1 + (Shock + x))$, where x is the smallest number such that all values are non-negative, and $\log(Trend + 1)$ for firm and industry long-term performance, with executive-firm fixed effects. Panel B shows bias-corrected estimates of coefficients estimated in Panel A. This is done using the split-panel jackknife bias correction by Dhaene and Jochmans (2015), where the panel is split into even and odd years. Standard errors clustered at executive-firm level via bootstrap (200 reps) are in parentheses. *** $p < .01$, ** $p < .05$, * $p < .1$. Full set of controls: Age, age squared, GDP, GDP growth, inflation, and year dummies used. Controls not reported. Data set used is the full unbalanced panel from BoardEx executive compensation database, ORBIS for market value variables, and span years 2003-2013 and 34 countries. GDP growth and GDP per capita are from the World Bank, and inflation is from the IMF database. Observations with zero total compensation or base salary are removed.

Table 2.7: MM-QR of log total compensation on firm and industry performance measures with asymmetry

	(1) Location	(2) Scale	(3) Q10	(4) Q30	(5) Q50	(6) Q70	(7) Q90
A: MM-QR							
Log $Shock_f$	0.2281	-0.1624**	0.4766**	0.3597**	0.2206	0.0989	-0.0113
	(0.1489)	(0.0674)	(0.2320)	(0.1737)	(0.1491)	(0.1351)	(0.1340)
$\mathbf{I}\{S_f < 0\}\times$ Log S_f	-0.0283***	0.0006	-0.0291***	-0.0287***	-0.0283***	-0.0279***	-0.0275***
	(0.0020)	(0.0010)	(0.0026)	(0.0023)	(0.0021)	(0.0023)	(0.0024)
Log $Trend_f$	0.1968***	0.0026	0.1928***	0.1947***	0.1969***	0.1989***	0.2007***
	(0.0170)	(0.0086)	(0.0247)	(0.0189)	(0.0175)	(0.0182)	(0.0203)
Log $Shock_s$	-0.0661**	-0.0113	-0.0489*	-0.0570**	-0.0667**	-0.0751**	-0.0828**
	(0.0275)	(0.0123)	(0.0290)	(0.0258)	(0.0292)	(0.0308)	(0.0344)
Log $Trend_s$	-0.0353	-0.0141	-0.0137	-0.0239	-0.0359*	-0.0465**	-0.0561**
	(0.0222)	(0.0114)	(0.0283)	(0.0283)	(0.0215)	(0.0233)	(0.0241)
B: Jackknife							
Log $Shock_f$	0.2281	-0.2198***	0.5891**	0.3736**	0.2141	0.0867	-0.1172
	(0.1489)	(0.0674)	(0.2320)	(0.1737)	(0.1491)	(0.1351)	(0.1340)
$\mathbf{I}\{S_f < 0\}\times$ Log S_f	-0.0283***	0.0010	-0.0299***	-0.0290***	-0.0282***	-0.0277***	-0.0267***
	(0.0020)	(0.0010)	(0.0026)	(0.0023)	(0.0021)	(0.0023)	(0.0024)
Log $Trend_f$	0.1968***	0.0021	0.1933***	0.1954***	0.1969***	0.1982***	0.2001***
	(0.0170)	(0.0086)	(0.0247)	(0.0189)	(0.0175)	(0.0182)	(0.0203)
Log $Shock_s$	-0.0661**	-0.0269**	-0.0220	-0.0484*	-0.0679**	-0.0834***	-0.1083***
	(0.0275)	(0.0123)	(0.0290)	(0.0258)	(0.0292)	(0.0308)	(0.0344)
Log $Trend_s$	-0.0353	-0.0114	-0.0167	-0.0278	-0.0360*	-0.0426*	-0.0531**
	(0.0222)	(0.0114)	(0.0283)	(0.0283)	(0.0215)	(0.0233)	(0.0241)
N	25,636	25,636	25,636	25,636	25,636	25,636	25,636

Results show Method of Moments-Quantile Regressions (Panel A) of log(1+total compensation) on transformed firm and industry short-term performance, $\log(1 + (Shock + x))$, where x is the smallest number such that all values are non-negative, and $\log(Trend + 1)$ for firm and industry long-term performance, with executive-firm fixed effects. Panel B shows bias-corrected estimates of coefficients estimated in Panel A. This is done using the split-panel jackknife bias correction by Dhaene and Jochmans (2015), where the panel is split into even and odd years. Standard errors clustered at executive-firm level via bootstrap (200 reps) are in parentheses. *** $p < .01$, ** $p < .05$, * $p < .1$. Full set of controls: Age, age squared, GDP, GDP growth, inflation, and year dummies used. Controls not reported. Data set used is the full unbalanced panel from BoardEx executive compensation database, ORBIS for market value variables, and span years 2003-2013 and 34 countries. GDP growth and GDP per capita are from the World Bank, and inflation is from the IMF database. Observations with zero total compensation or base salary are removed. $\mathbf{I}\{S_f < 0\}$ takes value one if the shock in that year was negative, and zero otherwise.

Table 2.8: MM-QR of log total wealth on log of firm and industry performance measures unwinsorized data

	(1) Location	(2) Scale	(3) Q10	(4) Q30	(5) Q50	(6) Q70	(7) Q90
A: MM-QR							
Log $Shock_f$	0.9049***	0.1284	0.7045***	0.8025***	0.9192***	1.0125***	1.0897***
	(0.1724)	(0.0879)	(0.2486)	(0.2115)	(0.1779)	(0.1654)	(0.1940)
Log $Trend_f$	0.6711***	-0.0539***	0.7552***	0.7141***	0.6651***	0.6259***	0.5935***
	(0.0512)	(0.0164)	(0.0456)	(0.0497)	(0.0511)	(0.0573)	(0.0614)
Log $Shock_s$	-0.0147	-0.0012	-0.0128	-0.0137	-0.0148	-0.0157	-0.0165
	(0.0351)	(0.0165)	(0.0446)	(0.0367)	(0.0378)	(0.0354)	(0.0359)
Log $Trend_s$	-0.0235	-0.0093	-0.0090	-0.0161	-0.0246	-0.0313	-0.0369
	(0.0483)	(0.0200)	(0.0623)	(0.0551)	(0.0518)	(0.0453)	(0.0521)
B: Jackknife							
Log $Shock_f$	0.9049***	0.0165	0.8773***	0.8943***	0.9069***	0.9172***	0.9303***
	(0.1724)	(0.0879)	(0.2486)	(0.2115)	(0.1779)	(0.1654)	(0.1940)
Log $Trend_f$	0.6711***	-0.0801***	0.8051***	0.7227***	0.6615***	0.6118***	0.5481***
	(0.0512)	(0.0164)	(0.0456)	(0.0497)	(0.0511)	(0.0573)	(0.0614)
Log $Shock_s$	-0.0147	0.0466***	-0.0925**	-0.0447	-0.0092	0.0197	0.0567
	(0.0351)	(0.0165)	(0.0446)	(0.0367)	(0.0378)	(0.0354)	(0.0359)
Log $Trend_s$	-0.0235	0.0065	-0.0344	-0.0277	-0.0228	-0.0187	-0.0136
	(0.0483)	(0.0200)	(0.0623)	(0.0551)	(0.0518)	(0.0453)	(0.0521)
N	22,075	22,075	22,075	22,075	22,075	22,075	22,075

Results show Method of Moments-Quantile Regressions (Panel A) of log(1+total wealth) on transformed firm and industry short-term performance, $\log(1 + (Shock + x))$, where x is the smallest number such that all values are non-negative, and $\log(Trend + 1)$ for firm and industry long-term performance, with executive-firm fixed effects. Panel B shows bias-corrected estimates of coefficients estimated in Panel A. This is done using the split-panel jackknife bias correction by Dhaene and Jochmans (2015), where the panel is split into even and odd years. Standard errors clustered at executive-firm level via bootstrap (200 reps) are in parentheses. $^{***}\ p < .01$, $^{**}\ p < .05$, $^{*}\ p < .1$. Full set of controls: Age, age squared, GDP, GDP growth, inflation, and year dummies used. Controls not reported. Data set used is the full unbalanced panel from BoardEx executive compensation database, ORBIS for market value variables, and span years 2003-2013 and 34 countries. GDP growth and GDP per capita are from the World Bank, and inflation is from the IMF database. Observations with zero total compensation or base salary are removed.

Table 2.9: Winsorized MM-QR of log total compensation on firm and industry performance measures

	(1) Location	(2) Scale	(3) Q10	(4) Q30	(5) Q50	(6) Q70	(7) Q90
A: MM-QR							
$Shock_f$	0.0075***	-0.0027***	0.0117***	0.0097***	0.0074***	0.0054***	0.0036**
	(0.0017)	(0.0009)	(0.0028)	(0.0022)	(0.0017)	(0.0014)	(0.0015)
$Trend_f$	0.0086***	0.0030***	0.0040	0.0062**	0.0088***	0.0110***	0.0130***
	(0.0024)	(0.0011)	(0.0034)	(0.0029)	(0.0027)	(0.0023)	(0.0020)
$Shock_s$	0.0013	-0.0006	0.0023	0.0018	0.0013	0.0008	0.0004
	(0.0024)	(0.0012)	(0.0033)	(0.0027)	(0.0026)	(0.0023)	(0.0026)
$Trend_s$	-0.0013	-0.0014	0.0008	-0.0002	-0.0014	-0.0024	-0.0034
	(0.0023)	(0.0010)	(0.0031)	(0.0026)	(0.0023)	(0.0022)	(0.0027)
B: Jackknife							
$Shock_f$	0.0075***	-0.0035***	0.0134***	0.0099***	0.0072***	0.0053***	0.0020
	(0.0017)	(0.0009)	(0.0028)	(0.0022)	(0.0017)	(0.0014)	(0.0015)
$Trend_f$	0.0086***	0.0043***	0.0014	0.0058**	0.0090***	0.0114***	0.0154***
	(0.0024)	(0.0011)	(0.0034)	(0.0029)	(0.0027)	(0.0023)	(0.0020)
$Shock_s$	0.0013	-0.0011	0.0032	0.0020	0.0012	0.0006	-0.0005
	(0.0024)	(0.0012)	(0.0033)	(0.0027)	(0.0026)	(0.0023)	(0.0026)
$Trend_s$	-0.0013	-0.0013	0.0009	-0.0004	-0.0014	-0.0022	-0.0034
	(0.0023)	(0.0010)	(0.0031)	(0.0026)	(0.0023)	(0.0022)	(0.0027)
N	25,636	25,636	25,636	25,636	25,636	25,636	25,636

Results show Method of Moments-Quantile Regressions (Panel A) of log(1+total compensation) on firm and industry short-term performance, and firm and industry long-term performance, with executive-firm fixed effects. Panel B shows bias-corrected estimates of coefficients estimated in Panel A. This is done using the split-panel jackknife bias correction by Dhaene and Jochmans (2015), where the panel is split into even and odd years. Standard errors clustered at executive-firm level via bootstrap (200 reps) are in parentheses. *** $p < .01$, ** $p < .05$, * $p < .1$. Full set of controls: Age, age squared, GDP, GDP growth, inflation, and year dummies used. Controls not reported. Data set used is the full unbalanced panel from BoardEx executive compensation database, ORBIS for market value variables, and span years 2003-2013 and 34 countries. GDP growth and GDP per capita are from the World Bank, and inflation is from the IMF database. Observations with zero total compensation or base salary are removed.

Table 2.10: Winsorized MM-QR of log total compensation on firm and industry performance measures with asymmetry

	(1) Location	(2) Scale	(3) Q10	(4) Q30	(5) Q50	(6) Q70	(7) Q90
A: MM-QR							
$Shock_f$	0.0040	-0.0004	0.0046	0.0043	0.0040	0.0037	0.0035
	(0.0035)	(0.0022)	(0.0056)	(0.0047)	(0.0034)	(0.0027)	(0.0030)
$\mathbf{I}\{S_f < 0\} \times S_f$	0.0074	-0.0049	0.0150	0.0114	0.0071	0.0035	0.0001
	(0.0068)	(0.0041)	(0.0103)	(0.0092)	(0.0066)	(0.0051)	(0.0058)
$Trend_f$	0.0089***	0.0028***	0.0046	0.0066**	0.0090***	0.0112***	0.0131***
	(0.0025)	(0.0011)	(0.0040)	(0.0029)	(0.0024)	(0.0021)	(0.0022)
$Shock_s$	0.0013	-0.0007	0.0024	0.0019	0.0013	0.0008	0.0004
	(0.0024)	(0.0012)	(0.0033)	(0.0027)	(0.0024)	(0.0026)	(0.0023)
$Trend_s$	-0.0014	-0.0014	0.0007	-0.0003	-0.0015	-0.0025	-0.0034
	(0.0023)	(0.0011)	(0.0028)	(0.0026)	(0.0023)	(0.0022)	(0.0028)
B: Jackknife							
$Shock_f$	0.0040	0.0010	0.0024	0.0034	0.0041	0.0046*	0.0056*
	(0.0035)	(0.0022)	(0.0056)	(0.0047)	(0.0034)	(0.0027)	(0.0030)
$\mathbf{I}\{S_f < 0\} \times S_f$	0.0074	-0.0099**	0.0239**	0.0139	0.0066	0.0010	-0.0081
	(0.0068)	(0.0041)	(0.0103)	(0.0092)	(0.0066)	(0.0051)	(0.0058)
$Trend_f$	0.0089***	0.0060***	-0.0012	0.0049*	0.0094***	0.0128***	0.0184***
	(0.0025)	(0.0011)	(0.0040)	(0.0029)	(0.0024)	(0.0021)	(0.0022)
$Shock_s$	0.0013	-0.0011	0.0032	0.0021	0.0012	0.0006	-0.0004
	(0.0024)	(0.0012)	(0.0033)	(0.0027)	(0.0024)	(0.0026)	(0.0023)
$Trend_s$	-0.0014	-0.0013	0.0007	-0.0006	-0.0015	-0.0022	-0.0034
	(0.0023)	(0.0011)	(0.0028)	(0.0026)	(0.0023)	(0.0022)	(0.0028)
N	25,636	25,636	25,636	25,636	25,636	25,636	25,636

Results show Method of Moments-Quantile Regressions (Panel A) of log(1+total compensation) on firm and industry short-term performance, and firm and industry long-term performance, with executive-firm fixed effects. Panel B shows bias-corrected estimates of coefficients estimated in Panel A. This is done using the split-panel jackknife bias correction by Dhaene and Jochmans (2015), where the panel is split into even and odd years. Standard errors clustered at executive-firm level via bootstrap (200 reps) are in parentheses. *** $p < .01$, ** $p < .05$, * $p < .1$. Full set of controls: Age, age squared, GDP, GDP growth, inflation, and year dummies used. Controls not reported. Data set used is the full unbalanced panel from BoardEx executive compensation database, ORBIS for market value variables, and span years 2003-2013 and 34 countries. GDP growth and GDP per capita are from the World Bank, and inflation is from the IMF database. Observations with zero total compensation or base salary are removed. $\mathbf{I}\{S_f < 0\}$ takes value one if the shock in that year was negative, and zero otherwise. Compensation, shock and trend variables are winsorized at the 1st and 99th percentiles.

Table 2.11: Winsorized MM-QR of log total wealth on firm and industry performance measures

	(1) Location	(2) Scale	(3) Q10	(4) Q30	(5) Q50	(6) Q70	(7) Q90
A: MM-QR							
$Shock_f$	0.0112***	0.0026**	0.0071**	0.0090***	0.0114***	0.0134***	0.0149***
	(0.0022)	(0.0012)	(0.0035)	(0.0026)	(0.0020)	(0.0023)	(0.0022)
$Trend_f$	0.0207***	0.0013	0.0186***	0.0196***	0.0208***	0.0218***	0.0226***
	(0.0042)	(0.0014)	(0.0049)	(0.0047)	(0.0042)	(0.0041)	(0.0039)
$Shock_s$	0.0053	-0.0047***	0.0127***	0.0092**	0.0049	0.0014	-0.0014
	(0.0035)	(0.0017)	(0.0048)	(0.0041)	(0.0034)	(0.0031)	(0.0034)
$Trend_s$	0.0024	-0.0017	0.0050	0.0038	0.0022	0.0010	-0.0000
	(0.0040)	(0.0018)	(0.0052)	(0.0044)	(0.0039)	(0.0037)	(0.0044)
B: Jackknife							
$Shock_f$	0.0112***	0.0019*	0.0079**	0.0099***	0.0114***	0.0126***	0.0140***
	(0.0022)	(0.0012)	(0.0035)	(0.0026)	(0.0020)	(0.0023)	(0.0022)
$Trend_f$	0.0207***	0.0025*	0.0165***	0.0190***	0.0210***	0.0226***	0.0245***
	(0.0042)	(0.0014)	(0.0049)	(0.0047)	(0.0042)	(0.0041)	(0.0039)
$Shock_s$	0.0053	-0.0049***	0.0135***	0.0086**	0.0048	0.0016	-0.0020
	(0.0035)	(0.0017)	(0.0048)	(0.0041)	(0.0034)	(0.0031)	(0.0034)
$Trend_s$	0.0024	-0.0006	0.0033	0.0028	0.0023	0.0020	0.0015
	(0.0040)	(0.0018)	(0.0052)	(0.0044)	(0.0039)	(0.0037)	(0.0044)
N	22,075	22,075	22,075	22,075	22,075	22,075	22,075

Results show Method of Moments-Quantile Regressions (Panel A) of log(1+total wealth) on firm and industry short-term performance, and firm and industry long-term performance, with executive-firm fixed effects. Panel B shows bias-corrected estimates of coefficients estimated in Panel A. This is done using the split-panel jackknife bias correction by Dhaene and Jochmans (2015), where the panel is split into even and odd years. Standard errors clustered at executive-firm level via bootstrap (200 reps) are in parentheses. *** $p < .01$, ** $p < .05$, * $p < .1$. Full set of controls: Age, age squared, GDP, GDP growth, inflation, and year dummies used. Controls not reported. Data set used is the full unbalanced panel from BoardEx executive compensation database, ORBIS for market value variables, and span years 2003-2013 and 34 countries. GDP growth and GDP per capita are from the World Bank, and inflation is from the IMF database. Observations with zero total compensation or base salary are removed. Compensation, shock and trend variables are winsorized at the 1st and 99th percentiles.

Table 2.12: MM-QR of log total compensation on firm and industry performance measures

	(1) Location	(2) Scale	(3) Q10	(4) Q30	(5) Q50	(6) Q70	(7) Q90
A: MM-QR							
$Shock_f$	0.0046***	-0.0014**	0.0069***	0.0058***	0.0046***	0.0035***	0.0025**
	(0.0012)	(0.0006)	(0.0018)	(0.0014)	(0.0012)	(0.0010)	(0.0012)
$Trend_f$	0.0049**	0.0025**	0.0011	0.0029	0.0050**	0.0069***	0.0085***
	(0.0021)	(0.0010)	(0.0030)	(0.0024)	(0.0025)	(0.0020)	(0.0020)
$Shock_s$	0.0008	-0.0008	0.0021	0.0015	0.0007	0.0001	-0.0005
	(0.0022)	(0.0010)	(0.0028)	(0.0023)	(0.0023)	(0.0020)	(0.0023)
$Trend_s$	0.0015	-0.0011	0.0032	0.0024	0.0014	0.0006	-0.0002
	(0.0020)	(0.0009)	(0.0027)	(0.0023)	(0.0020)	(0.0020)	(0.0024)
B: Jackknife							
$Shock_f$	0.0046***	-0.0020***	0.0079***	0.0059***	0.0045***	0.0034***	0.0015
	(0.0012)	(0.0006)	(0.0018)	(0.0014)	(0.0012)	(0.0010)	(0.0012)
$Trend_f$	0.0049**	0.0033***	-0.0005	0.0028	0.0051**	0.0070***	0.0101***
	(0.0021)	(0.0010)	(0.0030)	(0.0024)	(0.0025)	(0.0020)	(0.0020)
$Shock_s$	0.0008	-0.0014	0.0030	0.0017	0.0007	-0.0001	-0.0014
	(0.0022)	(0.0010)	(0.0028)	(0.0023)	(0.0023)	(0.0020)	(0.0023)
$Trend_s$	0.0015	-0.0010	0.0031	0.0021	0.0014	0.0009	-0.0001
	(0.0020)	(0.0009)	(0.0027)	(0.0023)	(0.0020)	(0.0020)	(0.0024)
N	25,636	25,636	25,636	25,636	25,636	25,636	25,636

Results show Method of Moments-Quantile Regressions (Panel A) of log(1+total compensation) on firm and industry short-term performance, and firm and industry long-term performance, with executive-firm fixed effects. Panel B shows bias-corrected estimates of coefficients estimated in Panel A. This is done using the split-panel jackknife bias correction by Dhaene and Jochmans (2015), where the panel is split into even and odd years. Standard errors clustered at executive-firm level via bootstrap (200 reps) are in parentheses. *** $p < .01$, ** $p < .05$, * $p < .1$. Full set of controls: Age, age squared, GDP, GDP growth, inflation, and year dummies used. Controls not reported. Data set used is the full unbalanced panel from BoardEx executive compensation database, ORBIS for market value variables, and span years 2003-2013 and 34 countries. GDP growth and GDP per capita are from the World Bank, and inflation is from the IMF database. Observations with zero total compensation or base salary are removed.

Table 2.13: MM-QR of log total compensation on firm and industry performance measures

	(1) Location	(2) Scale	(3) Q10	(4) Q30	(5) Q50	(6) Q70	(7) Q90
A: MM-QR							
$Shock_f$	0.0041**	-0.0005	0.0049*	0.0045*	0.0041**	0.0037**	0.0034
	(0.0019)	(0.0013)	(0.0029)	(0.0023)	(0.0018)	(0.0017)	(0.0021)
$\mathbf{I}\{S_f < 0\} \times S_f$	0.0012	-0.0022	0.0046	0.0030	0.0012	-0.0005	-0.0020
	(0.0039)	(0.0025)	(0.0060)	(0.0051)	(0.0039)	(0.0033)	(0.0040)
$Trend_f$	0.0050**	0.0023**	0.0014	0.0031	0.0051**	0.0069***	0.0085***
	(0.0022)	(0.0010)	(0.0034)	(0.0024)	(0.0020)	(0.0021)	(0.0024)
$Shock_s$	0.0008	-0.0009	0.0021	0.0015	0.0007	0.0001	-0.0005
	(0.0022)	(0.0010)	(0.0029)	(0.0024)	(0.0021)	(0.0023)	(0.0021)
$Trend_s$	0.0014	-0.0011	0.0031	0.0023	0.0014	0.0006	-0.0002
	(0.0020)	(0.0009)	(0.0024)	(0.0023)	(0.0020)	(0.0020)	(0.0026)
B: Jackknife							
$Shock_f$	0.0041**	0.0006	0.0030	0.0037	0.0041**	0.0045***	0.0051**
	(0.0019)	(0.0013)	(0.0029)	(0.0023)	(0.0018)	(0.0017)	(0.0021)
$\mathbf{I}\{S_f < 0\} \times S_f$	0.0012	-0.0044*	0.0085	0.0041	0.0009	-0.0016	-0.0057
	(0.0039)	(0.0025)	(0.0060)	(0.0051)	(0.0039)	(0.0033)	(0.0040)
$Trend_f$	0.0050**	0.0048***	-0.0029	0.0018	0.0053***	0.0081***	0.0126***
	(0.0022)	(0.0010)	(0.0034)	(0.0024)	(0.0020)	(0.0021)	(0.0024)
$Shock_s$	0.0008	-0.0014	0.0031	0.0017	0.0007	-0.0001	-0.0014
	(0.0022)	(0.0010)	(0.0029)	(0.0024)	(0.0021)	(0.0023)	(0.0021)
$Trend_s$	0.0014	-0.0008	0.0028	0.0020	0.0014	0.0009	0.0001
	(0.0020)	(0.0009)	(0.0024)	(0.0023)	(0.0020)	(0.0020)	(0.0026)
N	25,636	25,636	25,636	25,636	25,636	25,636	25,636

Results show Method of Moments-Quantile Regressions (Panel A) of log(1+total compensation) on firm and industry short-term performance, and firm and industry long-term performance, with executive-firm fixed effects. Panel B shows bias-corrected estimates of coefficients estimated in Panel A. This is done using the split-panel jackknife bias correction by Dhaene and Jochmans (2015), where the panel is split into even and odd years. Standard errors clustered at executive-firm level via bootstrap (200 reps) are in parentheses. *** $p < .01$, ** $p < .05$, * $p < .1$. Full set of controls: Age, age squared, GDP, GDP growth, inflation, and year dummies used. Controls not reported. Data set used is the full unbalanced panel from BoardEx executive compensation database, ORBIS for market value variables, and span years 2003-2013 and 34 countries. GDP growth and GDP per capita are from the World Bank, and inflation is from the IMF database. Observations with zero total compensation or base salary are removed. $\mathbf{I}\{S_f < 0\}$ takes value one if the shock in that year was negative, and zero otherwise.

Table 2.14: MM-QR of log total wealth on firm and industry performance measures

	(1) Location	(2) Scale	(3) Q10	(4) Q30	(5) Q50	(6) Q70	(7) Q90
A: MM-QR							
$Shock_f$	0.0055***	0.0017**	0.0028	0.0041***	0.0056***	0.0069***	0.0079***
	(0.0014)	(0.0008)	(0.0021)	(0.0016)	(0.0014)	(0.0016)	(0.0016)
$Trend_f$	0.0150***	0.0012	0.0131***	0.0140***	0.0151***	0.0160***	0.0167***
	(0.0032)	(0.0013)	(0.0039)	(0.0036)	(0.0031)	(0.0033)	(0.0029)
$Shock_s$	0.0064**	-0.0034**	0.0117***	0.0091***	0.0061*	0.0036	0.0016
	(0.0031)	(0.0015)	(0.0043)	(0.0035)	(0.0031)	(0.0028)	(0.0031)
$Trend_s$	0.0042	-0.0019	0.0071	0.0057	0.0040	0.0026	0.0015
	(0.0034)	(0.0016)	(0.0045)	(0.0037)	(0.0033)	(0.0032)	(0.0038)
B: Jackknife							
$Shock_f$	0.0055***	0.0006	0.0045**	0.0051***	0.0055***	0.0059***	0.0064***
	(0.0014)	(0.0008)	(0.0021)	(0.0016)	(0.0014)	(0.0016)	(0.0016)
$Trend_f$	0.0150***	0.0016	0.0122***	0.0139***	0.0152***	0.0162***	0.0174***
	(0.0032)	(0.0013)	(0.0039)	(0.0036)	(0.0031)	(0.0033)	(0.0029)
$Shock_s$	0.0064**	-0.0026*	0.0108**	0.0082**	0.0062**	0.0045	0.0025
	(0.0031)	(0.0015)	(0.0043)	(0.0035)	(0.0031)	(0.0028)	(0.0031)
$Trend_s$	0.0042	-0.0013	0.0064	0.0051	0.0040	0.0032	0.0022
	(0.0034)	(0.0016)	(0.0045)	(0.0037)	(0.0033)	(0.0032)	(0.0038)
N	22,075	22,075	22,075	22,075	22,075	22,075	22,075

Results show Method of Moments-Quantile Regressions (Panel A) of log(1+total wealth) on firm and industry short-term performance, and firm and industry long-term performance, with executive-firm fixed effects. Panel B shows bias-corrected estimates of coefficients estimated in Panel A. This is done using the split-panel jackknife bias correction by Dhaene and Jochmans (2015), where the panel is split into even and odd years. Standard errors clustered at executive-firm level via bootstrap (200 reps) are in parentheses. *** $p < .01$, ** $p < .05$, * $p < .1$. Full set of controls: Age, age squared, GDP, GDP growth, inflation, and year dummies used. Controls not reported. Data set used is the full unbalanced panel from BoardEx executive compensation database, ORBIS for market value variables, and span years 2003-2013 and 34 countries. GDP growth and GDP per capita are from the World Bank, and inflation is from the IMF database. Observations with zero total compensation or base salary are removed.

Chapter 3

The role of preferences, attitudes, and personality traits in labor market matching *

3.1 Introduction

The matching process between firms and employees in the labor market is a major topic in economics. A bad match between employer and employee may lead to ex-post sorting out of a firm and high costs from rehiring (Eeckhout, 2018). Previous literature has identified firm size, worker skills as well as worker and firm productivity as important drivers of this matching process, for example by Abowd et al. (1999); Lindenlaub (2017). But the role of attitudes, preferences and personality in the matching process is much less explored, despite them gaining increasing interest in the field (Falk et al., 2018).

Previous empirical research is mainly based on lab studies showing, for instance, the effect of individual attitudes and personality traits on contract choices between subjects (Bartling et al., 2009; Dohmen and Falk, 2011). Lazear et al. (2012) find sorting explains lab participants' sharing behavior better than demographics or education. But Levitt and List (2007) argue that generalizing from lab experiments on social preferences is difficult, making cross-validation with field and survey evidence necessary. In a field experiment, Bellemare and Shearer (2010) provide evidence for risk-sorting of workers, where risk-tolerant workers sort into high-risk firms. Previous studies using representative survey data on the role of preferences and personality for matching are scarce. Focusing on one variable, Grund and Sliwka (2010) show that risk preferences cross-sectionally correlate with sorting into compensation schemes. We address this gap by using representative matched employer-employee data that includes a variety of measures of individual attitudes, preferences and personality to answer the question whether employers affect workers' behavioral characteristics. In detail, we ask: Do time-constant, firm-specific characteristics correlate with worker preferences and personality? And if yes, how strong is the explanatory power relative to other control variables, i.e. are they important drivers of an employer-employee job match?

3.2 Data and methodology

Our data are from the Linked Personnel Panel (LPP), a longitudinal employer-employee survey. The survey is representative of German private establishments with at least 50 employees (in the first survey wave).[2] The stength of the LPP, is that is contains employee-level information, for example about attitudes, preferences and personality, which is linked to establishment-level information, for example on structural characteristics.

The LPP contains information on more than 7,000 randomly drawn employees aged between 18 and 74 working in 700 to 1,200 establishments in three survey waves 2012, 2014 and 2016.[3] As outcome variables, we use a series of validated measures capturing fundamental determinants of human behavior (individual attitudes, preferences and personality traits), which are common in behavioral economics and applied psychology research. In detail, we include the Big Five personality traits, risk attitude, reciprocity, altruism, time preferences (discounting), trust, affective commitment to the organization, inequity aversion, and helpfulness.

---

* This chapter is based on Haylock and Kampkötter (2019b)

[2] DOI: 10.5164/IAB.LPP1617.de.en.v1.

[3] As this chapter was written, and published prior to the availability of the fourth LPP survey wave, it only includes the first three waves of data

Our survey measures of individual attitudes and preferences have been experimentally validated, which is to say they have been shown to be very suitable predictors for actual behavior in incentivized laboratory experiments. These measures have been applied, for instance, in the Global Preference Survey or the German Socio-Economic Panel Study (Dohmen et al., 2009, 2011; Falk et al., 2016, 2018). For measuring commitment and the Big Five personality traits, we apply validated and commonly applied constructs. Section 3.5 reports the origin and translated wording of validated survey items. Table 3.1 reports summary statistics for our outcome variables. Summary statistics of independent variables on demographics, education, and job and establishment characteristics are shown in Table 3.3 of Appendix section 3.6.

Table 3.1: Summary statistics of individual outcome variables

Variable	Mean	SD	Min	Max
Attitudes and preferences				
Risk attitude	5.69	1.83	0	10
Trust	3.29	0.69	1	5
Helpfulness	4.27	0.71	1	5.5
Altruism	7.66	1.52	0	10
Positive reciprocity	1.47	0.61	1	5
Negative reciprocity	4.09	0.99	1	5
Affective commitment	3.71	0.88	1	5
Inequity aversion	2.50	1.03	1	5
Time preference (discounting)	2.52	1.18	1	5
Big Five personality traits				
Extraversion	3.70	0.73	1	5
Neuroticism	2.7	0.78	1	5
Conscientiousness	4.38	0.48	1.33	5
Openess to experience	3.67	0.64	1	5
Agreeableness	4.07	0.58	1	5

To identify to what extent matching is driven by worker attitudes, preferences and personality, we run nested pooled OLS regressions where the dependent variable is an employee i's attitude, preference or personality trait. In each specification, we add a new set of variables and observe the change in adjusted R^2 as well as the F-statistic of the added group. Specification (1) only controls for education (secondary and tertiary education levels), specification (2) then adds job characteristics (white collar, managerial responsibility, part time, monthly net wage, permanent contract), specification (3) further adds individual demographics (age, female, permanent relationship, household size, health status), specification (4) adds industry, region, and size fixed effects for each establishment j (X_j) as well as year fixed effects for each wave t (T_t). Finally, specification (5) adds establishment fixed effects (while excluding industry, size and region fixed effects).

Hence, our final regression specification is the following:

$$y_{it} = \alpha + \beta_1 educ_{it} + \beta_2 job_{it} + \beta_3 demog_{it} + \beta_4 X_j + \beta_5 T_t + \beta_6 establishment_j + \varepsilon_{it} \quad (3.1)$$

Our final specification tests whether the adjusted R^2 will significantly increase when adding establishment fixed effects. If the incremental explanatory power of time-constant establishment characteristics is higher than all other control variables that can potentially explain differences in individual outcome variables, firm characteristics would constitute an important match-specific component.[4]

[4] Studies also analyzing the effects of adding firm fixed effects into individual-level regressions include, for instance, Kampkötter (2015) and Grund and Hofmann (2019).

3.3 Results

Table 3.2 reports incremental changes in adjusted (adj.) R^2 when adding the different sets of controls for the variety of individual outcome variables. Note that the main effects do not change if the independent variables are added in a different order.[5] Column (5) shows the incremental change in adjusted R^2 when adding establishment fixed effects. To further illustrate the influence of establishment characteristics in explaining total variation in our behavioral outcome variables, column (6) quantifies the relative importance of establishment fixed effects, by dividing incremental adj. R^2 from column (5) by total adj. R^2 from all previously included covariates. The results show that the incremental change in adjusted R^2 is highest for establishment fixed effects for all of the attitudes and preferences outcome variables (column (5)), suggesting that time-constant firm characteristics explain the largest amount of variation in these outcomes. For all personality traits except openness, we observe a lower, but still reasonably high importance of establishment fixed effects in explaining total variation in behavioral outcomes. Looking at column (6), establishment fixed effects explain helping behavior (67%), trust (62%), positive reciprocity (60%), risk attitude (54%) as well as openness to experience (51%) best. This supports the notion that time constant establishment characteristics seem to play an important role in determining the worker-firm match. Further, helping is explained best by establishment fixed effects, relative to other outcomes, suggesting there is a large firm specific black-box to be explored.

Table 3.2: Incremental changes in adjusted R^2 when adding further covariates

	(1) Educ.	(2) + Job	(3) + Demographics	(4) + Industry, region, size, and year FE	(5) + Establishment FE	(6) ΔR^2 adj.	(7) ΔR^2 adj. (in %)	(8) N
Dep. var.:								
Attitudes and pref.								
Risk attitude	0.006**	0.013**	0.007**	0.005**	0.037**	0.068	54.4	13,599
Trust	0.003**	0.001	0.006**	0.019**	0.049**	0.079	62.0	13,577
Helping	0.002**	0.003**	0.013**	0.003**	0.040**	0.060	66.7	13,552
Altruism	0.002*	0.001*	0.021**	0.001	0.024**	0.049	49.0	6,508
Positive recipr.	0.003**	0.000	0.003**	0.000	0.009**	0.015	60.0	6,516
Negative recipr.	0.002	0.006**	0.014**	0.004**	0.023**	0.049	46.9	6,498
Commitment	0.012**	0.045**	0.039**	0.008**	0.086**	0.191	45.0	13,461
Inequity aversion	0.005**	0.011**	0.010**	0.007**	0.017**	0.051	33.3	10,482
Time preference	0.016**	0.005**	0.003**	0.003**	0.018**	0.045	40.0	6,501
Big Five								
Extroversion	0.001*	0.011**	0.026**	0.004**	0.014**	0.056	25.0	10,502
Neuroticism	0.007**	0.017**	0.065**	0.003**	0.015**	0.107	14.0	10,520
Conscientiousness	0.022**	0.006**	0.025**	0.002**	0.014**	0.069	20.3	10,520
Openness	0.005**	0.005**	0.006**	0.002**	0.019**	0.037	51.4	10,397
Agreeableness	0.005**	0.008**	0.014**	0.002**	0.016**	0.046	34.8	10,510

Results show the incremental change in the adjusted R^2 from regression results of preferences, attitudes, and Big Five on covariates, additionally included in a step-wise fashion in each column. The stars show the significance of an F-test of the set of additional covariates included in the respective column. Data used are individual responses from waves 1-3 of the Linked Personnel Panel. Standard errors clustered at the establishment level. (***) p<0.01,(**) p<0.05,(*) p<0.1.

[5] In line with Lazear et al. (2012), this is tested using the partial adj. R^2 of each explanatory variable set, $(\mathrm{R}^2 - R_{(i)}{}^2)/(1 - R_{(i)}{}^2)$.

3.4 Discussion and conclusion

We provide first evidence of worker-firm matching based on preferences, attitudes and personality traits using representative matched employer-employee data. Time-constant establishment characteristics, holding individual and other firm characteristics constant, explain a significant proportion of total variance in a series of behavioral outcome variables.

A possible mechanism behind these findings is labor market sorting, which contributes to our understanding why employees are heterogeneous across firms regarding social preferences. If firms screen on employee preferences such as cooperation, they set contracts that will only attract the desired type of worker (Kosfeld and von Siemens, 2011). Further, individuals in the workplace can become more similar over time. This is particularly important in the case where preferences are to a certain degree endogenous. Further explanations might include the similar-to-me effect (Rand and Wexley, 1975), a selection of employees similar to the employers themselves. Similarly, peer effects could influence personal preferences and attitudes over time through learning and changes in behavior (Sacerdote, 2001). Further research is needed to explore the detailed channels-to what extent matches are driven by sorting, learning, similarity, peer effects, or other potential channels, which our data thus far do not allow to causally identify.

3.5 Description of behavioral outcome variables

Risk attitude is measured with the single item adapted from the individual questionnaire of the SOEP (Richter et al., 2013). The wording is: "Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?" on a 11-point Likert scale ranging from 0 (risk averse) to 10 (fully prepared to take risks).

Altruism is measured with a single-item construct using the following wording: "How do you assess your willingness to share with others without expecting anything in return?"
Please assess your willingness on a scale with 0 meaning: "not at all willing to share without expecting something in return" and 10 meaning: "very willing to share without expecting something in return". The values in between allow you to grade your assessment." (Falk et al., 2016, 2018). The respective answering scale ranges from 0 to 10.

Time preferences (discounting) are measured using the following two items on a five-point Likert scale ranging from 1 (does not apply to me at all) to 5 (applies to me perfectly) (modified from Falk et al. (2016, 2018):
A: I abstain from certain things today so I can afford more tomorrow.
B: I tend to procrastinate things even though it would be better to do them now.

Reciprocity (positive and negative) is measured using the following two items on a five-point Likert scale ranging from 1 (does not apply to me at all) to 5 (applies to me perfectly) (Falk et al., 2016, 2018):
A: If someone tries to harm me on purpose, I will try to pay them back in kind even if this is associated with costs for me.
B: If someone does me a favor, I am prepared to return it.

Trust is measured with two of three trust items included in the German SOEP Study (Naef and Schupp, 2009). Both items are measured on a five-point Likert scale from "totally agree" to "totally disagree".

Affective commitment to the organization is measured using the six-item short scale by Meyer et al. (1993) on a five-point Likert scale.

Inequity aversion is based on two items of the USS-8 justice sensitivity inventory (Schmitt et al., 2010). We apply one item from the victim and one item from the beneficiary scale that reflect disadvantageous and advantageous inequity.

Helpfulness (helping) asks whether employees help colleagues, and are helped by colleagues based on the following two items (5-point Likert scale ranging from 1 "Always" to 5 "(Almost) never"):
A: How often do you receive help and support from colleagues if required so?
B: How often do you offer help to your colleagues?

We measure **Big Five** personality traits with the 16-item version of the Big Five Inventory short scale developed for the Socio-Economic Panel (SOEP) (Gerlitz and Schupp, 2005; Lang et al., 2011) on a five-point Likert scale ranging from 1 (does not apply to me at all) to 5 (applies to me perfectly).

More details on the items, constructs and their internal validity are provided in Kampkötter et al. (2016).

3.6 Summary statistics of individual-level explanatory variables

Variable	**Mean**	**SD**	**Min**	**Max**
Education				
No qualification	0.005	0.07	0	1
Lower secondary school certificate	0.24	0.43	0	1
Intermediate secondary school certificate	0.44	0.50	0	1
University of applied sciences entrance qualification	0.10	0.30	0	1
University entrance diploma (A-level)	0.21	0.40	0	1
Another level of education	0.007	0.08	0	1
No training qualification	0.02	0.15	0	1
Apprenticeship	0.49	0.50	0	1
Vocational training within the education	0.10	0.30	0	1
Master craftsmen or technical college	0.20	0.40	0	1
University of applied sciences degree	0.09	0.28	0	1
University degree	0.10	0.29	0	1
Other training qualification	0.004	0.06	0	1
Job				
White collar employee	0.60	0.49	0	1
Leadership position	0.30	0.46	0	1
Part time contract	0.13	0.34	0	1
Monthly net income	2,239.86	1,330.47	25	60,000
Fixed-term contract	0.05	0.22	0	1
Individual demographics				
Age	46.25	10.46	18	69
Female	0.28	0.45	0	1
Health status	2.36	0.94	1	5
Permanent relationship	0.84	0.37	0	1
Household size	2.78	1.22	1	13
Industry				
Manufacturing industry	0.32	0.46	0	1
Metal and electrical industry, automotive sector	0.37	0.48	0	1
Commerce, traffic, communication	0.11	0.31	0	1
Company-related services, financial services	0.13	0.33	0	1
IT, communication and other services	0.07	0.26	0	1
Region				
North	0.17	0.37	0	1
East	0.27	0.44	0	1
South	0.26	0.44	0	1
West	0.30	0.46	0	1
Establishment size				
20 to 49	0.006	0.08	0	1
50 to 99	0.13	0.34	0	1
100 to 249	0.25	0.43	0	1
250 to 499	0.24	0.43	0	1
500 and more	0.37	0.48	0	1
Year				
2012	0.39	0.49	0	1
2014	0.34	0.47	0	1
2016	0.27	0.44	0	1

N=13,599 (largest sample available for one of the dependent variables)

Chapter 4
Helping and antisocial behavior in the workplace *

4.1 Introduction

If employees help each other, firms can benefit (Hamilton et al., 2003). If employees behave antisocially, this is harmful not only to employee motivation, but also to firm performance (Gangadharan et al., 2020). But what explains helping and antisocial behavior within firms, as well as observed differences across firms? Why do some firms benefit from a high willingness among employees to help each other, while other firms do not, or only less so? Does "good" leadership promote helping and reduce antisocial behavior among employees? How important are employees' personality traits, as measured by the Big Five, and economic preferences such as altruism and reciprocity in this relationship? In this paper, we analyze employee helping and antisocial behavior as an integral part of a firm's organizational culture and working climate (Alan et al., 2021). Based on a unique data set spanning many firms in different industries, we offer a comprehensive analysis of the organizational and individual foundations of these important forms of behavior in the workplace. In doing so, we use a rich set of validated, and commonly used constructs including data on employees' personality traits, economic preferences as well as leadership quality.

Our analysis is based on a novel, representative employer-employee panel data set of larger private establishments in Germany that is particularly appropriate to answer the above research questions.[2] First, our data contain explicit information from employees about mutual helping, as well as antisocial behavior in their firm, which allows us to construct reliable indices for both kinds of behavior at the firm level, in each survey wave. Second, the data include established and validated survey measures of the quality of leadership (focusing on supervisors' trust, understanding, and fairness), employees' personality traits (Big Five), as well as trust and economic preferences (in particular social preferences, but also risk attitude and time discounting) that enable us to investigate to what extent differences along these dimensions explain differences in helping and antisocial behavior. Our survey design allows us to differentiate between time-variant and time-constant drivers of helping and antisocial behavior. Leadership quality is elicited in every survey wave, because leaders often rotate between teams, and might develop and adjust their leadership skills over time. As a result, employees typically face leaders of different quality over their careers, which calls for a continuous measurement of leadership quality. On the other hand, personality traits and economic preferences are typically much less affected by workplace events, and are considered as rather stable within individuals over our measurement period (Cobb-Clark and Schurer, 2012; Schwaba and Bleidorn, 2018; Fitzenberger et al., 2021). Hence, this information is elicited only when an employee is surveyed for the first time. Importantly, our preference measures are based on survey items that have been behaviorally validated for Germany, i.e., that are predictive for economically relevant behavior both at the individual and organizational level (cf. Falk et al. (2016, 2018) and references therein). Finally, we exploit a rich set of employee- and establishment-level data to control for important differences in, for example, ability, task interdependencies, industry, firm size, and management practices (Bloom and Van Reenen, 2007).

To the best of our knowledge, our study is the first to combine such rich and complementary employer-employee data to uncover the economic and behavioral foundations of helping and antisocial behavior in the workplace. Our data cover the period from 2012 to 2019 in four survey waves, with about 800 to 1,200 firms and 6,500-

* This chapter is based on joint work with Patrick Kampkötter, Michael Kosfeld, and Ferdinand von Siemens, not yet published.

[2] In our data, an establishment denotes a regionally and economically separate unit, in which employees liable to social security work (Fischer et al., 2009). In principle, a firm might therefore comprise of several establishments. For simplicity and whenever there is no possibility of confusion, we use the terms "firm" and "establishment" interchangeably.

8,000 employees participating in each wave. Importantly, sampled firms are drawn in a stratified manner to ensure the sample is representative for firms with more than 50 employees in the private sector. A random sample of employees working within the surveyed establishments are interviewed outside their workplace, typically in the evening. Because we are interested in workplace cultures of helping and antisocial behavior in a given firm, we aggregate the data on the level of the establishment, i.e., our units of observation are establishment-level averages of a particular survey item (or measure constructed with multiple individual items) in a given wave.

Our results document a large variation in helping and antisocial behavior across firms. The distribution of helping is left-skewed, with the modal firm displaying a helping level of 4 (on a scale from 1 = "never or nearly never" to 5 = "always"). The distribution of antisocial behavior is right-skewed with a modal value of 2. As one might expect, helping and antisocial behavior are negatively correlated on the firm level, but the size of the correlation is modest ($\rho = -0.22$). In fact, there exists a sizable share of firm-wave observations (17%) that show above-median levels in *both* types of behavior. These observations indicate that helping and antisocial behavior are rather distinct concepts, and one is not just the opposite of the other.

But what explains the firm-level heterogeneity? Our first finding shows that leadership quality plays a major role. A one SD increase in leadership quality is associated with a 0.17 SD increase in helping in our preferred specification, and a 0.41 SD reduction in antisocial behavior, respectively. As our leadership measure varies over time, we can include firm fixed effects. Our results show that the association is robust to this. Next, differences in time-constant social preferences and Big Five personality traits are key predictors of helping and antisocial behavior in the workplace. For example, a one SD increase in the level of altruism or positive reciprocity in an establishment is associated with a 0.06 and 0.05 SD higher level of helping, respectively. With regard to antisocial behavior, we find that personality traits are relatively more important, in particular, neuroticism, and openness to experience play a significant role. Here, a one SD increase in each factor is associated with a 0.10 and 0.05 SD increase in antisocial behavior, respectively. Intriguingly, social preferences, while predicting helping, do not explain differences in antisocial behavior. Finally, employees' level of general trust is a strong and significant predictor of both helping and antisocial behavior in the workplace. A one SD higher level of trust is associated with a 0.14 SD increase in helping behavior, and a 0.13 SD decrease in antisocial behavior.

Our results connect and add to several important strands in the literature. With respect to interpersonal interactions in the workplace, a number of recent papers have emphasized the importance of workers' "social" or "people skills" for firm and labor market outcomes (Borghans et al., 2008, 2014; Deming, 2017). Based on personnel data from a large US high-tech firm, Hoffman and Tadelis (2021) show that managers with higher employee ratings in "people management skills" are associated with lower employee turnover, and rewarded for this in terms of higher promotion rates and larger salary increases. Englmaier et al. (2021) show in a large-scale field experiment that encouraging teams to select leaders, has positive effects on team performance. Our measure of leadership, considering a leader's fairness, trust, and understanding towards her employees, can be seen as a proxy for similar social skills to manage people in a given firm. We hence complement the above results, by showing in a representative panel of German firms that good people management significantly contributes to explaining firm-level differences in employees' helping and antisocial behavior in the workplace. As both behaviors are significantly associated with employee turnover intention (and other outcomes) in our data (results shown in Table 4.24 in Section 4.7.9), our results further suggest a potential mediating role of helping and antisocial behavior, or more generally workplace climate, in this association.

Next, with respect to the role of employee attitudes, seminal research by Heckman et al. (2006), Borghans et al. (2008), and Almlund et al. (2011) has shown that differences in personality traits, as measured by the Big Five, play a significant role in economic behavior, in particular in the labor market. We add to this research by showing that heterogeneity in the level of helping and antisocial behavior between firms can also be partly attributed to personality differences in their respective workforce. Becker et al. (2012) consider the role of economic preferences, as elicited in incentivized laboratory experiments, and find that measures of personality traits and economic preferences are rather complementary when it comes to explaining heterogeneity in economic behavior. Our results corroborate this view by identifying important empirical relations between personality traits and preferences on the one hand, and helping and antisocial behavior on the other hand. For example, our results show that trust and social preferences significantly explain helping, while only some personality traits correlate with this behavior. At the same time, personality (neuroticism, openness) explains antisocial behavior, whereas social preferences do not. This suggests that helping and antisocial behavior are distinct behaviors that are also influenced by different individual traits and preferences.

Our empirical findings are closely related to the theoretical analysis of sorting and self-selection of employees with heterogeneous social preferences in the labor market (Kosfeld and von Siemens, 2009, 2011). While our data do not allow us to identify sorting behavior explicitly, Haylock and Kampkötter (2019b), using the same data, show that the distribution of employee types across firms is consistent with self-selection according to employees' attitudes and preferences. Kosfeld and von Siemens (2009, 2011) then predict that firm outcomes in terms of cooperation and helping behavior among employees should differ, even within the same industry, and correlate with measures of social preferences on the firm level. This is exactly what the results in this paper show.[3]

Finally, the paper connects to a classic literature going back to at least Alchian and Demsetz (1972), Holmstrom (1982), FitzRoy and Kraft (1986), Drago and Turnbull (1988), Drago and Turnbull (1991), Itoh (1991, 1992), Kandel and Lazear (1992), and Rotemberg (1994) that investigates the role of team production, mutual helping, and cooperation among workers, focusing largely on the design of incentives to induce efficient effort and production decisions.[4] Lazear (1989) also considers the problem of antisocial behavior between workers, such as sabotage. This important theoretical work has been complemented by a number of empirical studies analyzing the effect of incentives on teamwork and cooperation with single firm case studies or employer-employee data sets (Drago and Garvey, 1998; Knez and Simester, 2001; Berger et al., 2011; Friebel et al., 2017; Deversi et al., 2020; Delfgaauw et al., 2022) or documenting differences in helping behavior between individual firms from a given industry (Gittell et al., 2004; Encinosa et al., 2007). Our study contributes to this literature by providing the first comprehensive empirical analysis of the determinants of helping and antisocial behavior in the workplace, based on a representative sample of firms from a large economy, including rich information on key firm- and employee-level variables. Besides employee incentives, our data consider robust information about employees' "cooperative attitudes", i.e., social preferences, enabling us to address theoretical predictions of Lazear (1989), Kandel and Lazear (1992), and more recently Kosfeld and von Siemens (2009, 2011).

The paper is organized as follows. Section 4.2 provides a detailed description of the data we use for our analysis. In Section 4.3 we document the evidence for helping and antisocial behavior across firms and derive our main empirical results. Finally, Section 4.4 concludes.

4.2 Data

4.2.1 The Linked Personnel Panel

Our results make use of a unique data set, the Linked Personnel Panel (LPP), which constitutes a longitudinal linked employer-employee panel data set that is representative for German establishments in the private sector with at least 50 employees liable to social security (Kampkötter et al., 2016; Haylock and Kampkötter, 2019a; Müller and

[3] Other work documenting a positive association between social preferences and the level of cooperation in a natural field context includes Rustagi et al. (2010). Krueger and Schkade (2008) show that workers who are more gregarious tend to be employed in jobs that involve more social interactions.

[4] Relatively more recent work includes Dur and Sol (2010) and Ishihara (2017).

Wolter, 2020).[5] The LPP links important variables on the establishment level with rich employee-level information comprising key worker as well as job characteristics. We analyze the four available survey waves from 2012/13, 2014/15, 2016/17, and 2018/19.

The employer survey of the LPP covers between 769 and 1,219 establishments per wave. The sampling started in 2012 with the establishment survey, which was drawn from the IAB establishment panel wave of 2011, a large-scale annual survey of nearly 16,000 German establishments. To ensure that the data set is representative, a stratified disproportionate random sampling approach was used, where establishments were randomly drawn from a matrix stratified by business sector, establishment size, and region.[6] From an adjusted gross sample of 1,705 establishments, 1,219 valid establishment interviews could be realized leading to a response rate of 72%.

The first LPP employee survey was then launched in December 2012 based on a selection of establishments that had been interviewed in the preceding LPP employer survey. The main selection criteria were the stated willingness of establishments to participate in the 2014 wave of the LPP employer survey, and a workforce of at least 50 employees liable to social security contributions as documented in the administrative data. As a result, the total population comprised 300,881 employees from 869 establishments, from which a sample of 37,831 employees was randomly drawn in a disproportionate manner, stratified by establishment size to include not more than 10% of an establishment's workforce in the survey. To avoid that the survey was dominated by few, large establishments, larger establishments had smaller sampling probabilities. Importantly, the sampling of employees based on administrative social security data mitigates possible selection effects often found in survey data research. In each wave, roughly between 6,500 and 7,500 individuals aged between 18 and 74 are interviewed at home via telephone (CATI) or web interface (CAWI). These interviews take place outside the work environment at different dates, typically in the evening, ensuring that respondents working in the same firm are interviewed independently of each other.

In later survey waves 2 to 4, the LPP sample consists of two groups. First, the employee survey primarily targets panel cases, i.e., individuals, who were surveyed in the previous wave and explicitly expressed their consent to be surveyed again. Furthermore, they need to work in an establishment with a valid LPP employer survey interview in the corresponding wave. The second group comprises a refreshment sample. Here, employees from panel establishments are oversampled in case only a few or no employee interviews were available in the previous wave. For employees whose establishments are new to the LPP survey, a sample is drawn as described above. The data show that, on average, 39 percent of employee-wave observations come from the refreshment sample.

In both surveys, response rates are comparatively high: 79% for the employer and 57% for the employee survey on average. Moreover, there are no significant selection effects on panel participation. This, together with its careful implementation ensuring representativity on the firm level, as well as the use of established survey items on the employee level, make the LPP an ideal data source for our research question. For a further detailed description of the LPP, see Kampkötter et al. (2016).

4.2.2 Survey measures

Our survey measures are based on a rich set of validated and commonly used constructs, either from experiments, surveys, or management research (Patterson et al., 2005; Kim and Leung, 2007; Falk et al., 2016, 2018). In the following, we describe these measures in detail and document their source. An overview on all employee- and establishment-level survey items including their original wording is presented in Tables 4.3 and 4.4 in the Appendix. Furthermore, summary statistics are provided in Table 4.5 in the Appendix.

[5] Besides public administration, charity organizations, agriculture, forestry, and fishery are also excluded. The data set is available via the Research Data Centre (FDZ) of the German Federal Employment Agency at the Institute for Employment Research (IAB). The DOI is: 10.5164/IAB.LPP1617.de.en.v1.

[6] Details on the sampling matrices are provided in Bellmann et al. (2015).

4.2.2.1 Helping and antisocial behavior

Helping is measured with two items in the LPP employee survey: *offering help* to colleagues (“How often do you offer your colleagues help?”), and *receiving help* from colleagues (“How often do you receive support or help from your colleagues if you ask?"). *Antisocial behavior* is measured with one item referring to corresponding behavior by both colleagues and superiors (“How often do you feel wrongly criticized, harassed or denounced by your colleagues or superiors?”). Both items of help received and antisocial behavior are based on the Copenhagen Psychosocial Questionnaire (COPSOQ), which has been used in more than 40 countries (Burr et al., 2019). For all three items, a 5-point Likert scale is applied with the following response categories ranging from 1 to 5: “never or nearly never”, “seldom”, “sometimes”, “often”, or “always”. *Helping* and *antisocial behavior* are asked repeatedly over all four survey waves of the LPP.

The data show that the two helping items are strongly and positively correlated. The Spearman correlation coefficient between help offered and help received at the individual level across all waves is 0.54, and significant at the 1% level. The correlation remains very stable over time, ranging from 0.52 to 0.55. A large share of respondents engage in *mutual help*, i.e. they both offer help when asked and receive help if they ask for it. Precisely, we observe a value of at least 4 (“often") in 86.7% of person-year observations with respect to help offered, 85.7% with respect to help received, and 82.3% with respect to the average of the two items. Since individual helping behavior is closely linked to the helping behaviors of those an employee interacts with, we take the equally-weighted average of the two helping items at the individual level as our *helping index* for respondent i in wave t.

The Spearman correlation between the helping index and antisocial behavior at the individual level is significant at the 1% level, but at a rather modest level of -0.26 across all waves. The correlation is stable over time, ranging from -0.23 to -0.27, all significant at the 1% level. Thus, while correlated, the two constructs of helping and antisocial behavior seem to measure substantively different dimensions of interpersonal behavior at the workplace. They are not direct mirrors of one another.

4.2.2.2 Leadership quality

Employees’ assessment of leadership quality in their establishment is also elicited repeatedly in each survey wave. Similar to Hoffman and Tadelis (2021), we measure good leadership by constructing an equally-weighted *leadership index*, in our case based on three items: *supervisor trust* and *supervisor understanding* are derived from the perceived supervisory support scale of the well-established Organizational Climate Questionnaire (Patterson et al., 2005), and *supervisor fairness* is from the interactional justice scale (Kim and Leung, 2007).

4.2.2.3 Personality traits, social preferences, and trust

To assess an employee’s personality, we apply the *Big Five personality traits* (Costa and McCrae, 1995), which are measured by the Big Five inventory short scale (16-item variant) as in the German Socio-Economic Panel (SOEP) (Gerlitz and Schupp, 2005). This scale has been used, for example, by Dohmen et al. (2008, 2010) and Becker et al. (2012).

Based on prior literature, we apply a set of economic preferences commonly used in behavioral economics research (Becker et al., 2012; Falk et al., 2016) that have been behaviorally validated for Germany both in the lab and in the field. We elicit *positive* and *negative reciprocity* (Falk and Fischbacher, 2006; Dohmen et al., 2009) by two items from the Preference Survey Module (PSM) of Falk et al. (2016), where one of these items also used in the SOEP. *Altruism* (Eckel and Grossman, 1996) is similarly elicited by a single item from the PSM. Finally, we measure *trust* (Fehr, 2009) by two of the three items that are commonly used in the SOEP (Dohmen et al., 2008; Naef and Schupp, 2009).

All personality traits and preferences are only asked once, when an employee is surveyed for the first time, as these variables are considered to be rather stable for working-age adults (Cobb-Clark and Schurer, 2012; Schwaba and Bleidorn, 2018; Fitzenberger et al., 2021). For the regression analyses, we impute these values into all subsequent waves where the same individual is observed.

4.2.2.4 Controls

Our empirical analysis takes into account a rich set of control variables on both the employee and the establishment level.

On the employee level, a key challenge is to control for differences in *employee ability* as these are likely to be associated with both individual demand and supply of helping. In addition to using information on employee education (see below), a notable feature of our data is that we have access to individual AKM/CHK fixed effects (Abowd et al., 1999; Card et al., 2013). These individual fixed effects, which have also been used by Bender et al. (2018), are calculated using the full sample of German social security data (the IAB Employment History File (BEH)) for the period 2010-2017 (hence, extending the procedure by Card et al. (2013)), which covers most of our sample period (Bellmann et al., 2020).[7] The estimated fixed effects are imputed across all individual observations prior to aggregation.

We next include information about an employee's level of *education*. Precisely, we include dummies for six school education categories (no school certificate, 9th grade (Hauptschule), 10th grade (Realschule), university of applied sciences entrance qualification (Fachhochschulreife), higher education entrance qualification (Abitur), and other) and seven vocational and university educational categories (none, apprenticeship, trade school, master craftsman, university of applied sciences degree, university degree, and other). Further, we control for *gender*, *age* categories (under 30, 30 to 40, 40 to 50, and above 50), an indicator for having a *life partner*, and an indicator for *living alone*. We also control for employees' *risk attitude*, *time preferences*, and *self-efficacy*. Risk attitude is measured using a single SOEP item that has been shown to predict risk-taking in experimental lottery choices (Dohmen et al., 2011). Time preferences are operationalized via two items from the PSM (Falk et al., 2016). Self-efficacy is elicited by the ASKU self-efficacy scale from Beierlein et al. (2013), which includes three items measured on a five-point Likert scale.

To control for differences in employees' job characteristics, we include individual-level information about *task interdependencies*, which is elicited by two items asking whether an employee's tasks depend on the input of colleagues, and whether colleagues' tasks depend on an employee's own task fulfillment. Finally, we include an *interview method* dummy (CATI vs. CAWI).

On the establishment level, our controls include relevant information about *industry* (manufacturing; metal, electronics, and automotive; retail, logistics, and media; company-related and financial services; IT, communication, and other services), *region* (east, south, west, north), *establishment size* (less than 50 employees, 50 to 99 employees, 100 to 249, 250 to 499, and more than 500)[8], as well as *ownership* type (family firm, management, investor or dispersed ownership, (partly) state-owned, and other). In addition, we include survey wave fixed effects.

Next, establishment managers provide information about the use of *human resource management (HRM) practices* in each wave of the employer survey. These items closely follow the spirit of recent, large-scale management surveys such as the World Management Survey or the Management and Organizational Practices Survey (Bloom and Van Reenen, 2007, 2010; Bloom et al., 2010, 2013; Bender et al., 2018; Bloom et al., 2019). We include

[7] Detailed information about the LPP-ADIAB is provided via DOI: 10.5164/IAB.FDZD.1907.en.v1.

[8] Some of the smallest establishments may shrink in size over time and are then allocated to the category "less than 50 employees".

dummy variables for the existence of performance appraisal systems, written target agreements, employee feedback talks, (career) development plans, annual employee surveys, the use of a code of conduct, and the importance of a good working atmosphere for retaining employees. We further control for an index variable called consequence management, which is the equally-weighted average of four items measuring the way in which poor performance is addressed and the type of consequences that result from persistent underperformance.

Finally, a potentially relevant management practice is *performance pay*. We construct variables both for managerial and non-managerial employees taking the value one if the establishment uses variable pay that depends on firm performance and/or team performance and/or individual performance, and zero otherwise. We also control for the share of non-managerial staff that receive performance pay to proxy for the general importance of incentive pay in a given establishment.[9]

4.3 Results

In our empirical analysis, we aggregate all individual-level variables on the establishment level for each survey wave. We use the terms *establishment* and *firm* interchangeably whenever there is no possibility of confusion. Hence, our unit of observation is establishment, or firm, f in survey wave t. These equally-weighted establishment-wave-averages are calculated for establishments with at least three observations per wave.[10] By doing so, we not only reduce measurement error that may occur at the individual level, such as common method bias, but we explicitly take an organizational-culture perspective on the manifestation and foundations of helping and antisocial behavior. All variables of interest are standardized with mean zero and unit variance at the firm level in the regression analyses.

We first document the variation in helping and antisocial behavior across firms based on our representative data. We then come to our main research question: What explains these firm-level differences in helping and antisocial behavior? In the remainder of this chapter, we present our empirical strategy, followed by the main results.

4.3.1 Helping and antisocial behavior across firms

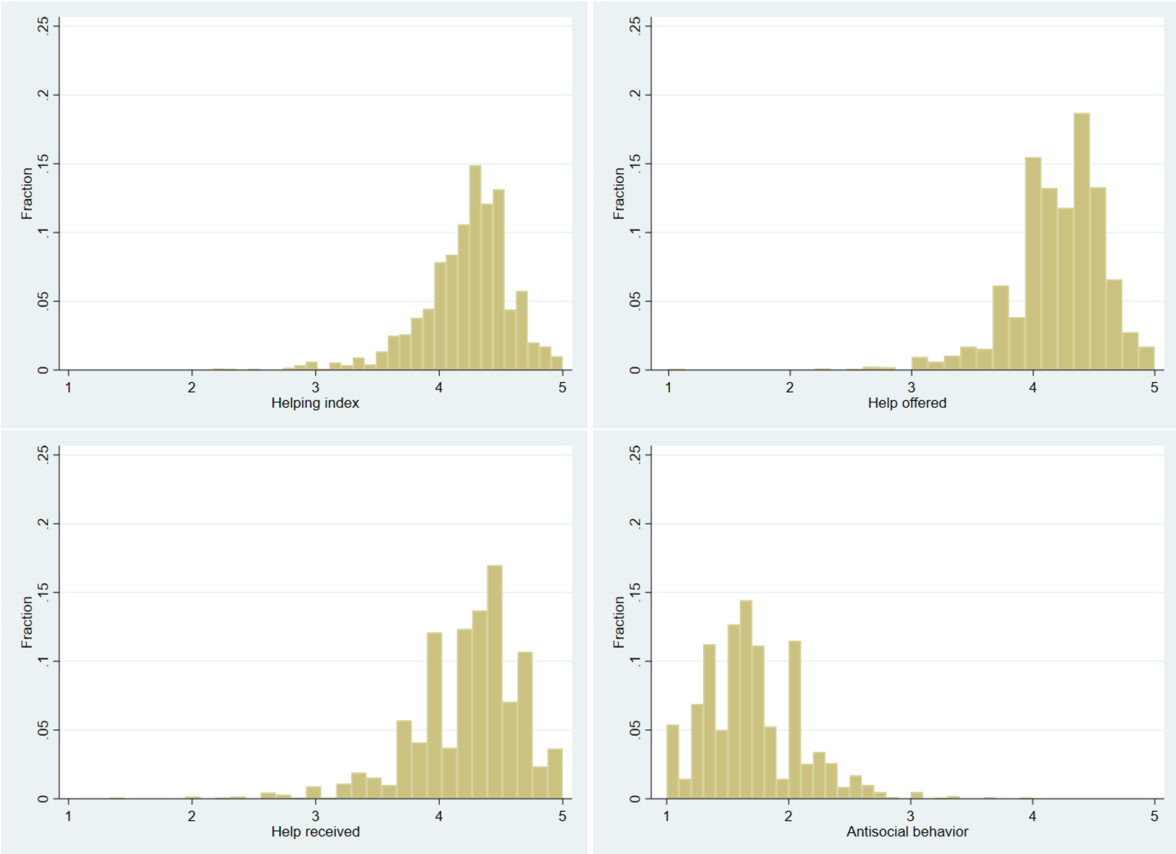
Figure 4.1 shows the distribution of firm-wave averages of our helping index, the single items on help offered and help received, as well as antisocial behavior. The distribution of the helping index and the underlying helping items is left-skewed, with the modal firm lying between 4 (“often”) and 5 (“always”). As the figure shows, there is substantial heterogeneity in helping behavior across firms. For antisocial behavior, the distribution is right-skewed and most firm averages range between 1 (“never or nearly never") and 2 (“seldom”). Still, the observed heterogeneity seems striking.[11]

[9] We also include dummies for all missing covariates at the firm level, most of which result from items not elicited in the first survey wave. In these cases, the associated variables are set to the same value for the missing values. Thus, our data make use of all firm-wave observations where helping or antisocial behavior was measured from survey waves 1 to 4.

[10] The number of employees sampled per establishment is increasing in establishment size. At the 10th percentile, we observe about 3 employees per establishment-wave cell, at the 50th percentile 6, at the 90th percentile 18, and at the 99th percentile 38 employees. The average number of employees per establishment-wave cell is 9.

[11] In Figures 4.6 and 4.7 in the Appendix, we show the distributions of the helping index and antisocial behavior by wave. For helping, Kolmogorov-Smirnov tests reveal significant differences between waves 1 and 2 ($p < 0.0000$), as well as between waves 3 and 4 ($p < 0.0000$). The difference in distributions between waves 2 and 3 is not statistically significant ($p = 0.688$). The data show that helping cultures become more heterogeneous in the population of firms over time. For antisocial behavior, we observe a somewhat similar pattern, where the distribution significantly changes from waves 3 to 4 ($p < 0.000$), but stays similar from waves 1 to 2 ($p = 0.212$) and 2 to 3 ($p = 0.324$).

Fig. 4.1: Distribution of establishment-wave averages of the helping index, help offered, help received, and antisocial behavior

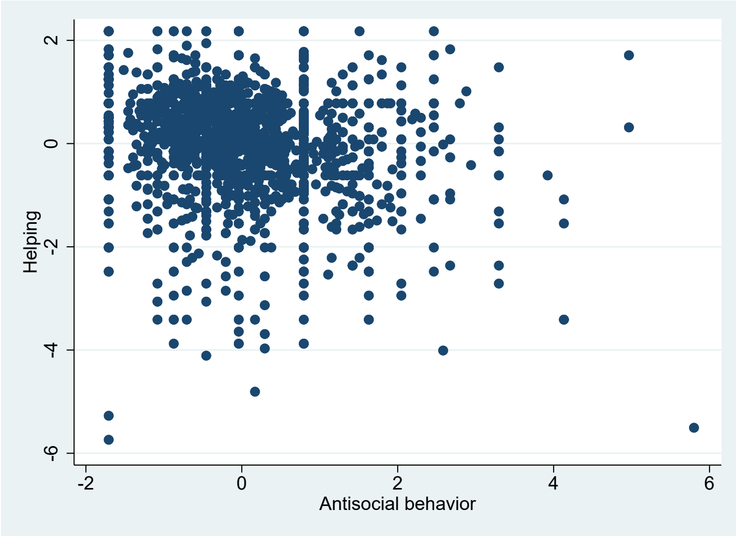


Source: Linked Personnel Panel, waves 1 to 4 (regression sample). Survey items are shown in Table 4.3. $N = 2,002$ per histogram. Only establishments with at least three employee respondents per wave used. One observation is the average response within one establishment-wave cell.

Similar to our findings on the individual level, the firm-level data reveal that helping and antisocial behavior are rather distinct concepts, not just opposites of each other. While helping and antisocial behavior are significantly negatively correlated at the firm level, the correlation size is also rather modest (Spearman $\rho = -0.22$ across all waves, $p < 0.000$). Figure 4.2 shows a scatter plot of both outcome measures. As can be seen, there is considerable variation in antisocial behavior for all levels of helping. Further analysis in the Appendix confirms this impression. Table 4.6 reports a contingency table of median splits for helping and antisocial behavior at the firm-wave level to identify different clusters of firms. The analysis reveals that firms with above-median helping are more likely to have below-median antisocial behavior and vice versa (Pearson Chi-squared and Fisher's exact tests significant at $p < 0.000$). Nevertheless, in 35% of all firm-year observations in which helping is above the median, antisocial behavior is also above the median (corresponding to 17% of all firm-wave observations).[12]

[12] We provide further descriptive information in Tables 4.7 to 4.8 in the Appendix. Distributional plots of our main variables, pooled together, and by survey wave, are shown in Figures 4.3 to 4.12 in the Appendix.

Fig. 4.2: Scatter plot between helping index (std.) and antisocial behavior (std.) at the firm-wave level



Source: Linked Personnel Panel, waves 1 to 4. N=2,002. Only establishments with at least 3 employee respondents used. One observation is the average response within one establishment-wave cell. Survey items are shown in Table 4.3.

4.3.2 *Empirical strategy*

What explains the observed heterogeneity in helping and antisocial behavior across firms? To answer this question, we pursue the following empirical strategy. We first regress the standardized firm-level outcome variables ($Y_{f,t}$) in a firm f at wave t on each of our three main groups of explanatory variables, using the following regression equation:

$$Y_{f,t} = \alpha + \beta_1 EXPLANATORY_{f,t} + \beta_2 CONTROLS_{f,t} + \theta_t + \varepsilon_{f,t}.$$

The main independent firm-level variables ($EXPLANATORY_{f,t}$) are either leadership quality ($LEADERSHIP_{f,t}$), Big Five personality traits ($PERSONALITY_{f,t}$), or economic preferences and trust ($PREFERENCES_{f,t}$). The outcome variables are either helping ($HELP_{f,t}$) or antisocial behavior ($ASB_{f,t}$). Unless stated otherwise, all continuous independent variables are standardized firm-wave averages. The regressions are weighted by the number of observations per firm-wave cell (i.e., the number of employees contributing to the firm-wave average), with standard errors clustered at the firm level. All regressions include our full set of control variables ($CONTROLS_{f,t}$) and survey wave fixed-effects (θ_t). We next simultaneously include all three groups of explanatory variables and estimate the following regression equation in the full model:

$$\begin{aligned} Y_{f,t} = \alpha + \beta_1 LEADERSHIP_{f,t} + \beta_2 PERSONALITY_{f,t} \\ + \beta_3 PREFERENCES_{f,t} + \beta_4 CONTROLS_{f,t} + \theta_t + \varepsilon_{f,t}. \end{aligned}$$

Therefore, we can study whether leadership, personality traits, economic preferences, and trust have distinct additional explanatory power. This type of analysis also shows whether different constructs, such as personality traits, social preferences, or leadership quality are complements or substitutes (Becker et al., 2012).

We consider two main robustness checks. First, we add the lagged dependent variable (LDV), $Y_{f,t-1}$, to mitigate reverse causation (simultaneity bias). The latter could arise from time-varying unobserved characteristics, such as the sorting of workers into firms based on helping or antisocial behavior (Angrist and Pischke, 2008; Beckmann and Kräkel, 2022). The LDV approach also serves to further reduce potential common method bias. Second, as we elicit leadership quality in every survey wave, we run a firm fixed-effects regression to study the link between leadership quality and our outcome variables, while accounting for time-invariant omitted variables at the firm level. We cannot conduct analogous firm fixed-effects regressions to examine the explanatory power of personality traits, social preferences, and trust. The reason is that we elicit these characteristics only the first time an individual joins the employee survey. Hence, any firm-level differences here would require substantial changes in the workforce composition, which we do not observe. Even though we observe attrition in the survey on the employee side, refreshers are not new employees. Hence, if there is a sorting of workers based on preferences and personality, which is rather time constant across firms, as Haylock and Kampkötter (2019b) find, we are unlikely to have enough variation in preferences, personality, and trust over time to use a firm-fixed effects estimation strategy with these explanatory variables.

4.3.3 Explaining helping across firms

Table 4.1 contains our results for helping in the workplace as the dependent variable. We first study whether leadership quality positively correlates with helping in firms. The results reported in column (1) show that leadership quality is positively associated with subordinates' cooperative behavior. Recall that leadership is always answered from the perspective of an employee about his or her supervisor. The effect size is substantial: a one SD increase in leadership quality is associated with a 0.22 SD increase in helping. The specification in column (4) indicates that, although the coefficient is slightly smaller, leadership quality remains positively and significantly correlated with helping even when simultaneously including personality traits, social preferences, and trust in the regression. The specification reported in column (5) shows that the link between leadership and helping remains substantial in size, and significant when also including lagged helping, to reduce reverse causation, and potentially reduce common method bias. Finally, column (6) reports the results of our firm fixed-effects regression. The results in the table indicate that a within-firm increase in leadership quality by one SD is associated with a 0.20 SD increase in helping behavior. Our findings strongly suggest that leadership quality is an essential driver of helping in firms. These results are consistent with Kosfeld (2020) and Hoffman and Tadelis (2021), who find that good leadership induces cooperation among followers and reduces employee turnover. Our data also support the argument that good leadership might signal to followers that helping is profitable, or an advantageous social norm (Hermalin, 1998; Potters et al., 2007; Sliwka, 2007; Danilov and Sliwka, 2017) and that leaders may sustain helping as an equilibrium within groups by fairly punishing defectors, and rewarding helpers (Kosfeld and Rustagi, 2015).

We next consider the association between employees' personality traits and helping. The results reported in column (2) document higher average helping in firms with higher average agreeableness, extroversion, and openness, although the latter is only marginally significant. The coefficients are substantial in size, but smaller compared to the coefficient on leadership quality: A one SD increase in agreeableness is associated with a 0.12 SD increase in helping, a one SD increase in extroversion is associated with a 0.07 SD increase in helping, and a one SD increase in openness is associated with a 0.05 SD increase in helping. Neuroticism is negatively correlated with helping, where a one SD increase in average neuroticism is associated with a 0.06 SD decrease in helping. These results are very intuitive, as, for instance, laboratory experiments show that agreeableness is associated with cooperative tendencies like helping (Kagel and McGee, 2014; Proto et al., 2019). Extroverted employees are more communicative and socialize more, and openness to experience relates to curiosity and willingness to engage in team processes. Therefore, both extroversion and openness should promote helping. Neurotic individuals are overly concerned with envy, are insecure, and are generally worried about themselves; this likely reduces the willingness to engage in mutual helping with colleagues.

Although the above findings are intuitive, the relationship between personality traits and helping is not quite robust. The results reported in column (4) show that the negative correlation between neuroticism and helping is halved, and no longer statistically significant if we include leadership quality, social preferences, and trust in the regression. The estimated coefficient for agreeableness is smaller, although still positively and significantly correlated.

Further including the lagged dependent variable, the findings reported in column (5) reveal that neuroticism and agreeableness are no longer significantly correlated with helping. We instead observe a weakly significant positive link between conscientiousness and helping, a relationship we do not find in any other specifications.

What could explain that neuroticism and agreeableness are no longer significantly correlated with helping once we include the lagged dependent variable in the regression? In this specification, we lose all data points for the first wave of helping as the outcome. The resulting loss in statistical power might explain why the correlation between agreeableness and helping is no longer significant. But also the magnitudes (although not the signs) of the estimated coefficients change once including the lagged dependent variable. It is also possible that the relationship between personality traits and helping is different in the reduced sample compared to the overall sample. To further study this possibility, we run the specification without lagged dependent variable reported in columns (1) to (4), but with the smaller sample used in our regression with the lagged dependent variable as reported in column (5). In column (4) of Table 4.22, shown in Section 4.7.8, we find weakly significant positive correlations of helping with conscientiousness and agreeableness, similar to those in our specification with the lagged dependent variable, but no significant link with neuroticism. The findings on personality traits in our specification with the lagged dependent variable thus seem to be driven by both a loss in statistical power, and a slightly different composition of the sample of considered firms. We conclude that the association between personality traits and helping in firms is not quite robust, at least not as robust as the relationship between leadership and helping.

We next investigate whether social preferences and trust are associated with helping in firms. The results reported in column (3) show that positive reciprocity, altruism, and trust correlate positively with firm-level helping behavior. The effect sizes are once again substantial and comparable to leadership quality. In particular, a one SD increase in trust is associated with a 0.20 SD increase in helping, a one SD increase in positive reciprocity is associated with a 0.06 SD increase in helping, and a one SD increase in altruism is associated with a 0.08 SD increase in helping. The results reported in columns (4) and (5) show that the results remain qualitatively the same when including leadership quality, personality traits, and the lagged dependent variable in the regressions.

The strong correlations between helping and social preferences, and between helping and trust are intuitive, and support existing findings from laboratory and field experiments. Concerning social preferences, positive reciprocity generally describes the willingness to return a favor (Dohmen et al., 2009), and altruism captures the unconditional willingness to support others (Andreoni, 1990; Fehr and Fischbacher, 2003). Both types of social preferences should be positively related to helping and cooperation by definition. Field studies indeed show that prosocial preferences are associated with various kinds of helping behavior (Rustagi et al., 2010; Falk et al., 2018). Further, the theoretical arguments in Kosfeld and von Siemens (2009, 2011) suggest that employees with social preferences self-select into specific firms, creating corporate cultures of cooperation. There exists indeed ample laboratory evidence showing a relation between group formation, social preferences, and voluntary cooperative outcomes (Gächter and Thöni, 2005; Brekke et al., 2011; Gürerk et al., 2014). Our evidence complements these laboratory findings by showing that social preferences are associated with cooperative cultures in the corporate context of representative German firms. Concerning trust, the existing literature suggests that trust should be positively associated with cooperation in general and in organizations (see La Porta et al. (1997) and the references therein). For example, trust is necessary to offer help if and only if one expects help to be returned, and employees might only ask for help if they trust others to do the work properly. Trust differs from social preferences because it additionally captures the beliefs of an individual about being able to rely on others, as well as betrayal aversion (Fehr, 2009). Gächter et al. (2012) and Miettinen et al. (2020) indeed document that beliefs about others' cooperation is a positive predictor for one's own cooperation.

Table 4.1: Determinants of mutual helping

Dep. variable:	Helping index (std.)					
	(1)	(2)	(3)	(4)	(5)	(6)
Leadership	0.2209***			0.1678***	0.1475***	0.2028***
	(0.0233)			(0.0236)	(0.0344)	(0.0534)
Conscientiousness		0.0147		0.0254	0.0669*	
		(0.0259)		(0.0254)	(0.0362)	
Extroversion		0.0661**		0.0515**	0.0270	
		(0.0258)		(0.0255)	(0.0357)	
Neuroticism		-0.0637**		-0.0315	-0.0348	
		(0.0255)		(0.0256)	(0.0346)	
Openness		0.0494*		0.0258	0.0050	
		(0.0277)		(0.0264)	(0.0378)	
Agreeableness		0.1196***		0.0813***	0.0415	
		(0.0263)		(0.0257)	(0.0384)	
Trust			0.1957***	0.1388***	0.1218***	
			(0.0290)	(0.0294)	(0.0408)	
Positive recipr.			0.0568**	0.0538**	0.1054***	
			(0.0227)	(0.0227)	(0.0340)	
Negative recipr.			-0.0050	0.0190	0.0016	
			(0.0251)	(0.0250)	(0.0390)	
Altruism			0.0752***	0.0581**	0.0892**	
			(0.0255)	(0.0244)	(0.0383)	
Base controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual ability	Yes	Yes	Yes	Yes	Yes	Yes
Job controls	Yes	Yes	Yes	Yes	Yes	Yes
HRM practices	Yes	Yes	Yes	Yes	Yes	Yes
Lag helping	No	No	No	No	Yes	No
Firm fixed effects	No	No	No	No	No	Yes
Adj. R^2	0.19	0.18	0.19	0.23	0.28	0.38
Obs.	2,002	2,002	2,002	2,002	973	2,002

The dependent variable “Helping index” is an index containing the standardized firm-wave average of two helping items. All continuous independent variables of interest are standardized with mean zero and unit variance. We analytically weight observations in the OLS regression by cell size (the number of individual observations per firm-wave cell). Our set of base controls includes time and risk preferences, self-efficacy, school education, vocational and university education, gender, age categories, partner, living alone, interview method, survey wave, industry, region, establishment size, and ownership type. Job controls include the logarithm of the monthly net wage, white collar, task interdependence, management position, part-time, and fixed-term work contract. Individual ability is proxied by the firm-wave average of individual AKM fixed effects from the LIAB data, as calculated from 2010 to 2017. Standard errors clustered on the establishment level in parentheses. The symbols *,**, and *** represent significance levels of 10%, 5%, and 1%, respectively.

4.3.4 Explaining antisocial behavior across firms

Table 4.2 contains our results for antisocial behavior in the workplace as the dependent variable. We first study whether leadership quality is related to antisocial behavior in firms. Looking at the results reported in column (1), we observe a large estimated coefficient for leadership quality. A one SD larger average leadership quality is associated with a 0.46 SD lower antisocial behavior in a given firm. As with helping, the relationship between leadership and antisocial behavior appears to be very robust. The results reported in columns (4) and (5) show that the coefficient on leadership always remains highly significant and very similar in size when including personality traits, social preferences, trust, and the lagged dependent variable in the regressions. Our last specification with firm fixed-effects, reported in column (6), confirms that leadership quality remains negatively associated with antisocial behavior even in a within-firm analysis.

Our results on leadership are consistent with the idea that good leaders reduce antisocial behavior by punishing unwanted, and rewarding good behavior (Kosfeld, 2020). Such leadership behavior helps to create a good team climate, which in turn reduces antisocial behavior among employees (Alan et al., 2021). Note that the coefficient size for leadership quality is larger than when explaining mutual help. One possible reason for this result could be that the questionnaire item measuring antisocial behavior specifically includes superiors as perpetrators, while those measuring helping only focus on colleagues.

We next investigate the link between employees' personality traits and antisocial behavior. The results reported in column (2) show that extroversion and agreeableness are negatively correlated with antisocial behavior, while neuroticism is positively correlated with antisocial behavior. In detail, a one SD higher level of extroversion is associated with a 0.05 SD lower level of antisocial behavior, and a one SD higher level of agreeableness is associated with a 0.10 SD lower level of antisocial behavior. On the other hand, a one SD higher level of neuroticism is associated with a 0.16 SD higher level of antisocial behavior. Our results are again very intuitive. In particular, we would expect less antisocial behavior among more agreeable employees. Neuroticism is a likely candidate to explain the level of antisocial behavior in a given firm, because it could easily be associated with misunderstandings between employees. The results reported in columns (4) and (5) show that the link between antisocial behavior and neuroticism is very robust. The relevant coefficient remains similar in size and always highly significant in all specifications. As with helping, the link between the other personality traits and antisocial behavior appears to be less robust. Concerning agreeableness, we find that the negative relationship with antisocial behavior becomes much smaller and only marginally significant when including leadership, social preferences, and trust in our specification. The coefficient is no longer significant once we include the lagged dependent variable. Instead, we observe a somewhat puzzling positive relationship between openness and antisocial behavior in the specifications reported in columns (4) and (5).

We finally consider the link between preferences, trust, and antisocial behavior. Social preferences seem differently linked to antisocial behavior than to helping. The results reported in column (3) suggest that only trust is significantly correlated with antisocial behavior. A one SD increase in trust is associated with a 0.25 SD decrease in antisocial behavior. This result seems plausible, because colleagues who do not trust each other are more likely to end up in conflict. For example, if an employee distrusts her colleagues and consequently behaves in a controlling and uncooperative way, this could lead to a vicious cycle of antisocial behavior. The results reported in columns (4) and (5) show that the negative relationship between antisocial behavior and trust remains significant if we include leadership quality, personality traits, social preferences, and the lagged dependent variable in the regressions.

Table 4.2: Determinants of antisocial behavior

Dep. variable:	Antisocial behavior (std.)					
	(1)	(2)	(3)	(4)	(5)	(6)
Leadership	-0.4567***			-0.4144***	-0.4073***	-0.4474***
	(0.0213)			(0.0209)	(0.0284)	(0.0510)
Conscientiousness		0.0352		0.0269	0.0560	
		(0.0260)		(0.0231)	(0.0350)	
Extroversion		-0.0465*		-0.0217	-0.0065	
		(0.0247)		(0.0213)	(0.0297)	
Neuroticism		0.1637***		0.1010***	0.1441***	
		(0.0262)		(0.0241)	(0.0329)	
Openness		0.0147		0.0486**	0.0688**	
		(0.0256)		(0.0223)	(0.0319)	
Agreeableness		-0.1047***		-0.0426*	-0.0511	
		(0.0273)		(0.0244)	(0.0335)	
Trust			-0.2513***	-0.1281***	-0.1122***	
			(0.0255)	(0.0224)	(0.0321)	
Positive recipr.			-0.0017	-0.0177	-0.0359	
			(0.0238)	(0.0209)	(0.0319)	
Negative recipr.			0.0224	0.0132	0.0393	
			(0.0251)	(0.0220)	(0.0322)	
Altruism			-0.0296	-0.0134	0.0151	
			(0.0260)	(0.0219)	(0.0338)	
Base controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual ability	Yes	Yes	Yes	Yes	Yes	Yes
Job controls	Yes	Yes	Yes	Yes	Yes	Yes
HRM practices	Yes	Yes	Yes	Yes	Yes	Yes
Lag antisocial behavior	No	No	No	No	Yes	No
Firm fixed effects	No	No	No	No	No	Yes
Adj. R^2	0.29	0.13	0.15	0.31	0.36	0.42
Obs.	2,002	2,002	2,002	2,002	973	2,002

The dependent variable is the standardized firm-wave average of antisocial behavior. All continuous independent variables of interest are standardized with mean zero and unit variance. We analytically weight observations in the OLS regression by cell size (the number of individual observations per firm-wave cell). Our set of base controls includes time and risk preferences, self-efficacy, school education, vocational and university education, gender, age categories, partner, living alone, interview method, survey wave, industry, region, establishment size, and ownership type. Job controls include the logarithm of the monthly net wage, white collar, task interdependence, management position, part-time, and fixed-term work contract. Individual ability is proxied by the firm-wave average of individual AKM fixed effects from the LIAB data, as calculated from 2010 to 2017. Standard errors clustered on the establishment level in parentheses. The symbols *,**, and *** represent significance levels of 10%, 5%, and 1%, respectively.

4.3.5 *Robustness tests*

We conduct several robustness checks, and present the results in Section 4.7. In detail, we average observations across all survey waves in our pooled analysis, control for lagged organizational climate measures, analyze items from the helping index separately, and also analyze items from the leadership index separately. We also exclude missing indicators to ensure this does not drive results. In the following, we briefly present the robustness checks separately for our two dependent variables, helping and antisocial behavior.

4.3.5.1 Helping

First, we calculate firm-level averages of all variables over the whole time period, as pooling all observations per firm might remove idiosyncratic measurement error and lead to more precise estimates. Hence, we use one observation per establishment here. This might be particularly important in the case of leadership quality, since it could be the case that the large coefficient sizes for leadership quality are mainly driven by larger variation due to its repeated measurement. However, Table 4.12 shows that our main results are broadly confirmed.

It could also be that our correlations are spuriously driven by other elements of the firm's culture, for example, it could be that more committed employees on average help more, but are also more trusting persons. We can rule out that this is driven by various other proxies of organizational climate. In Table 4.14 the main results are robust to controlling for a series of lagged firm-level organizational climate variables elicited at the individual level, and aggregated to the firm level, including employee engagement, job satisfaction, affective commitment, turnover intention, and sick days. We also control for the AKM firm fixed effect. These survey items used for each variable are in Table 4.3.

In Tables 4.10 and 4.11, we separate the helping index into its two original survey items used to measure helping behavior, *help given*, and *help received*. Surprisingly, the results are nearly identical for both items, which again justifies their aggregation and shows that no single item drives our main results.

Next, we exclude employee sample refreshers from the employee survey from the analysis to check whether our results are driven by the refresher employees that replace survey attritors. This tests whether results are potentially driven by unobserved differences in characteristics of refreshers and original survey respondents. As Table 4.16 shows, the results again are very similar to our baseline specification, although trust does lose significance in column 5, the lagged dependent variable specification.

We also use a balanced panel of firms to address a potential selectivity bias at the firm level. Here, we only keep firms in the data set that have been continuously surveyed in all four waves of the LPP. The balanced panel has the benefit of checking whether any developments that we observe over time are driven by a different composition of firms in the sample. The refreshment sample of firms used to construct the full sample aims to re-balance the initial sample, with respect to observable firm characteristics. It could be that for variables not considered in re-balancing, for example the establishment age, the developments we observe are just driven by changes in the composition of firms with respect to these variables. Doing this exercise, we get an idea about the magnitude of the bias when looking at the balanced sample. Of course, this also introduces a survivor bias (surviving firms are not representative). As Table 4.18 shows, positive reciprocity loses significance and reduces in size, but other coefficients remain similar in size in most specifications. Additionally, agreeableness stays significant across all specifications and large in size. Altruism loses significance in column (5), but stays relatively similar in size.

Excluding observations with missing values in Table 4.20 shows very robust results regarding all coefficients, and that results are not driven by including indicators of missing covariates. Last, in column (1) of Table 4.9, we regress the helping index on each item from the leadership index separately. This shows that primarily supervisor fairness and supervisor understanding are the main drivers of helping within the leadership index. Further, supervisor fairness and supervisor understanding are correlated with both help given and help received, as columns (3) and (4) show. To conclude, our findings regarding the link between personality traits, trust, social preferences, leadership, and helping behavior appear to be moderately robust across a broad range of alternative specifications.

4.3.5.2 Antisocial behavior

Similarly to the previous subsection, we run the same robustness checks for antisocial behavior and again, our main results remain unchanged. The pooled analysis in Table 4.13 shows very similar coefficient sizes and significance of neuroticism, agreeableness, trust, and leadership to the main specification. Some coefficients also become larger and increase in significance.

Similar to above, in Table 4.15 the main results are robust to controlling for the series of lagged firm-level outcome variables (engagement, job satisfaction, affective commitment, turnover intention, sick days, and the AKM firm fixed effect). In Table 4.17, where we exclude employee sample refreshers from the analysis, coefficients of neuroticism, trust and leadership are all significant and similar in magnitude. When using a balanced panel of firms in Table 4.19, coefficient sizes and significance levels for neuroticism, trust, and leadership remain largely unchanged. Excluding observations with missing values in Table 4.21 shows very robust results regarding all coefficients.

Assessing which elements of the leadership index drive the results, we regress antisocial behavior on the individual leadership items in Table 4.9 of the Appendix. This shows that each item drives antisocial behavior to some extent, but that supervisor fairness is largest in magnitude. This is intuitive since antisocial behavior also mentions treatment by the supervisor in the survey item. Nevertheless, supervisor trust and understanding also play a substantial role.

Altogether, these robustness checks support the interpretation that leadership, along with personality play and important role in explaining antisocial behavior in the workplace.

4.4 Conclusion

This paper uses unique linked employer-employee data to document the heterogeneity, and the determinants of firm-level measures of helping and antisocial behavior, two important proxies of collective action in the workplace. Our data are representative of a large, developed economy and cover a substantial period. The survey applies experimentally validated items to measure social preferences, and they use other established measures from validated scales to measure trust, personality traits, and leadership quality. Finally, the data provide rich information on further employee and firm characteristics, including human resource management practices. As far as we know, we are the first to combine such rich and complementary employer-employee data to uncover the economic and behavioral foundations of helping and antisocial behavior in the workplace.

Our results document considerable variation in helping and antisocial behavior across firms. Although these behaviors are negatively correlated, many firms exhibit both high levels of helping and antisocial behavior, suggesting that these are distinct constructs. Moreover, altruism, positive reciprocity, and trust are most important for explaining prosocial behavior in a workplace context across a representative sample of firms. Negative reciprocity does not play a substantial role here. Employees' personality traits also matter, especially neuroticism to explain antisocial behavior. In addition, leadership quality adds significant explanatory power and is strongly associated with more helping and in particular less antisocial behavior.

Summarizing, we provide the first representative evidence that personality traits, preferences, trust, and leadership quality are essential explanations for the observed differences in helping and antisocial behavior across firms. Although our findings are correlative, they corroborate lab evidence that helping and antisocial behavior matter, and provide additional external validity to a growing body of evidence on the importance of preferences, personality, and leadership in a broad range of settings. Our findings further indicate that selecting the right employees is vital for promoting helping and curbing antisocial behavior in organizations. However, good leadership also appears to have a substantial additional explanatory power in explaining workplace behavior. Employee selection is not everything; even if a firm can get the right workers together, a leader can substantially affect these behaviors on top of personality and preferences.

4.5 Descriptive statistics

Table 4.3: Survey items of individual-level variables

Survey item (or index)	Repeatedly measured	Exact wording of item(s)	Scale
Helping index	Yes	A: “How often do you receive support or help from your colleagues if you ask?" B: “How often do you offer your colleagues help?”	Index (5-point)
Antisocial behavior	Yes	“How often do you feel wrongly criticized, harassed or denounced by your colleagues or superiors?”	5-point
Leadership index	Yes	A: “Supervisors show that they trust in those they manage.” B: “Supervisors show an understanding of the people who work for them.” C: “The way my supervisor treats me is fair.”	5-point
Positive reciprocity	No	“If someone does me a favor, I am prepared to return it."	5-point
Negative reciprocity	No	“If I am treated very unjustly, I will take revenge at the first occasion, even if there is a cost to do so."	5-point
Altruism	No	“How do you assess your willingness to share with others without expecting anything in return?"	11-point
Trust	No	A: “Nowadays, one can’t rely on anybody.” (R) B: “On the whole, one can trust people.”	5-point
Big 5: Openness to experience	No	“I see myself as someone who..." A: “is original, comes up with new ideas.” B: “values artistic, aesthetic experiences.” “C: has an active imagination.” D: “is eager for knowledge.”	5-point
Big 5: Extroversion	No	“I see myself as someone who...” A: “is communicative, talkative.” B: “is reserved.” (R) C: “is outgoing, sociable:”	5-point
Big 5: Conscientiousness	No	“I see myself as someone who...” A: “does a thorough job.” B: “tends to be lazy.” (R) C: “does things effectively and efficiently.”	5-point
Big 5: Agreeableness	No	“I see myself as someone who...” A: “is sometimes somewhat rude to others.” (R) B: “has a forgiving nature.” C: “is considerate and kind to others.”	5-point
Big 5: Neuroticism	No	“I see myself as someone who...” A: “worries a lot.” B: “gets nervous easily” C: “is relaxed, handles stress well” (R)	5-point
Job satisfaction	Yes	“How satisfied are you with your job?”	11-point
Work engagement	Yes	A: “At my work, I feel bursting with energy.” B: “At my job, I feel strong and vigorous.” C: “When I get up in the morning, I feel like going to work.” D: “I am enthusiastic about my job.” E: “My job inspires me.” F: “I am proud on the work that I do.” G: “I feel happy when I am working intensely.” H: I am immersed in my work.” I: “I get carried away when I’m working.”	5-point
Affective commitment	Yes	A: “I would be very happy to spend the rest of my career with this organization.” B: “This organization has a great deal of personal meaning for me.” C: “I really feel as if this organization’s problems are my own.” D: “I do not feel a strong sense of “belonging” to my organization.” (R) E: “I do not feel “emotionally attached” to this organization. ” (R) F: “I do not feel like “part of the family” at my organization.” (R)	5-point
Turnover intention	Yes	“How many times in the past 12 months have you thought about changing your job?”	5-point
Sick days	Yes	“How many days did you not work during the previous 12 months because you were ill? Please indicate all sick days, not only those for which you have received a doctor’s certificate of incapacity to work.”	Range from 1-230 days
Risk attitude	No	“How do you see yourself: Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?”	11-point
Time preference	No	“A: I abstain from certain things today so I can afford more tomorrow.” “B: I tend to procrastinate things even though it would be better to do them now.” (R)	5-point

(R): Reverse-coded

Table 4.4: Survey items of establishment-level variables

Survey item	Repeatedly measured	Exact wording	Scale
Feedback talk	Yes	“Are structured employee feedback talks conducted at least once a year in your establishment?"	Yes/No
Performance appraisal	Yes	“Does your establishment evaluate the performance of employees at least once a year by their respective supervisors?"	Yes/No
Written target agreement	Yes	Does your establishment use written target agreements?	Yes/No
Consequence management	Yes	"How do you and your managers deal with employees whose performance you are not satisfied with? A: Supervisors openly discuss the problems with the employee concerned. B: HR development measures are offered in a targeted manner to address performance problems. C: If performance problems persist, the company looks for another position in the firm. D: Employees who consistently perform poorly are dismissed or encouraged to leave the company."	Index (5-point Likert scale)
Development plan	Yes	“Does your establishment use (career) development plans for employees?"	Yes/No
Employee survey	Yes	“Are employee surveys conducted regularly in your establishment?"	Yes/No
Working atmosphere	Yes	“In your opinion, how important is the following aspect for retaining your employees in your company: General working atmosphere"	5-point Likert scale
Code of conduct	Yes	“Does your establishment have a written strategy for promoting diversity and equal opportunities in the workforce with regard to characteristics such as gender, age, nationality, culture, religion or sexual orientation?"	Yes/No
Share non-manag. PP	Yes	“What percentage of non-managerial employees receive a variable compensation?"	Percent
Bonus to base	Yes	“How large is the average variable pay component measured as a percentage of the fixed base salary when targets have been met? Please distinguish between managerial and non-managerial employees."	Percent
Pay mix	Yes	“What percentage of the variable compensation of the two employee groups (managerial and non-managerial) is, on average, based on the criteria company-wide performance, team/divisional performance, and individual performance?"	Percent (adding up to 100%)

Table 4.5: Summary statistics (weighted by number of employees)

	Obs.	Mean	S.D.	Min.	Max.	Med.
Number of employees (unweighted)	2,002	9.2	17.2	3	536	6
Helping index	2,002	4.2	0.3	2.2	5	4.3
Help offered	2,002	4.2	0.3	1	5	4.2
Help received	2,002	4.3	0.3	1.3	5	4.3
Antisocial behavior	2,002	1.7	0.3	1	4	1.6
Sick days	2,002	12.3	9.4	0	95	9.6
AKM ind. $FE_{2010-2017}$	1,996	4.2	0.2	3.3	5.2	4.2
Trust	2,002	3.5	0.3	1.9	4.6	3.5
Positive reciprocity	1,793	4.5	0.2	2	5	4.5
Negative reciprocity	1,793	1.9	0.4	1	5	1.9
Altruism	1,793	7.7	0.6	4	10	7.7
Risk tolerance	2,002	5.6	0.6	1.8	8.5	5.6
Time preference	1,792	3	0.3	1	5	3
Conscientiousness	2,002	4.4	0.2	3.2	5	4.3
Extroversion	2,002	3.7	0.3	2.3	4.9	3.7
Neuroticism	2,002	2.7	0.3	1.3	4	2.7
Openness	2,002	3.6	0.2	2.3	4.9	3.7
Agreeableness	2,002	4	0.2	2.9	5	4
Self efficacy	1,793	4.2	0.2	3	5	4.2
Leadership	2,002	3.8	0.3	1.9	5	3.8
Supervisor trust	2,002	3.8	0.4	1.3	5	3.8
Supervisor understanding	2,002	3.7	0.4	1.2	5	3.8
Supervisor fairness	2,002	3.9	0.4	1.3	5	4
Feedback talk	2,002	0.8	0.4	0	1	1
Development plan	2,001	0.6	0.5	0	1	1
Performance appraisal	2,001	0.7	0.4	0	1	1
Employee survey	2,000	0.5	0.5	0	1	1
Code of conduct	2,001	0.5	0.5	0	1	0
Low performer	1,992	3.5	0.7	1	5	3.5
Target agreement	1,644	0.9	0.3	0	1	1
Work climate important	1,998	4.3	0.7	1	5	4
Ind. PP empl.	2,002	0.4	0.5	0	1	0
Team PP empl.	2,002	0.2	0.4	0	1	0
Firm PP empl.	2,002	0.3	0.5	0	1	0
Ind. PP man.	2,002	0.5	0.5	0	1	0
Team PP man.	2,002	0.3	0.5	0	1	0
Firm PP man.	2,002	0.5	0.5	0	1	1
Share staff PP	2,002	33	42.5	0	100	4
Depend on me	2,002	3.8	0.5	1	5	3.8
Depend on others	2,002	3.4	0.5	1	5	3.4
White-collar	2,002	0.6	0.3	0	1	0.7
Management position	2,002	0.3	0.2	0	1	0.3
Part-time	2,002	0.1	0.2	0	1	0.1
Log monthly net wage	2,001	7.7	0.4	4	9.9	7.7
Fixed-term work contract	2,002	0	0.1	0	1	0
Female	2,002	0.3	0.2	0	1	0.2
Under 30 years	2,002	0.1	0.1	0	0.8	0.1
30-40 years old	2,002	0.2	0.1	0	1	0.2
40-50 years old	2,002	0.3	0.2	0	1	0.3
Partner	2,002	0.8	0.1	0	1	0.9
Lives alone	2,002	0.1	0.1	0	1	0.1
No school certificate	2,002	0	0	0	0.3	0
9th grade	2,002	0.2	0.2	0	1	0.2
10th grade	2,002	0.4	0.2	0	1	0.4
Univ. of app. sc. entrance qual.	2,002	0.1	0.1	0	1	0.1
University entrance qualification	2,002	0.2	0.2	0	1	0.2
Other school education	2,002	0	0	0	1	0
No further education	2,002	0	0.1	0	1	0
Apprenticeship	2,002	0.5	0.2	0	1	0.5

Continued on next page

Table 4.5 – *Continued from previous page*

	Obs.	Mean	S.D.	Min.	Max.	Med.
Trade school	2,002	0.1	0.1	0	1	0.1
Master craftsman	2,002	0.2	0.2	0	1	0.2
Univ. of appl. sciences	2,002	0.1	0.1	0	1	0.1
University	2,002	0.1	0.1	0	1	0.1
Other further education	2,002	0	0	0	0.3	0
North	2,002	0.2	0.4	0	1	0
East	2,002	0.2	0.4	0	1	0
South	2,002	0.3	0.4	0	1	0
West	2,002	0.3	0.5	0	1	0
Under 50	2,002	0	0.1	0	1	0
50-99	2,002	0.1	0.3	0	1	0
100-249	2,002	0.2	0.4	0	1	0
250-499	2,002	0.2	0.4	0	1	0
Above 500	2,002	0.4	0.5	0	1	0
Manufact.	2,002	0.3	0.5	0	1	0
Metal	2,002	0.4	0.5	0	1	0
Commerce	2,002	0.1	0.3	0	1	0
Bus./fin. serv.	2,002	0.1	0.3	0	1	0
IT/oth. serv.	2,002	0	0.2	0	1	0
Health/Social	2,002	0	0.2	0	1	0
Family owned	2,002	0.4	0.5	0	1	0
Management-owned	2,002	0.1	0.4	0	1	0
Financial inv./Dispersed ownership	2,002	0.2	0.4	0	1	0
State-owned	2,002	0	0.2	0	1	0
Other ownership type	2,002	0.2	0.4	0	1	0
CATI	2,002	0.9	0.3	0	1	1
Wave 1	2,002	0.3	0.5	0	1	0
Wave 2	2,002	0.3	0.4	0	1	0
Wave 3	2,002	0.2	0.4	0	1	0
Wave 4	2,002	0.2	0.4	0	1	0

Table 4.6: Contingency table of helping index and ASB

	ASB≤median	ASB>median	Total
Helping≤median	472 (47%)	531 (53%)	1,003 (100%)
Helping>median	653 (65%)	346 (35%)	999 (100%)
Total observations	1,125	877	2,002

The table shows a two-way contingency table of above and below median helping index and antisocial behavior (ASB) at the firm-wave level. Survey items are shown in Table 4.3. The top number in each cell is the frequency and the bottom number is the row percentage. Source: Linked Personnel Panel, waves 1 to 4. One observation of the helping (ASB) index is the firm-wave average.

Table 4.7: Summary statistics of outcome variables by industry sector

Helping				
Sector	N	Mean	Med.	S.D.
Manufact.	671	4.25	4.28	0.28
Metal	661	4.24	4.25	0.25
Commerce	267	4.19	4.25	0.37
Bus./fin. serv.	265	4.23	4.29	0.32
IT/oth. serv.	77	4.25	4.29	0.31
Health/Social	81	4.22	4.25	0.28
Total	2,022	4.24	4.25	0.28
Antisocial behavior				
Sector	N	Mean	Med.	S.D.
Manufact.	671	1.68	1.63	0.33
Metal	661	1.67	1.67	0.28
Commerce	267	1.69	1.67	0.36
Bus./fin. serv.	265	1.59	1.52	0.32
IT/oth. serv.	77	1.60	1.55	0.32
Health/Social	81	1.72	1.70	0.33
Total	2,022	1.66	1.64	0.31
CV helping				
Sector	N	Mean	Med.	S.D.
Manufact.	670	0.15	0.14	0.06
Metal	661	0.16	0.16	0.05
Commerce	267	0.18	0.16	0.08
Bus./fin. serv.	265	0.17	0.15	0.08
IT/oth. serv.	76	0.15	0.14	0.07
Health/Social	80	0.15	0.14	0.06
Total	2,019	0.16	0.15	0.06
CV antisocial behavior				
Sector	N	Mean	Med.	S.D.
Manufact.	670	0.46	0.47	0.13
Metal	661	0.48	0.50	0.11
Commerce	267	0.50	0.51	0.15
Bus./fin. serv.	265	0.46	0.47	0.15
IT/oth. serv.	77	0.46	0.43	0.11
Health/Social	80	0.45	0.47	0.13
Total	2,020	0.47	0.48	0.13

Analytically weighted summary statistics of firm-wave averages of helping index and antisocial behavior and firm-wave level coefficients of variation (CV) of helping index and antisocial behavior across sectors. Source: Linked Personnel Panel, waves 1 to 4. Survey items shown in Table 4.3. Weights used are the number of observations in each establishment-wave cell, which give rise to one observation.

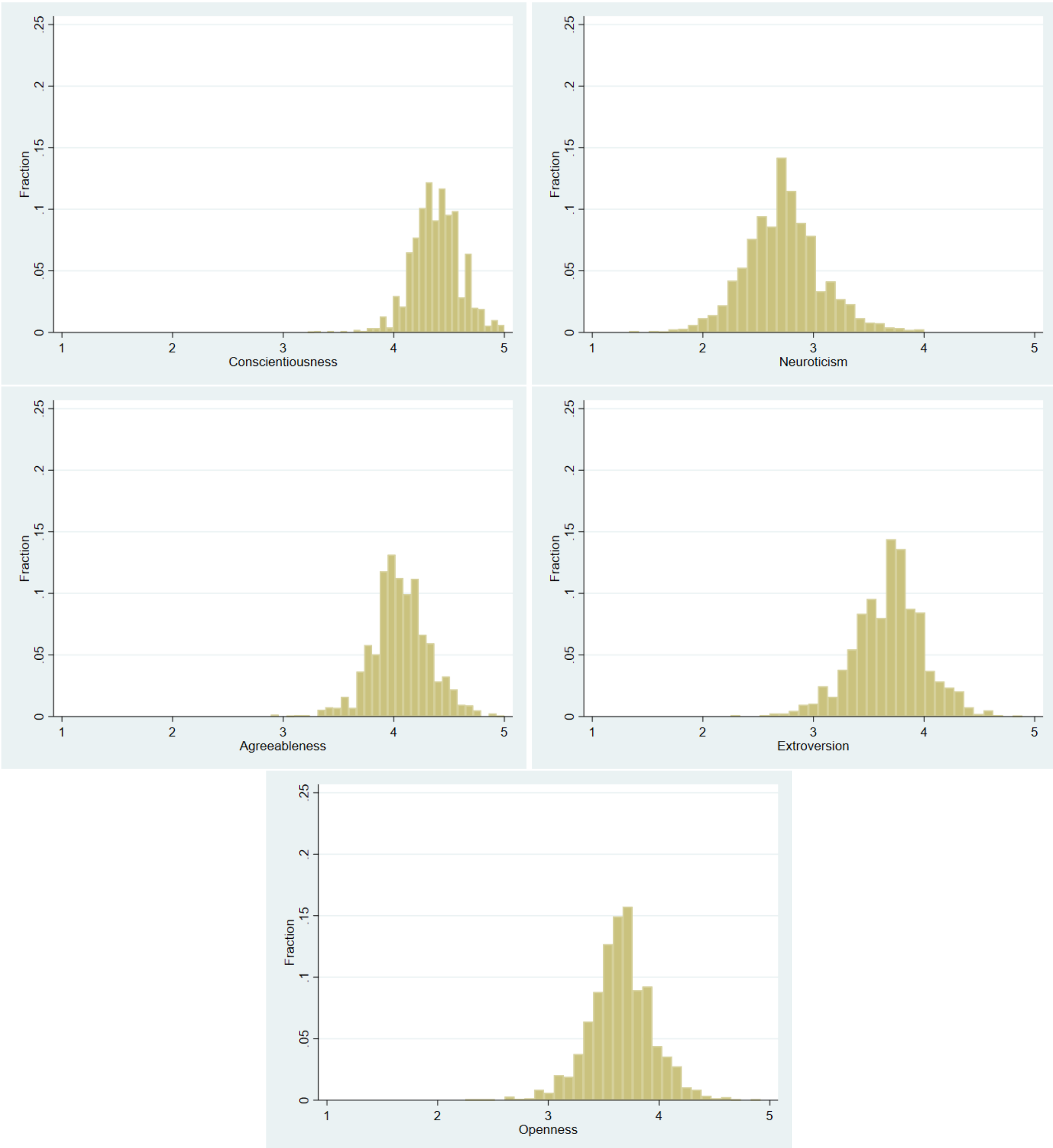
Table 4.8: Weighted average of outcome variables by sector and survey wave

Helping				
Sector	Wave 1	Wave 2	Wave 3	Wave 4
Manufact.	4.30	4.24	4.26	4.13
Metal	4.34	4.26	4.25	4.09
Commerce	4.31	4.13	4.19	3.96
Bus./fin. serv.	4.27	4.22	4.25	4.10
IT/oth. serv.	4.35	4.28	4.22	4.09
Health/Social	4.30	4.20	4.24	4.02
Total	4.31	4.23	4.25	4.10
Antisocial behavior				
Sector	Wave 1	Wave 2	Wave 3	Wave 4
Manufact.	1.69	1.66	1.65	1.73
Metal	1.64	1.66	1.65	1.75
Commerce	1.69	1.72	1.58	1.76
Bus./fin. serv.	1.63	1.57	1.48	1.67
IT/oth. serv.	1.52	1.67	1.57	1.63
Health/Social	1.74	1.61	1.86	1.68
Total	1.66	1.66	1.63	1.73
CV helping				
Sector	Wave 1	Wave 2	Wave 3	Wave 4
Manufact.	0.15	0.16	0.15	0.16
Metal	0.15	0.16	0.16	0.18
Commerce	0.16	0.20	0.17	0.20
Bus./fin. serv.	0.16	0.17	0.16	0.18
IT/oth. serv.	0.16	0.14	0.13	0.18
Health/Social	0.15	0.16	0.14	0.17
Total	0.15	0.16	0.16	0.17
CV antisocial behavior				
Sector	Wave 1	Wave 2	Wave 3	Wave 4
Manufact.	0.45	0.46	0.45	0.44
Metal	0.48	0.48	0.48	0.49
Commerce	0.51	0.51	0.46	0.49
Bus./fin. serv.	0.48	0.47	0.45	0.44
IT/oth. serv.	0.45	0.45	0.45	0.50
Health/Social	0.44	0.47	0.47	0.42
Total	0.47	0.47	0.46	0.48

Analytically weighted averages of firm-wave helping index and antisocial behavior as well as firm-wave level coefficients of variation (CV) of helping index and antisocial behavior, across industry sectors and time. Source: Linked Personnel Panel, waves 1 to 4. Survey items shown in Table 4.3. Weights used are number of observations in each establishment-wave cell, which give rise to one observation. Number of firm-wave cells is $N = 2,002$.

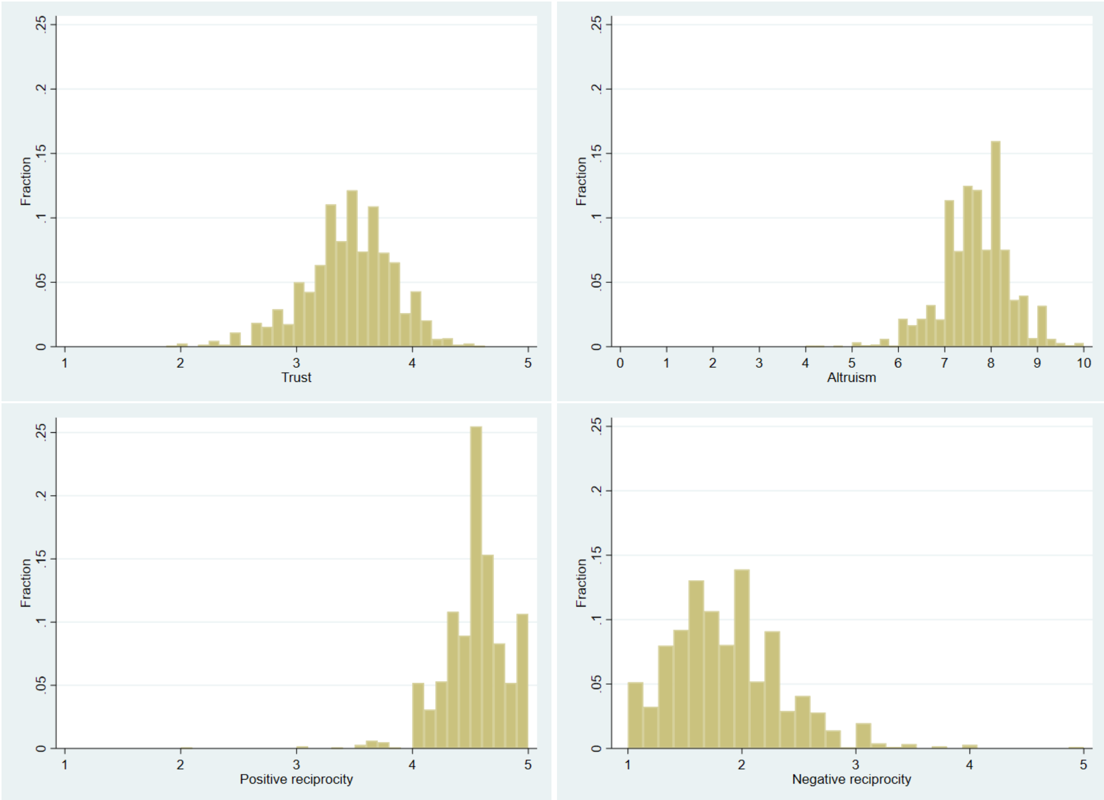
4.6 Distribution of main variables

Fig. 4.3: Distribution of establishment-wave average Big Five personality traits



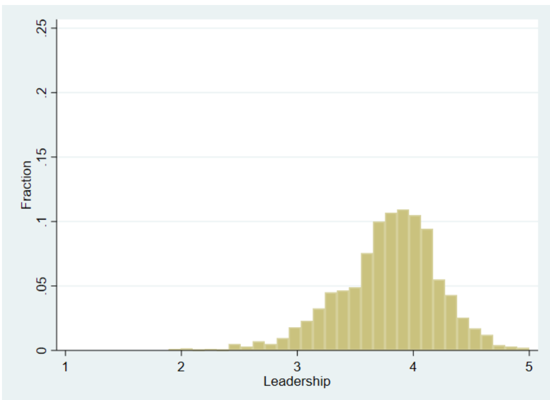
Source: Linked Personnel Panel, waves 1 to 4. N=2,002 in total. Wave 1: N=724, wave 2: N=541, wave 3: N=451, wave 4: N=286. Only establishments with at least 3 employee respondents used. One observation is the average response within one establishment-wave cell. Survey items shown in Table 4.3.

Fig. 4.4: Distribution of establishment-wave average trust and social preferences



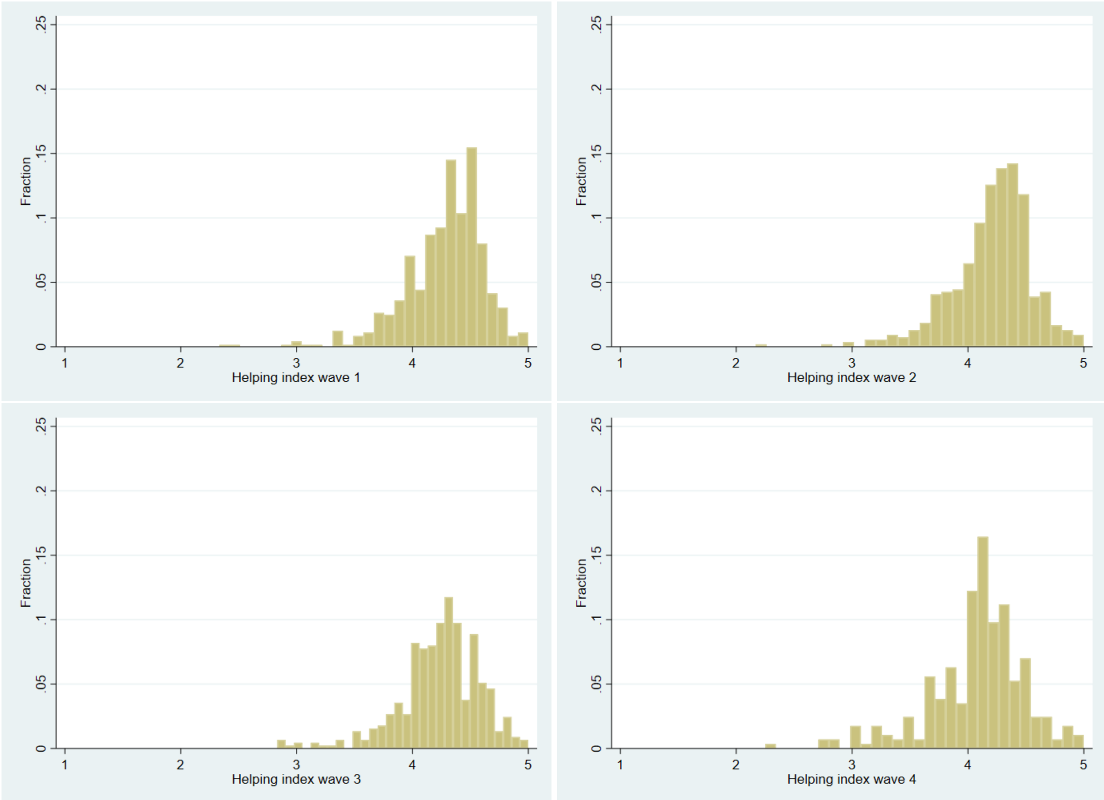
Source: Linked Personnel Panel, waves 1 to 4. N=2,002 in total. Wave 1: N=724, wave 2: N=541, wave 3: N=451, wave 4 N=286, except positive reciprocity, negative reciprocity, and altruism (each wave 1: N=515, wave 2: N=541, wave 3: N=451, wave 4 N=286). Only establishments with at least 3 employee respondents per wave used. One observation is the average response within one establishment-wave cell. Survey items shown in Table 4.3.

Fig. 4.5: Distribution of establishment-wave average leadership index



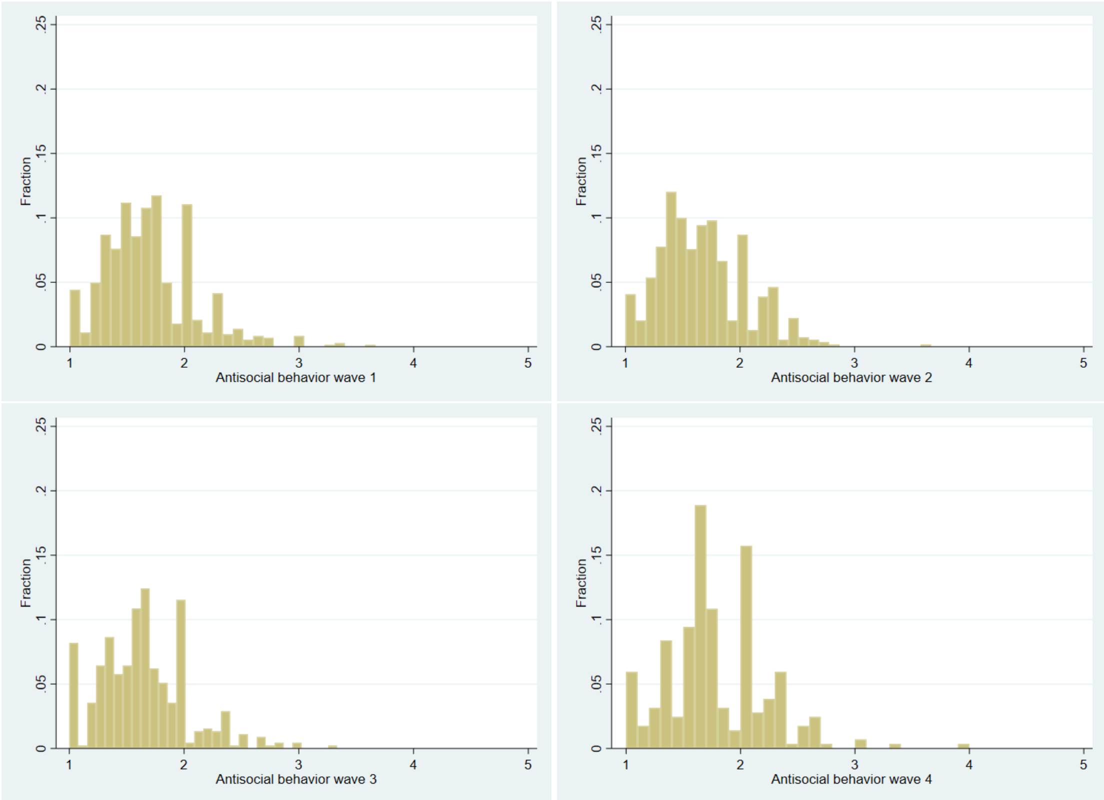
The leadership index is the equally weighted average of three items, supervisor trust, supervisor understanding and supervisor fairness. Source: Linked Personnel Panel, waves 1 to 4. N=2,002 in total. Wave 1: N=724, wave 2: N=541, wave 3: N=451, wave: 4 N=286. Only establishments with at least 3 employee respondents per wave used. One observation is the average response within one establishment-wave cell. Survey items shown in Table 4.3.

Fig. 4.6: Distribution of establishment-wave average helping index by wave



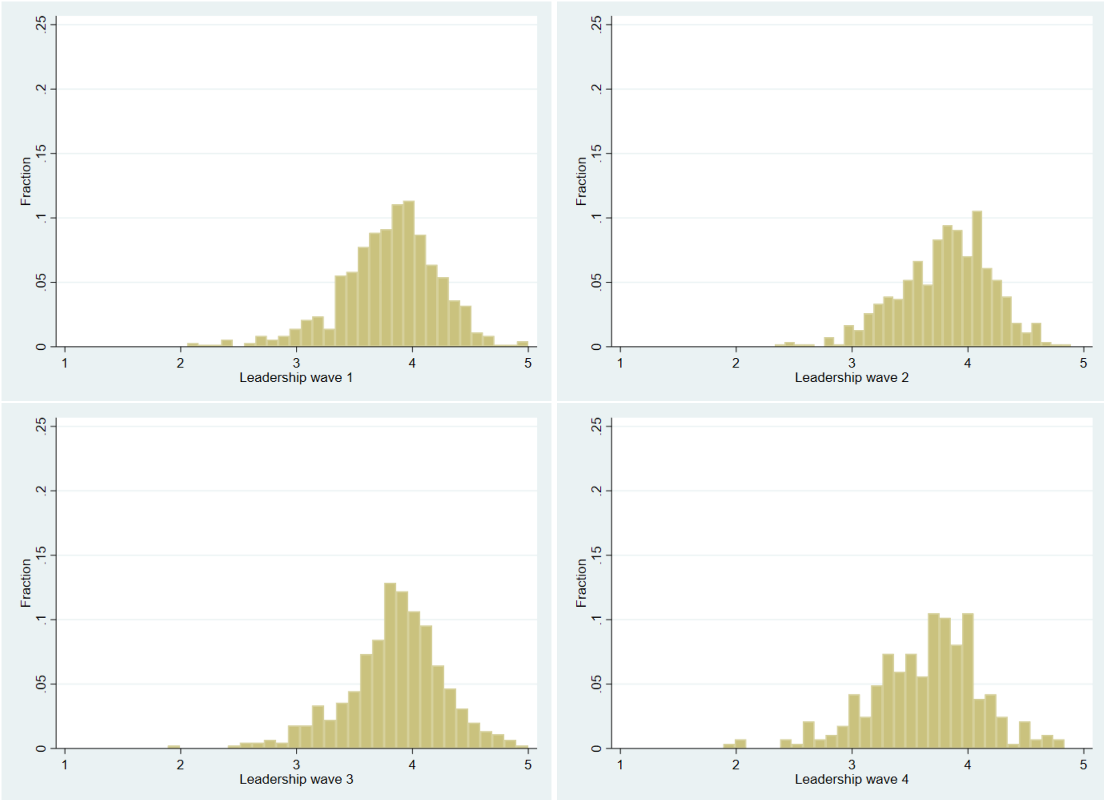
Source: Linked Personnel Panel, waves 1 to 4. N=2,002 in total. Wave 1: N=724, wave 2: N=541, wave: 3 N=451, wave: 4 N=286. Only establishments with at least 3 employee respondents per wave used. One observation is the average response within one establishment-wave cell. Survey items shown in Table 4.3.

Fig. 4.7: Distribution of establishment-wave average antisocial behavior of establishments by wave



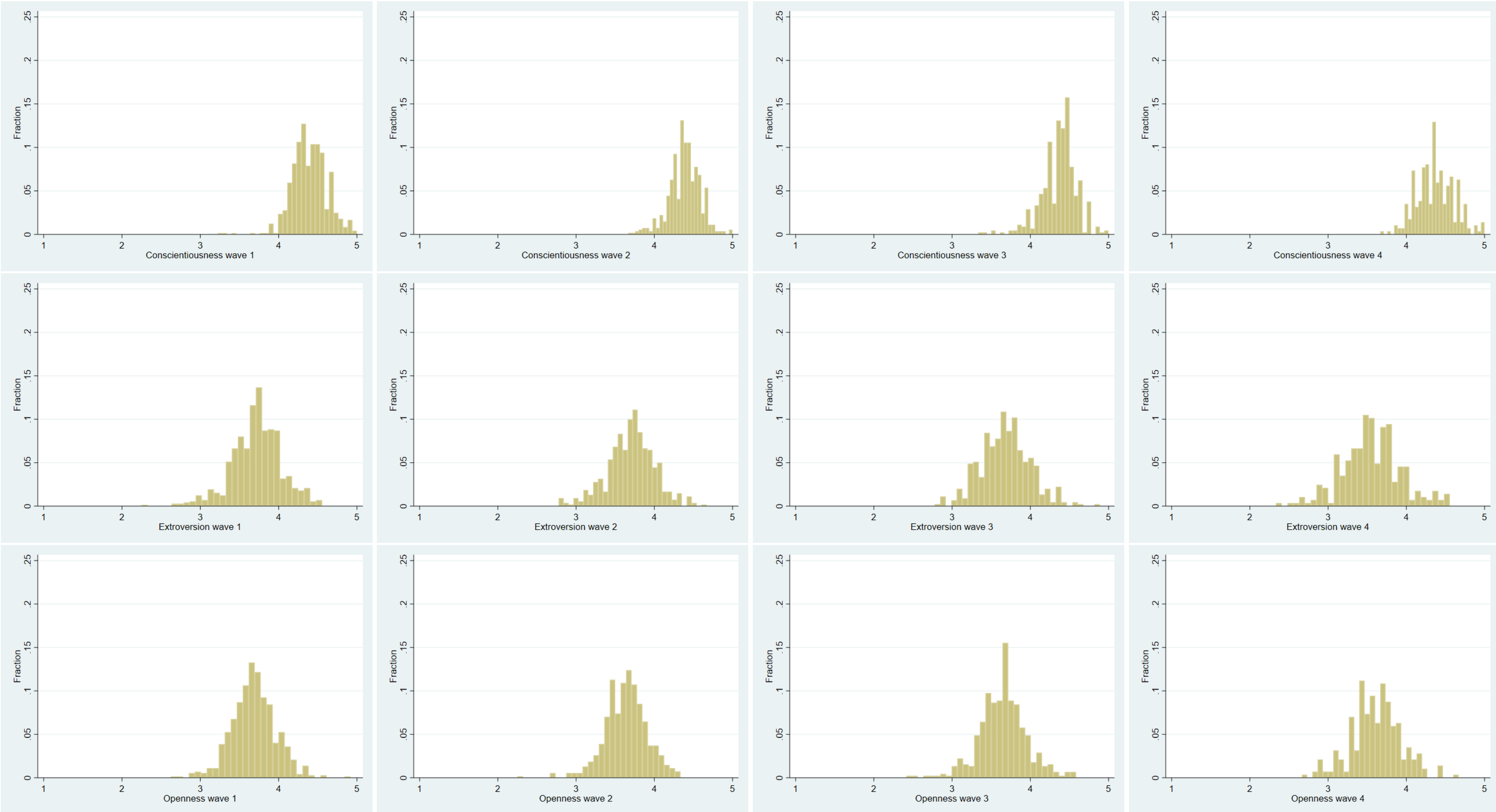
Source: Linked Personnel Panel, waves 1 to 4. N=2,002 in total. In wave1 N=724, wave 2 N=541, in wave 3 N=451, in wave 4 N=286. Only establishments with at least 3 employee respondents per wave used. One observation is the average response within one establishment-wave cell. Survey items shown in Table 4.3.

Fig. 4.8: Distribution of establishment-wave average leadership index by wave



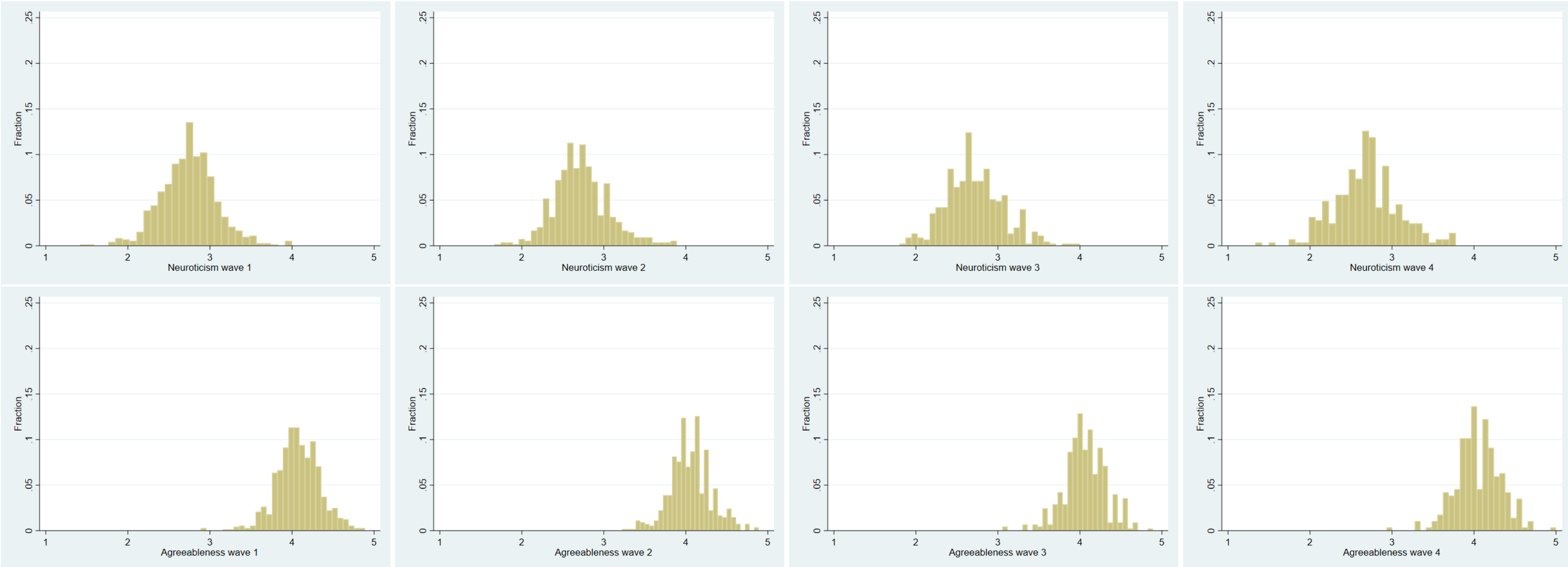
Source: Linked Personnel Panel, waves 1 to 4. N=2,002 in total. Wave 1: N=724, wave 2: N=541, wave: 3 N=451, wave: 4 N=286. Only establishments with at least 3 employee respondents per wave used. One observation is the average response within one establishment-wave cell. Survey items shown in Table 4.3.

Fig. 4.9: Distribution of establishment-wave average conscientiousness, extroversion, and openness to experience by wave



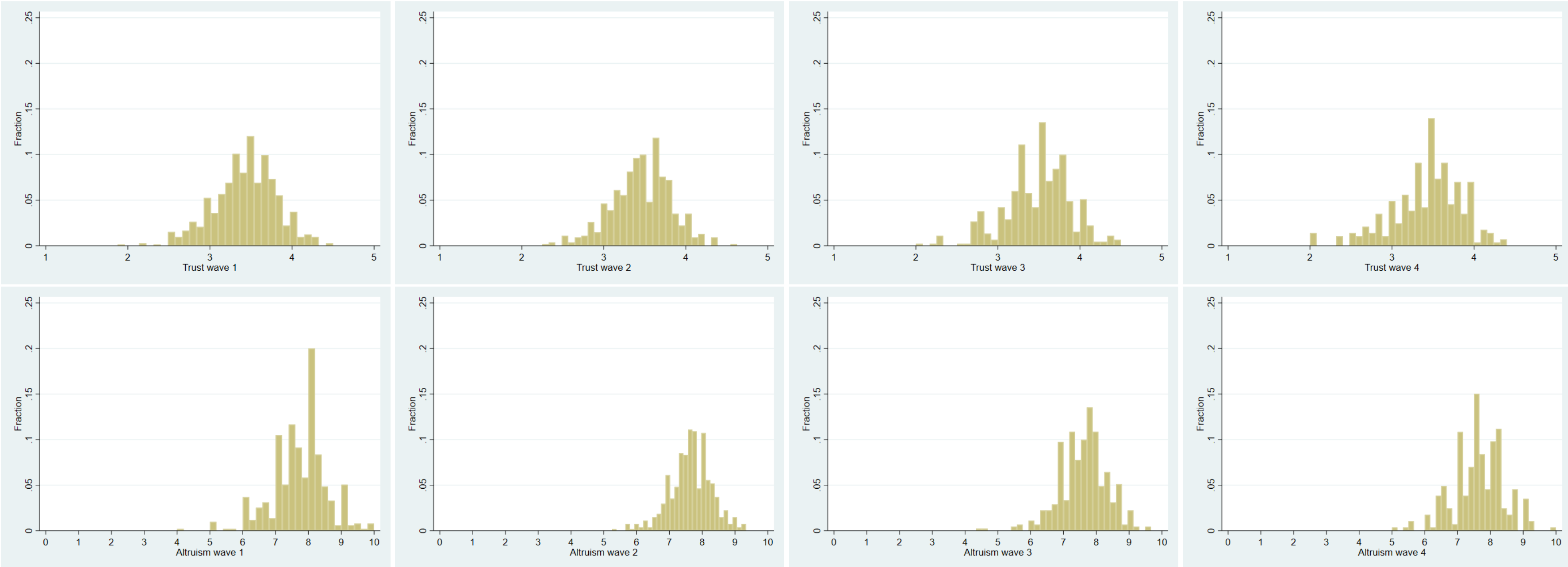
Source: Linked Personnel Panel, waves 1 to 4. N=2,002 in total. Wave 1: N=724, wave 2: N=541, wave 3: N=451, wave 4: N=286. Only establishments with at least 3 employee respondents used. One observation is the average response within one establishment-wave cell. Survey items shown in Table 4.3.

Fig. 4.10: Distribution of establishment-wave average neuroticism and agreeableness by wave



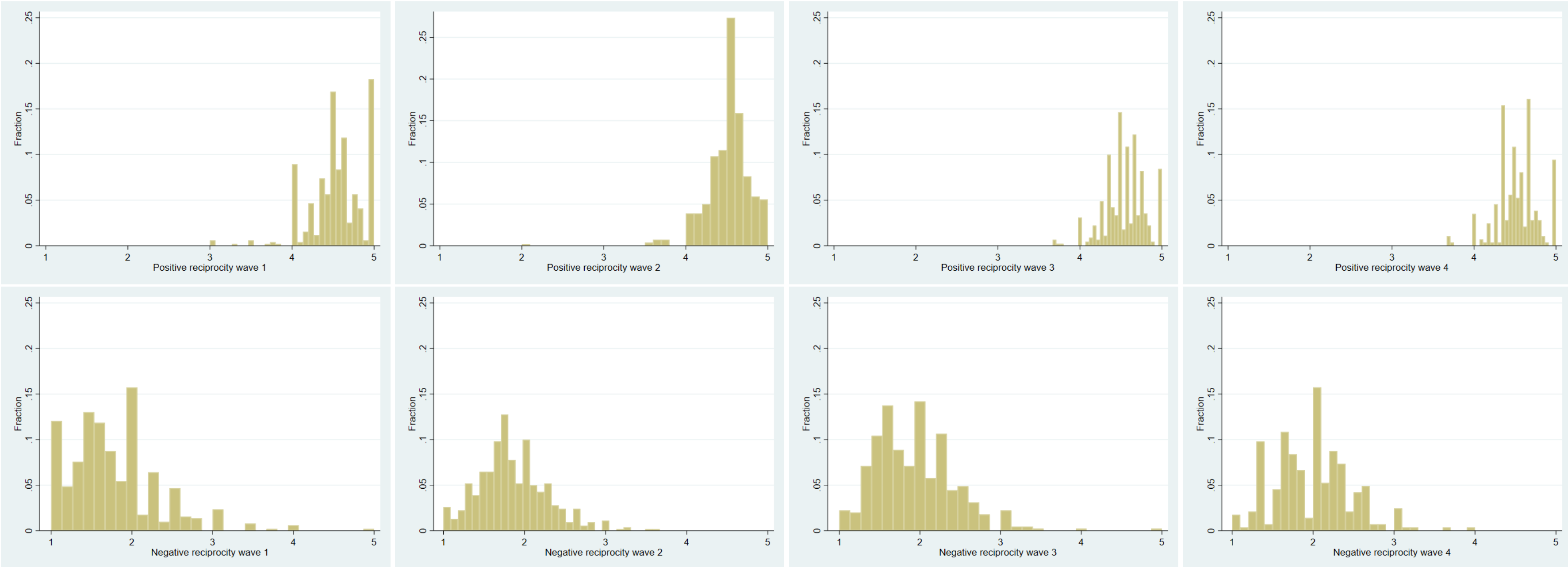
Source: Linked Personnel Panel, waves 1 to 4. N=2,002 in total. Wave 1: N=724, wave 2: N=541, wave: 3 N=451, wave: 4 N=286. Only establishments with at least 3 employee respondents used. One observation is the average response within one establishment-wave cell. Survey items shown in Table 4.3.

Fig. 4.11: Distribution of establishment-wave average trust and altruism by wave



Source: Linked Personnel Panel, waves 1 to 4. N=2,002 in total. Wave 1: N=724, wave 2: N=541, wave 3: N=451, wave: 4 N=286, except positive reciprocity, negative reciprocity, and altruism (each wave 1: N=515, wave 2: N=541, wave 3: N=451, wave 4: N=286), and time preferences (in wave1 N=515, wave 2 N=541, in wave 3 N=450, in wave 4 N=286). Only establishments with at least 3 employee respondents per wave used. One observation is the average response within one establishment-wave cell. Survey items shown in Table 4.3.

Fig. 4.12: Distribution of establishment-wave average positive reciprocity and negative reciprocity by wave



Source: Linked Personnel Panel, waves 1 to 4. N=2,002 in total. Wave 1: N=724, wave 2: N=541, wave 3: N=451, wave: 4 N=286, except positive reciprocity, negative reciprocity, and altruism (each wave 1: N=515, wave 2: N=541, wave 3: N=451, wave 4: N=286), and time preferences (in wave1 N=515, wave 2 N=541, in wave 3 N=450, in wave 4 N=286). Only establishments with at least 3 employee respondents per wave used. One observation is the average response within one establishment-wave cell. Survey items shown in Table 4.3.

4.7 Additional results

4.7.1 Leadership items

Table 4.9: Leadership items

Dep. variable:	Helping (1)	ASB (2)	Help given (3)	Help received (4)
Supervisor trust	-0.0537*	-0.0858***	-0.0560*	-0.0401
	(0.0307)	(0.0302)	(0.0313)	(0.0305)
Supervisor fairness	0.1240***	-0.2770***	0.0781***	0.1383***
	(0.0254)	(0.0248)	(0.0276)	(0.0244)
Supervisor understanding	0.1323***	-0.1331***	0.0730**	0.1548***
	(0.0310)	(0.0319)	(0.0331)	(0.0306)
Base controls	Yes	Yes	Yes	Yes
Preferences	Yes	Yes	Yes	Yes
Personality	Yes	Yes	Yes	Yes
AKM FE	Yes	Yes	Yes	Yes
Adj. R^2	0.23	0.32	0.18	0.23
Obs.	2,002	2,002	2,002	2,002

Results show regressions of mutual helping, ASB, and individual helping items on the leadership items used in the main leadership index in Table 4.3. All continuous independent variables are standardized with mean zero and unit variance. We analytically weight observations in the OLS regression by cell size (the number of individual observations per firm-wave cell). Our set of base controls includes time and risk preferences, self-efficacy, school education, vocational and university education, gender, age categories, partner, living alone, interview method, survey wave, industry, region, establishment size, and ownership type. Standard errors clustered on the establishment level in parentheses. The symbols *,**, and *** represent significance levels of 10%, 5%, and 1%, respectively.

4.7.2 Individual helping items

Table 4.10: Determinants of help given

Dep. variable:	Help given (std.)					
	(1)	(2)	(3)	(4)	(5)	(6)
Leadership	0.1186***			0.0771***	0.0609*	0.1121*
	(0.0237)			(0.0238)	(0.0349)	(0.0577)
Conscientiousness		0.0652**		0.0704***	0.0946**	
		(0.0263)		(0.0264)	(0.0371)	
Extroversion		0.0628**		0.0530**	0.0254	
		(0.0255)		(0.0256)	(0.0336)	
Neuroticism		-0.0360		-0.0237	-0.0564	
		(0.0257)		(0.0262)	(0.0351)	
Openness		0.0764***		0.0614**	0.0439	
		(0.0270)		(0.0266)	(0.0379)	
Agreeableness		0.0810***		0.0599**	0.0128	
		(0.0267)		(0.0272)	(0.0394)	
Trust			0.1056***	0.0753***	0.0746*	
			(0.0285)	(0.0287)	(0.0423)	
Positive recipr.			0.0695***	0.0602**	0.1083***	
			(0.0238)	(0.0235)	(0.0355)	
Negative recipr.			-0.0073	0.0186	0.0087	
			(0.0243)	(0.0245)	(0.0386)	
Altruism			0.0711***	0.0558**	0.0918**	
			(0.0259)	(0.0254)	(0.0390)	
Base controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual ability	Yes	Yes	Yes	Yes	Yes	Yes
Job controls	Yes	Yes	Yes	Yes	Yes	Yes
HRM practices	Yes	Yes	Yes	Yes	Yes	Yes
Lag helping	No	No	No	No	Yes	No
Firm fixed effects	No	No	No	No	No	Yes
Adj. R^2	0.14	0.16	0.15	0.17	0.20	0.31
Obs.	2,002	2,002	2,002	2,002	973	2,002

The dependent variable "Help given" is the standardized firm-wave average of the associated survey item. All continuous independent variables are standardized with mean zero and unit variance. We analytically weight observations in the OLS regression by cell size (the number of individual observations per firm-wave cell). Our set of base controls includes time and risk preferences, self-efficacy, school education, vocational and university education, gender, age categories, partner, living alone, interview method, survey wave, industry, region, establishment size, and ownership type. Job controls include the logarithm of the monthly net wage, white collar, task interdependence, management position, part-time, and fixed-term work contract. Standard errors clustered on the establishment level in parentheses. The symbols *,**, and *** represent significance levels of 10%, 5%, and 1%, respectively.

Table 4.11: Determinants of help received

Dep. variable:	Help received (std.)					
	(1)	(2)	(3)	(4)	(5)	(6)
Leadership	0.2624***			0.2108***	0.1916***	0.2393***
	(0.0231)			(0.0234)	(0.0344)	(0.0527)
Conscientiousness		-0.0305		-0.0175	0.0300	
		(0.0258)		(0.0249)	(0.0361)	
Extroversion		0.0541**		0.0385	0.0246	
		(0.0260)		(0.0252)	(0.0373)	
Neuroticism		-0.0757***		-0.0334	-0.0067	
		(0.0254)		(0.0251)	(0.0332)	
Openness		0.0154		-0.0103	-0.0311	
		(0.0276)		(0.0261)	(0.0375)	
Agreeableness		0.1271***		0.0822***	0.0601*	
		(0.0261)		(0.0251)	(0.0348)	
Trust			0.2295***	0.1614***	0.1346***	
			(0.0290)	(0.0292)	(0.0405)	
Positive recipr.			0.0336	0.0368	0.0812**	
			(0.0230)	(0.0232)	(0.0341)	
Negative recipr.			-0.0028	0.0145	-0.0051	
			(0.0257)	(0.0255)	(0.0379)	
Altruism			0.0627**	0.0477**	0.0703*	
			(0.0248)	(0.0236)	(0.0367)	
Base controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual ability	Yes	Yes	Yes	Yes	Yes	Yes
Job controls	Yes	Yes	Yes	Yes	Yes	Yes
HRM practices	Yes	Yes	Yes	Yes	Yes	Yes
Lag helping	No	No	No	No	Yes	No
Firm fixed effects	No	No	No	No	No	Yes
Adj. R^2	0.19	0.15	0.18	0.22	0.28	0.36
Obs.	2,002	2,002	2,002	2,002	973	2,002

The dependent variable “Help received” is the standardized firm-wave average of the associated survey item. All continuous independent variables are standardized with mean zero and unit variance. We analytically weight observations in the OLS regression by cell size (the number of individual observations per firm-wave cell). Our set of base controls includes time and risk preferences, self-efficacy, school education, vocational and university education, gender, age categories, partner, living alone, interview method, survey wave, industry, region, establishment size, and ownership type. Job controls include the logarithm of the monthly net wage, white collar, task interdependence, management position, part-time, and fixed-term work contract. Standard errors clustered on the establishment level in parentheses. The symbols *,**, and *** represent significance levels of 10%, 5%, and 1%, respectively.

4.7.3 Pooled analysis

Table 4.12: Determinants of mutual helping, pooled analysis

Dep. variable:	Helping index (std.)			
	(1)	(2)	(3)	(4)
Leadership	0.2427***			0.1888***
	(0.0312)			(0.0316)
Conscientiousness		0.0251		0.0279
		(0.0366)		(0.0355)
Extroversion		0.0506		0.0342
		(0.0344)		(0.0342)
Neuroticism		-0.0387		-0.0005
		(0.0315)		(0.0317)
Openness		0.0440		0.0141
		(0.0355)		(0.0330)
Agreeableness		0.1229***		0.0867***
		(0.0345)		(0.0333)
Trust			0.2199***	0.1706***
			(0.0374)	(0.0383)
Positive recipr.			0.0806***	0.0759**
			(0.0298)	(0.0303)
Negative recipr.			-0.0026	0.0219
			(0.0303)	(0.0295)
Altruism			0.0937***	0.0749**
			(0.0335)	(0.0329)
Base controls	Yes	Yes	Yes	Yes
Individual ability	Yes	Yes	Yes	Yes
Job controls	Yes	Yes	Yes	Yes
HRM practices	Yes	Yes	Yes	Yes
Adj. R^2	0.20	0.17	0.20	0.24
Obs.	1,003	1,003	1,003	1,003

The dependent variable Helping Index is an index containing the standardized pooled average of two items. All continuous independent variables are standardized with mean zero and unit variance. We analytically weight observations in the OLS regression by cell size (the number of individual observations per firm on average over all waves in this pooled analysis). Our set of base controls includes time and risk preferences, self-efficacy, school education, vocational and university education, gender, age categories, partner, living alone, interview method, survey wave, industry, region, establishment size, and ownership type. Job controls include the logarithm of the monthly net wage, white collar, task interdependence, management position, part-time, and fixed-term work contract. Standard errors clustered on the establishment level in parentheses. The symbols *,**, and *** represent significance levels of 10%, 5%, and 1%, respectively. This is a pooled analysis, where the firm averages of all variables are constructed using all survey waves.

Table 4.13: Determinants of antisocial behavior, pooled analysis

Dep. variable:	Antisocial behavior (std.)			
	(1)	(2)	(3)	(4)
Leadership	-0.4932***			-0.4550***
	(0.0284)			(0.0284)
Conscientiousness		0.0341		0.0424
		(0.0357)		(0.0306)
Extroversion		-0.0221		0.0007
		(0.0323)		(0.0276)
Neuroticism		0.1830***		0.1166***
		(0.0340)		(0.0305)
Openness		0.0131		0.0510*
		(0.0330)		(0.0281)
Agreeableness		-0.1076***		-0.0576*
		(0.0363)		(0.0315)
Trust			-0.2396***	-0.1095***
			(0.0327)	(0.0288)
Positive recipr.			-0.0226	-0.0394
			(0.0311)	(0.0261)
Negative recipr.			0.0156	-0.0022
			(0.0308)	(0.0258)
Altruism			-0.0299	-0.0140
			(0.0334)	(0.0271)
Base controls	Yes	Yes	Yes	Yes
Individual ability	Yes	Yes	Yes	Yes
Job controls	Yes	Yes	Yes	Yes
HRM practices	Yes	Yes	Yes	Yes
Adj. R^2	0.36	0.18	0.18	0.39
Obs.	1,003	1,003	1,003	1,003

The dependent variable is the standardized pooled firm average of antisocial behavior. All continuous independent variables are standardized with mean zero and unit variance. We analytically weight observations in the OLS regression by cell size (the number of individual observations per firm on average over all waves in this pooled analysis). Our set of base controls includes time and risk preferences, self-efficacy, school education, vocational and university education, gender, age categories, partner, living alone, interview method, survey wave, industry, region, establishment size, and ownership type. Job controls include the logarithm of the monthly net wage, white collar, task interdependence, management position, part-time, and fixed-term work contract. Standard errors clustered on the establishment level in parentheses. The symbols *,**, and *** represent significance levels of 10%, 5%, and 1%, respectively. This is a pooled analysis, where the firm averages of all variables are constructed using all survey waves.

4.7.4 Additional controls

Table 4.14: Determinants of mutual helping, additional controls

Dep. variable:	Helping index (std.)					
	(1)	(2)	(3)	(4)	(5)	(6)
Leadership	0.1907***			0.1424***	0.1322***	0.1661**
	(0.0369)			(0.0377)	(0.0361)	(0.0837)
Conscientiousness		0.0681*		0.0725*	0.0692*	
		(0.0398)		(0.0392)	(0.0366)	
Extroversion		0.0380		0.0258	0.0220	
		(0.0377)		(0.0380)	(0.0362)	
Neuroticism		-0.0500		-0.0236	-0.0261	
		(0.0374)		(0.0380)	(0.0349)	
Openness		0.0258		-0.0026	0.0053	
		(0.0432)		(0.0417)	(0.0386)	
Agreeableness		0.1129***		0.0627	0.0398	
		(0.0396)		(0.0406)	(0.0384)	
Trust			0.1740***	0.1406***	0.1192***	
			(0.0442)	(0.0449)	(0.0408)	
Positive recipr.			0.1221***	0.1109***	0.1113***	
			(0.0359)	(0.0362)	(0.0342)	
Negative recipr.			-0.0350	-0.0012	-0.0002	
			(0.0391)	(0.0417)	(0.0391)	
Altruism			0.1129***	0.0948**	0.0909**	
			(0.0414)	(0.0413)	(0.0384)	
Base controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual ability	Yes	Yes	Yes	Yes	Yes	Yes
Job controls	Yes	Yes	Yes	Yes	Yes	Yes
HRM practices	Yes	Yes	Yes	Yes	Yes	Yes
Lag helping	No	No	No	No	Yes	No
Firm fixed effects	No	No	No	No	No	Yes
Lagged org. outcomes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.20	0.19	0.22	0.24	0.28	0.50
Obs.	973	973	973	973	973	973

This specification additionally controls for the lag of firm-level employee outcomes. The dependent variable “Helping index” is an index containing the standardized firm-wave average of two items. All continuous independent variables are standardized with mean zero and unit variance. We analytically weight observations in the OLS regression by cell size (the number of individual observations per firm-wave cell). Our set of base controls includes time and risk preferences, self-efficacy, school education, vocational and university education, gender, age categories, partner, living alone, interview method, survey wave, industry, region, establishment size, and ownership type. Job controls include the logarithm of the monthly net wage, white collar, task interdependence, management position, part-time, and fixed-term work contract. Standard errors clustered on the establishment level in parentheses. The symbols *,**, and *** represent significance levels of 10%, 5%, and 1%, respectively.

Table 4.15: Determinants of antisocial behavior, additional controls

Dep. variable:	Antisocial behavior (ASB) (std.)					
	(1)	(2)	(3)	(4)	(5)	(6)
Leadership	-0.4586***			-0.4134***	-0.4115***	-0.4308***
	(0.0306)			(0.0303)	(0.0296)	(0.0871)
Conscientiousness		0.0486		0.0490	0.0523	
		(0.0391)		(0.0361)	(0.0351)	
Extroversion		0.0036		-0.0091	-0.0111	
		(0.0340)		(0.0307)	(0.0294)	
Neuroticism		0.2112***		0.1560***	0.1409***	
		(0.0374)		(0.0344)	(0.0325)	
Openness		0.0385		0.0639*	0.0608*	
		(0.0374)		(0.0333)	(0.0322)	
Agreeableness		-0.1260***		-0.0624*	-0.0571*	
		(0.0356)		(0.0346)	(0.0333)	
Trust			-0.2339***	-0.1218***	-0.1123***	
			(0.0363)	(0.0329)	(0.0323)	
Positive recipr.			-0.0242	-0.0300	-0.0310	
			(0.0415)	(0.0331)	(0.0327)	
Negative recipr.			0.0576*	0.0335	0.0332	
			(0.0346)	(0.0335)	(0.0323)	
Altruism			-0.0023	0.0184	0.0182	
			(0.0389)	(0.0350)	(0.0339)	
Base controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual ability	Yes	Yes	Yes	Yes	Yes	Yes
Job controls	Yes	Yes	Yes	Yes	Yes	Yes
HRM practices	Yes	Yes	Yes	Yes	Yes	Yes
Lag antisocial behavior	No	No	No	No	Yes	No
Firm fixed effects	No	No	No	No	No	Yes
Lagged org. outcomes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.30	0.19	0.19	0.35	0.36	0.50
Obs.	973	973	973	973	973	973

The dependent variable is the standardized firm-wave average of antisocial behavior. All continuous independent variables are standardized with mean zero and unit variance. We analytically weight observations in the OLS regression by cell size (the number of individual observations per firm-wave cell). Our set of base controls includes time and risk preferences, self-efficacy, school education, vocational and university education, gender, age categories, partner, living alone, interview method, survey wave, industry, region, establishment size, and ownership type. Job controls include the logarithm of the monthly net wage, white collar, task interdependence, management position, part-time, and fixed-term work contract. Standard errors clustered on the establishment level in parentheses. The symbols *,**, and *** represent significance levels of 10%, 5%, and 1%, respectively.

4.7.5 No individual sample refreshers

Table 4.16: Determinants of mutual helping, no sample refreshers

Dep. variable:	Helping index (std.)					
	(1)	(2)	(3)	(4)	(5)	(6)
Leadership	0.2197***			0.1703***	0.1452***	0.2003***
	(0.0294)			(0.0304)	(0.0419)	(0.0660)
Conscientiousness		-0.0336		-0.0108	0.0051	
		(0.0311)		(0.0303)	(0.0387)	
Extroversion		0.0491		0.0289	-0.0018	
		(0.0334)		(0.0330)	(0.0393)	
Neuroticism		-0.0584**		-0.0218	-0.0223	
		(0.0291)		(0.0293)	(0.0345)	
Openness		0.0519		0.0239	0.0170	
		(0.0374)		(0.0354)	(0.0447)	
Agreeableness		0.1465***		0.1133***	0.1231***	
		(0.0327)		(0.0322)	(0.0387)	
Trust			0.1705***	0.1087***	0.0567	
			(0.0334)	(0.0330)	(0.0361)	
Positive recipr.			0.0590*	0.0673**	0.1189***	
			(0.0308)	(0.0290)	(0.0423)	
Negative recipr.			-0.0102	0.0162	0.0186	
			(0.0295)	(0.0298)	(0.0406)	
Altruism			0.1291***	0.1041***	0.1494***	
			(0.0285)	(0.0274)	(0.0364)	
Base controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual ability	Yes	Yes	Yes	Yes	Yes	Yes
Job controls	Yes	Yes	Yes	Yes	Yes	Yes
HRM practices	Yes	Yes	Yes	Yes	Yes	Yes
Lag helping	No	No	No	No	Yes	No
Firm fixed effects	No	No	No	No	No	Yes
Adj. R^2	0.15	0.15	0.16	0.20	0.24	0.24
Obs.	1,703	1,703	1,703	1,703	877	1,703

The dependent variable Helping Index is an index containing the standardized firm-wave average of two items. All continuous independent variables are standardized with mean zero and unit variance. We analytically weight observations in the OLS regression by cell size (the number of individual observations per firm-wave cell). Our set of base controls includes time and risk preferences, self-efficacy, school education, vocational and university education, gender, age categories, partner, living alone, interview method, survey wave, industry, region, establishment size, and ownership type. Job controls include the logarithm of the monthly net wage, white collar, task interdependence, management position, part-time, and fixed-term work contract. Standard errors clustered on the establishment level in parentheses. The symbols *,**, and *** represent significance levels of 10%, 5%, and 1%, respectively. Sample refreshers are excluded from this analysis.

Table 4.17: Determinants of antisocial behavior, no sample refreshers

Dep. variable:	Antisocial behavior (std.)					
	(1)	(2)	(3)	(4)	(5)	(6)
Leadership	-0.3928***			-0.3429***	-0.3017***	-0.3226***
	(0.0286)			(0.0276)	(0.0344)	(0.0562)
Conscientiousness		0.0056		-0.0219	-0.0112	
		(0.0341)		(0.0312)	(0.0405)	
Extroversion		-0.0695**		-0.0377	-0.0234	
		(0.0302)		(0.0282)	(0.0362)	
Neuroticism		0.1969***		0.1310***	0.1439***	
		(0.0339)		(0.0310)	(0.0368)	
Openness		0.0025		0.0290	0.0317	
		(0.0313)		(0.0287)	(0.0376)	
Agreeableness		-0.0727**		-0.0307	-0.0516	
		(0.0327)		(0.0326)	(0.0408)	
Trust			-0.2430***	-0.1287***	-0.1193***	
			(0.0351)	(0.0312)	(0.0421)	
Positive recipr.			-0.0207	-0.0417	-0.0556	
			(0.0275)	(0.0259)	(0.0353)	
Negative recipr.			0.0227	0.0128	0.0198	
			(0.0315)	(0.0296)	(0.0384)	
Altruism			-0.0271	-0.0035	0.0411	
			(0.0300)	(0.0285)	(0.0373)	
Base controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual ability	Yes	Yes	Yes	Yes	Yes	Yes
Job controls	Yes	Yes	Yes	Yes	Yes	Yes
HRM practices	Yes	Yes	Yes	Yes	Yes	Yes
Lag antisocial behavior	No	No	No	No	Yes	No
Firm fixed effects	No	No	No	No	No	Yes
Adj. R^2	0.21	0.13	0.12	0.26	0.26	0.28
Obs.	1,703	1,703	1,703	1,703	877	1,703

The dependent variable is the standardized firm-wave average of antisocial behavior. All continuous independent variables are standardized with mean zero and unit variance. We analytically weight observations in the OLS regression by cell size (the number of individual observations per firm-wave cell). Our set of base controls includes time and risk preferences, self-efficacy, school education, vocational and university education, gender, age categories, partner, living alone, interview method, survey wave, industry, region, establishment size, and ownership type. Job controls include the logarithm of the monthly net wage, white collar, task interdependence, management position, part-time, and fixed-term work contract. Standard errors clustered on the establishment level in parentheses. The symbols *,**, and *** represent significance levels of 10%, 5%, and 1%, respectively. Sample refreshers are excluded from this analysis.

4.7.6 Balanced panel of firms

Table 4.18: Determinants of mutual helping, balanced panel of firms

Dep. variable:	Helping index (std.)					
	(1)	(2)	(3)	(4)	(5)	(6)
Leadership	0.2166***			0.1429***	0.1282***	0.2230***
	(0.0345)			(0.0362)	(0.0416)	(0.0570)
Conscientiousness		0.0271		0.0565	0.1012**	
		(0.0381)		(0.0362)	(0.0434)	
Extroversion		0.0760**		0.0544	0.0305	
		(0.0353)		(0.0343)	(0.0378)	
Neuroticism		-0.0461		-0.0104	0.0187	
		(0.0348)		(0.0349)	(0.0385)	
Openness		0.0338		0.0098	0.0059	
		(0.0414)		(0.0401)	(0.0470)	
Agreeableness		0.1608***		0.1215***	0.1097***	
		(0.0382)		(0.0364)	(0.0420)	
Trust			0.2197***	0.1592***	0.1228***	
			(0.0390)	(0.0388)	(0.0438)	
Positive recipr.			-0.0069	-0.0040	0.0563	
			(0.0323)	(0.0315)	(0.0438)	
Negative recipr.			-0.0122	0.0351	0.0743	
			(0.0334)	(0.0326)	(0.0454)	
Altruism			0.0986***	0.0807**	0.0653	
			(0.0356)	(0.0348)	(0.0459)	
Base controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual ability	Yes	Yes	Yes	Yes	Yes	Yes
Job controls	Yes	Yes	Yes	Yes	Yes	Yes
HRM practices	Yes	Yes	Yes	Yes	Yes	Yes
Lag helping	No	No	No	No	Yes	No
Firm fixed effects	No	No	No	No	No	Yes
Adj. R^2	0.26	0.26	0.26	0.30	0.34	0.45
Obs.	924	924	924	924	623	924

The dependent variable Helping Index is an index containing the standardized firm-wave average of two items. All continuous independent variables are standardized with mean zero and unit variance. We analytically weight observations in the OLS regression by cell size (the number of individual observations per firm-wave cell). Our set of base controls includes time and risk preferences, self-efficacy, school education, vocational and university education, gender, age categories, partner, living alone, interview method, survey wave, industry, region, establishment size, and ownership type. Job controls include the logarithm of the monthly net wage, white collar, task interdependence, management position, part-time, and fixed-term work contract. Standard errors clustered on the establishment level in parentheses. The symbols *,**, and *** represent significance levels of 10%, 5%, and 1%, respectively. This analysis only uses a balanced panel of firms.

Table 4.19: Determinants of antisocial behavior, balanced panel of firms

Dep. variable:	Antisocial behavior (std.)					
	(1)	(2)	(3)	(4)	(5)	(6)
Leadership	-0.4733***			-0.4067***	-0.3762***	-0.4569***
	(0.0311)			(0.0305)	(0.0356)	(0.0555)
Conscientiousness		0.0788**		0.0567	0.0790*	
		(0.0396)		(0.0358)	(0.0458)	
Extroversion		-0.0933**		-0.0504	-0.0270	
		(0.0367)		(0.0324)	(0.0381)	
Neuroticism		0.2256***		0.1609***	0.1961***	
		(0.0373)		(0.0353)	(0.0400)	
Openness		-0.0047		0.0318	0.0279	
		(0.0417)		(0.0355)	(0.0444)	
Agreeableness		-0.1304***		-0.0569	-0.0434	
		(0.0385)		(0.0364)	(0.0416)	
Trust			-0.3059***	-0.1333***	-0.1247***	
			(0.0391)	(0.0360)	(0.0480)	
Positive recipr.			-0.0010	-0.0412	-0.0330	
			(0.0369)	(0.0326)	(0.0453)	
Negative recipr.			0.0441	0.0219	0.0788*	
			(0.0380)	(0.0333)	(0.0440)	
Altruism			-0.0314	-0.0061	0.0747	
			(0.0345)	(0.0325)	(0.0467)	
Base controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual ability	Yes	Yes	Yes	Yes	Yes	Yes
Job controls	Yes	Yes	Yes	Yes	Yes	Yes
HRM practices	Yes	Yes	Yes	Yes	Yes	Yes
Lag antisocial behavior	No	No	No	No	Yes	No
Firm fixed effects	No	No	No	No	No	Yes
Adj. R^2	0.30	0.18	0.18	0.35	0.37	0.44
Obs.	924	924	924	924	623	924

The dependent variable is the standardized firm-wave average of antisocial behavior. All continuous independent variables are standardized with mean zero and unit variance. We analytically weight observations in the OLS regression by cell size (the number of individual observations per firm-wave cell). Our set of base controls includes time and risk preferences, self-efficacy, school education, vocational and university education, gender, age categories, partner, living alone, interview method, survey wave, industry, region, establishment size, and ownership type. Job controls include the logarithm of the monthly net wage, white collar, task interdependence, management position, part-time, and fixed-term work contract. Standard errors clustered on the establishment level in parentheses. The symbols *,**, and *** represent significance levels of 10%, 5%, and 1%, respectively. This analysis only uses a balanced panel of firms.

4.7.7 No missing indicators

Table 4.20: Determinants of mutual helping, no missing indicators

Dep. variable:	Helping index (std.)					
	(1)	(2)	(3)	(4)	(5)	(6)
Leadership	0.2159***			0.1627***	0.1571***	0.2400***
	(0.0272)			(0.0276)	(0.0354)	(0.0578)
Conscientiousness		0.0329		0.0350	0.0751*	
		(0.0304)		(0.0303)	(0.0394)	
Extroversion		0.0660**		0.0531*	0.0349	
		(0.0292)		(0.0291)	(0.0364)	
Neuroticism		-0.0710***		-0.0405	-0.0115	
		(0.0275)		(0.0272)	(0.0339)	
Openness		0.0366		0.0097	0.0044	
		(0.0309)		(0.0293)	(0.0401)	
Agreeableness		0.1321***		0.1003***	0.0575	
		(0.0294)		(0.0299)	(0.0407)	
Trust			0.1817***	0.1230***	0.1128***	
			(0.0327)	(0.0331)	(0.0423)	
Positive recipr.			0.0821***	0.0795***	0.1219***	
			(0.0243)	(0.0246)	(0.0356)	
Negative recipr.			-0.0021	0.0299	0.0135	
			(0.0254)	(0.0259)	(0.0419)	
Altruism			0.0649**	0.0444*	0.0673*	
			(0.0276)	(0.0268)	(0.0386)	
Base controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual ability	Yes	Yes	Yes	Yes	Yes	Yes
Job controls	Yes	Yes	Yes	Yes	Yes	Yes
HRM practices	Yes	Yes	Yes	Yes	Yes	Yes
Lag helping	No	No	No	No	Yes	No
Firm fixed effects	No	No	No	No	No	Yes
Adj. R^2	0.20	0.19	0.19	0.23	0.27	0.39
Obs.	1,494	1,494	1,494	1,494	848	1,494

The dependent variable “Helping index” is an index containing the standardized firm-wave average of two items. All continuous independent variables of interest are standardized with mean zero and unit variance. We analytically weight observations in the OLS regression by cell size (the number of individual observations per firm-wave cell). Our set of base controls includes time and risk preferences, self-efficacy, school education, vocational and university education, gender, age categories, partner, living alone, interview method, survey wave, industry, region, establishment size, and ownership type. Job controls include the logarithm of the monthly net wage, white collar, task interdependence, management position, part-time, and fixed-term work contract. Individual ability is proxied by the firm-wave average of individual AKM fixed effects from the LIAB data, as calculated from 2010 to 2017. Standard errors clustered on the establishment level in parentheses. The symbols *,**, and *** represent significance levels of 10%, 5%, and 1%, respectively. This sample does not include missing indicators, and thus removes any observations with at least one missing value.

Table 4.21: Determinants of antisocial behavior, no missing indicators

Dep. variable:	Antisocial behavior (ASB) (std.)					
	(1)	(2)	(3)	(4)	(5)	(6)
Leadership	-0.4650***			-0.4238***	-0.4001***	-0.4863***
	(0.0230)			(0.0230)	(0.0301)	(0.0586)
Conscientiousness		0.0251		0.0299	0.0643*	
		(0.0313)		(0.0277)	(0.0384)	
Extroversion		-0.0300		-0.0132	-0.0028	
		(0.0292)		(0.0254)	(0.0323)	
Neuroticism		0.1876***		0.1192***	0.1487***	
		(0.0295)		(0.0272)	(0.0345)	
Openness		0.0055		0.0493**	0.0604*	
		(0.0284)		(0.0245)	(0.0329)	
Agreeableness		-0.1163***		-0.0543*	-0.0536	
		(0.0315)		(0.0280)	(0.0358)	
Trust			-0.2322***	-0.1002***	-0.0978***	
			(0.0294)	(0.0256)	(0.0340)	
Positive recipr.			0.0110	-0.0083	-0.0290	
			(0.0253)	(0.0226)	(0.0335)	
Negative recipr.			0.0433	0.0267	0.0435	
			(0.0273)	(0.0244)	(0.0336)	
Altruism			-0.0316	-0.0107	0.0221	
			(0.0282)	(0.0246)	(0.0372)	
Base controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual ability	Yes	Yes	Yes	Yes	Yes	Yes
Job controls	Yes	Yes	Yes	Yes	Yes	Yes
HRM practices	Yes	Yes	Yes	Yes	Yes	Yes
Lag antisocial behavior	No	No	No	No	Yes	No
Firm fixed effects	No	No	No	No	No	Yes
Adj. R^2	0.30	0.14	0.15	0.32	0.35	0.44
Obs.	1,494	1,494	1,494	1,494	848	1,494

The dependent variable is the standardized firm-wave average of antisocial behavior. All continuous independent variables of interest are standardized with mean zero and unit variance. We analytically weight observations in the OLS regression by cell size (the number of individual observations per firm-wave cell). Our set of base controls includes time and risk preferences, self-efficacy, school education, vocational and university education, gender, age categories, partner, living alone, interview method, survey wave, industry, region, establishment size, and ownership type. Job controls include the logarithm of the monthly net wage, white collar, task interdependence, management position, part-time, and fixed-term work contract. Individual ability is proxied by the firm-wave average of individual AKM fixed effects from the LIAB data, as calculated from 2010 to 2017. Standard errors clustered on the establishment level in parentheses. The symbols *,**, and *** represent significance levels of 10%, 5%, and 1%, respectively. This sample does not include missing indicators, and thus removes any observations with at least one missing value.

4.7.8 Restricted sample

Table 4.22: Determinants of mutual helping, restricted sample

Dep. variable:	Helping index (std.)			
	(1)	(2)	(3)	(4)
Leadership	0.2282***			0.1699***
	(0.0351)			(0.0363)
Conscientiousness		0.0641		0.0700*
		(0.0402)		(0.0391)
Extroversion		0.0502		0.0352
		(0.0378)		(0.0378)
Neuroticism		-0.0757**		-0.0361
		(0.0376)		(0.0382)
Openness		0.0282		-0.0022
		(0.0430)		(0.0411)
Agreeableness		0.1235***		0.0696*
		(0.0400)		(0.0408)
Trust			0.1988***	0.1466***
			(0.0444)	(0.0452)
Positive recipr.			0.1120***	0.1015***
			(0.0363)	(0.0364)
Negative recipr.			-0.0341	0.0034
			(0.0401)	(0.0421)
Altruism			0.1144***	0.0939**
			(0.0422)	(0.0416)
Base controls	Yes	Yes	Yes	Yes
Individual ability	Yes	Yes	Yes	Yes
Job controls	Yes	Yes	Yes	Yes
HRM practices	Yes	Yes	Yes	Yes
Lag helping	No	No	No	No
Firm fixed effects	No	No	No	No
Adj. R^2	0.19	0.18	0.20	0.23
Obs.	973	973	973	973

The sample here is from the LDV estimation in main results. The dependent variable "Helping index" is an index containing the standardized firm-wave average of two helping items. All continuous independent variables of interest are standardized with mean zero and unit variance. We analytically weight observations in the OLS regression by cell size (the number of individual observations per firm-wave cell). Our set of base controls includes time and risk preferences, self-efficacy, school education, vocational and university education, gender, age categories, partner, living alone, interview method, survey wave, industry, region, establishment size, and ownership type. Job controls include the logarithm of the monthly net wage, white collar, task interdependence, management position, part-time, and fixed-term work contract. Individual ability is proxied by the firm-wave average of individual AKM fixed effects from the LIAB data, as calculated from 2010 to 2017. Standard errors clustered on the establishment level in parentheses. The symbols *,**, and *** represent significance levels of 10%, 5%, and 1%, respectively. The sample here is from the LDV estimation in main results.

Table 4.23: Determinants of antisocial behavior, restricted sample

Dep. variable:	Antisocial behavior (std.)			
	(1)	(2)	(3)	(4)
Leadership	-0.4800***			-0.4229***
	(0.0288)			(0.0289)
Conscientiousness		0.0543		0.0539
		(0.0404)		(0.0363)
Extroversion		-0.0084		-0.0088
		(0.0346)		(0.0309)
Neuroticism		0.2429***		0.1654***
		(0.0384)		(0.0353)
Openness		0.0421		0.0724**
		(0.0384)		(0.0330)
Agreeableness		-0.1304***		-0.0576
		(0.0372)		(0.0352)
Trust			-0.2661***	-0.1254***
			(0.0371)	(0.0328)
Positive recipr.			-0.0200	-0.0321
			(0.0405)	(0.0321)
Negative recipr.			0.0629*	0.0390
			(0.0359)	(0.0336)
Altruism			-0.0027	0.0165
			(0.0399)	(0.0350)
Base controls	Yes	Yes	Yes	Yes
Individual ability	Yes	Yes	Yes	Yes
Job controls	Yes	Yes	Yes	Yes
HRM practices	Yes	Yes	Yes	Yes
Lag antisocial behavior	No	No	No	No
Firm fixed effects	No	No	No	No
Adj. R^2	0.30	0.17	0.16	0.34
Obs.	973	973	973	973

The sample here is from the LDV estimation in main results. The dependent variable is the standardized firm-wave average of antisocial behavior. All continuous independent variables of interest are standardized with mean zero and unit variance. We analytically weight observations in the OLS regression by cell size (the number of individual observations per firm-wave cell). Our set of base controls includes time and risk preferences, self-efficacy, school education, vocational and university education, gender, age categories, partner, living alone, interview method, survey wave, industry, region, establishment size, and ownership type. Job controls include the logarithm of the monthly net wage, white collar, task interdependence, management position, part-time, and fixed-term work contract. Individual ability is proxied by the firm-wave average of individual AKM fixed effects from the LIAB data, as calculated from 2010 to 2017. Standard errors clustered on the establishment level in parentheses. The symbols *,**, and *** represent significance levels of 10%, 5%, and 1%, respectively.

4.7.9 Outcomes

Table 4.24: Mutual helping, antisocial behavior, and organizational outcomes

	(1) Engage-ment$_{t+1}$	(2) Commit-ment$_{t+1}$	(3) Job satis-faction$_{t+1}$	(4) Turnover intention$_{t+1}$	(5) Sick days$_{t+1}$	(6) Firm productivity$_{t+1}$
Panel A: Helping						
Helping	0.1117***	0.1188***	0.1557***	-0.1209***	0.0068	0.0538***
	(0.0383)	(0.0349)	(0.0372)	(0.0319)	(0.0341)	(0.0162)
Base controls	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.15	0.21	0.12	0.13	0.09	0.61
Obs.	1,267	1,267	1,267	1,267	1,267	1,003
Panel B: Antisocial behavior						
Antisocial beh.	-0.0958***	-0.1410***	-0.2114***	0.1990***	0.0683**	-0.0116
	(0.0307)	(0.0303)	(0.0311)	(0.0318)	(0.0313)	(0.0167)
Base controls	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.14	0.22	0.14	0.16	0.09	0.61
Obs.	1,267	1,267	1,267	1,267	1,267	1,003

This table shows regressions of several employee outcomes in $t+1$ on mutual helping and antisocial behavior. All continuous independent variables are standardized with mean zero and unit variance. We analytically weight observations in the OLS regression by cell size (the number of individual observations per firm-wave cell). The dependent variables in columns 1 to 5 are standardized firm-wave weighted averages of employee outcomes. In column 6 the dependent variable is the standardized AKM firm fixed effect estimated using the LIAB data as estimated from 2010 to 2017. The sample used is the main regression sample in columns 1-5, and the sample used in column 6 is collapsed at the firm level over time, since the AKM FE is only measured once per firm over the sample period, additionally including wave 1 observations. Our set of base controls includes time and risk preferences, self-efficacy, school education, vocational and university education, gender, age categories, partner, living alone, interview method, survey wave, industry, region, establishment size, and ownership type, as well as missing dummies. Standard errors (clustered on the establishment level) are reported in parentheses. The symbols *,**, and *** represent significance levels of 10%, 5%, and 1%, respectively.

Chapter 5
Improving the availability of unrelated stem cell donors: Evidence from a major donor registry *

5.1 Introduction

For many patients suffering from leukemia or other blood diseases who lack a related donor, a hematopoietic stem cell transplantation (HSCT) from a matching unrelated donor offers the best treatment and chance of survival. HSCTs have extended the lifespan of hundreds of thousands of patients worldwide and enhanced their quality of life (e.g., Gratwohl et al., 2015). However, the unavailability of potential stem cell donors who are in a position to make an actual donation to a patient is a critical challenge hampering the stem cell donation process and, consequentially, reducing the chances of survival for many patients.

Stem cell donation is a multi-stage process. First, individuals willing to donate stem cells sign up to a registry. Then, potential donors wait for a matching patient, which might take several years, if ever.[2] Once a matching patient is found, a request for confirmatory typing (CT; sometimes called verification typing) by a transplant center (typically via a central registry), is made to the potential donor (e.g., Bergstrom et al., 2009; Lacetera et al., 2014; Dasgupta, 2018; Heger et al., 2020). The CT stage is thus *the* crucial milestone in the process of actually becoming a stem cell donor.[3] Unfortunately, stem cell donor registries around the world are faced with considerable donor attrition at the CT stage. This attrition leads to delays in donor search and increases in the time to transplant for many patients, which can negatively affect survival rates (Lown et al., 2014). Moreover, attrition at CT causes inefficiencies in resource allocation because of the costs of recruiting and handling donors who ultimately are not available for transplantation (Anthias et al., 2020). For example, in the case of DKMS Germany, a major stem cell donor registry and the organization that provided the data for this study, the average attrition rate of potential donors from a CT request was 23% in 2018 (the final year of observation).[4] Thus, it is important for stem cell donor registry managers to understand whether, and how their recruitment and retention practices ultimately affect the availability of donors who are a match for a patient to follow through with the stem cell donation process.

In this paper, we focus on the challenge of stem cell donor registries to maintain the motivation and commitment of already registered, potential donors until they ultimately receive a CT request. The multi-stage process of stem cell collection is a unique setting that differs from other medical donations such as those of blood or plasma, where, typically, the commitment to donate is immediately followed by the actual donation. Stem cell donors, on the other hand, make a non-binding commitment to potentially donate somewhere in the (near or even distant) future.

* This chapter is based on joint work with Patrick Kampkötter, Mario Macis, Jürgen Sauter, Susanne Seitz, Alexander H. Schmidt, Robert Slonim, and Daniel Wiesen not yet published.

[2] The average lifetime donation probability for donors who register with DKMS Germany at age 18 amounts to around 4%, while this number can be much higher for potential donors with unique biological characteristics.

[3] The CT is a mandatory and decisive point for the potential donor to decide whether to actually follow through with a donation. At the CT stage, information is collected and tests are performed to confirm whether the donor is genetically suitable and medically eligible for HSCT and still willing to undergo a stem cell collection procedure.

[4] Attrition in Germany is lower compared to the UK, with 38% attrition rate in 2011 (Lown et al., 2014) and more recently 35% estimated among Europeans and 56% among non-Europeans (Anthias et al., 2020). Attrition rate in Germany is also lower compared to the US, with estimates of 36% (Gragert et al., 2014), 40% (Switzer et al., 2013), 47% (Lown et al., 2014), and most recently 50% estimated using a machine learning algorithm (Sivasankaran et al., 2018). This emphasizes the large medical and economic benefits from improving the availability of stem cell donors and makes it a pressing issue for donor registries.

Hence, the donation process is characterized by uncertainty about whether or not an actual donation opportunity will materialize and, if so, when. This setup makes self-commitment problems more likely.

We analyze a set of initiatives implemented by DKMS Germany, a leading stem cell donor registry with more than 10 million registered donors (Schmidt et al., 2020a), that attempt to address this challenge and/or may have an impact on donors' self-commitment. The initiatives are directed to individuals who have signed up to the donor registry, but who have not (yet) been requested for CT. Participation in the initiatives involves some costly action for the potential donor. Specifically, the registry asks donors to send in own biological material or schedule a blood draw for genetic retyping ("prospective" initiatives to obtain a higher resolution human leukocyte antigen (HLA) profile), and/or to report periods of unavailability in advance ("status update" initiatives). Low- and intermediate resolution HLA typing and donor unavailability at the CT stage are among the main reasons for registries' inability to provide suitable donors for patients. Retyping initiatives are helpful, because they allow the registry to obtain more accurate genetic information about potential donors. With status updates, the registry acquires important information about potential donors' (un)availability, which can be used to optimize donor prioritization. These initiatives can potentially have additional beneficial effects through at least two channels. First, voluntary participation in an initiative may signal a donor's level of commitment; if that is the case, participation might allow the registry to identify a select group of donors who are more likely to be readily available when called upon to make an actual donation. Second, being invited to an initiative might cause potential donors to become more committed or motivated, which may in turn have a positive effect on availability to make an actual donation. However, if the initiatives are perceived as burdensome by registered donors, they could potentially reduce motivation or commitment. Therefore, whether the initiatives have an overall positive or negative effect on availability at CT is an empirical question.

Understanding the impact of these initiatives can help answer important questions relevant for both stem cell donor registries and scholars interested in donation behavior. Despite the potential importance of such initiatives, empirical evidence on how they influence donor availability at the CT stage is scarce. We address this gap in the literature by empirically analyzing 91,479 CT requests issued to potential donors registered with DKMS Germany for the time period 2013 to 2018. The data include a rich set of donor- and registry-related variables that determine being invited to participate in the initiatives. We investigate whether the initiatives have an effect on donor availability at CT (direct effect) as well as whether donors who chose to actively participate in the initiatives are more likely to be available at CT (selection effect). Invitation to the initiatives was based entirely on observable, pre-determined (i.e., exogenous) characteristics of potential donors such as biological traits (e.g., age and sex) and genetic characteristics. This implies that invitees could not influence their probability of receiving an invitation letter to an initiative. Importantly, for cost reasons, not all potential candidates for receiving invitations by letter were invited, providing a large comparison group with substantial overlap in terms of observables. This means selection variables do not perfectly predict invitation, providing unsystematic variation in the invitations.

We perform the following empirical analyses. First, we evaluate the impact of retyping and status update initiatives on potential donors' CT availability. This analysis is unconditional on donor participation in these initiatives; in particular, we implement an intention-to-treat approach (ITT). Second, we analyze the predictive power of participation in the initiatives for CT availability, by exploiting information on the participation or non-participation of registered donors in these initiatives. Third, we estimate a local average treatment effect (LATE) using the invitation as an instrument for participation, to test whether there is an impact of initiatives on CT availability for the participants, net of observable selection effects.

We find that the CT availability of donors invited to a retyping initiative or to an initiative that involved both retyping and status update were 3.2 and 2.5 percentage points (pp), respectively, higher than of those who received no invitation (both significant at the one percent level). Since baseline attrition at the CT stage is 22.9%, this corresponds to a 14.0% (3.2/22.9) and 10.9% (2.5/22.9) reduction in attrition. However, simply asking for status updates without retyping did not yield significant effects. We also estimate predictive effects of participation, and find that participants in all of the initiatives are 13-15pp more likely to follow through with CT than non-participants.

The LATE estimates indicate that the retyping initiative and the retyping plus status update initiative led to 4.3pp and 8.2pp higher CT availability, respectively (both statistically significant at the one percent level); the status update initiative alone increased availability by 3.8pp, but this estimated effect was not statistically significant. Thus, the initiatives reduce attrition by between 16.6% (3.8/22.9) and 35.8% (8.2/22.9) for participants. Also, comparing the LATE to the selection effects for participants, our findings imply that between 40% and 60% of the availability increase of participants is due to the initiatives' causal impact, with the remaining portion being a selection effect. The results are robust to several robustness checks; in particular, using the methodology of Oster (2019), we conclude that our findings would still hold under reasonable assumptions about unobserved selection.

Our study contributes to several streams of literature. First, we contribute to the economics literature on medical donations. Unlike other medical donations such as blood and organs (e.g., Roth et al., 2004; Craig et al., 2017), the stem cell donation process varies significantly in several aspects. First, the point in time of stating the willingness to give is different from the point in time of the decision to actually donate. In case of whole blood (and plasma), for example, the donation takes place immediately after the registration and a health check (e.g., Wildman and Hollingsworth, 2009; Stutzer et al., 2011; Lacetera et al., 2012; Slonim et al., 2014; Goette and Stutzer, 2020). Second, for stem cell donations, it is not uncommon that several years lie between registration and a first CT request. Compared to organ donation, where most transplants are from cadavers, stem cell donations are from living donors (Grieco et al., 2018).[5] Finally, donating stem cells is typically a high-stakes decision with life-saving implications for the recipient. Also, donating stem cells via peripheral blood stem cell collection or extraction from the iliac crest is generally safe (Schmidt et al., 2017), but nevertheless carries larger risks and discomfort than blood donation.

In the context of designing volunteer markets, the blood donation literature has shown that creating a registry for blood donors has substantial benefits. Heger et al. (2020) find in a field experiment with blood donors that when there are shortages in volunteer markets for blood donation, creating a registry can increase responsiveness by 66%. However, with stem cell donation, the much longer and uncertain time horizon makes it difficult for registries to maintain donors' commitment and, ultimately, to guarantee their CT availability and donation. The initiatives we analyze in this paper attempt to address this problem, and take place during the potentially long interim stage after registration, but before an actual donation opportunity arises.

Findings on the effects of initiatives commonly conducted by stem cell donor registries, such as those we analyze in this paper, yield practical insights for the management of donor registries. In particular, we find that the association between CT availability and registered donors' participation in an initiative is positive and larger in magnitude than the negative association of non-participants. Donors' decision to participate in an initiative, which requires a status update, yields practical information to the registry which would otherwise not be available. Based on reported periods of future unavailability of participating donors, for example, the registry is able to block the donor for the corresponding time period and this donor will not be requested by search coordinators of transplant centers. This would obviously render the donation process more efficient. As a consequence, donor registries should be given the possibility to make such information available to search coordinators. By uncovering these empirical relations, our paper makes a novel contribution to the literature and is of direct relevance to donor registries. Besides blood donation, our findings may also offer insights to other contexts where altruistic individuals sign up to a registry expressing the intention to make a contribution should the need arise; examples include volunteer teaching (Coffman et al., 2017) and other volunteer work in general (Exley and Petrie, 2018).

Second, we contribute to the medical and health literature that assesses motivations of stem cell donors for attrition at the CT stage. While the number of potential donors enrolled in registries like the DKMS has increased

[5] For example, from January to June 2021, 21,061 organ transplants were performed in the United States of America, 17,821 of which came from deceased donors, and 3,240 of which came from living donors: (https://optn.transplant.hrsa.gov/data/, accessed on August 5, 2021). Of the 24,522 unrelated stem cell transplants performed 2014-2018 in the US, 14% were from cord blood, the rest being from peripheral blood after stem cell mobilization or from bone marrow. Data retrieved on August 5, 2021, from https://bloodstemcell.hrsa.gov/data/donation-and-transplantation-statistics/transplant-initiative-report.

steadily (see, for instance, Bergstrom et al., 2009; Schmidt et al., 2020a), it has been widely documented that a significant number of donors across many countries withdraw their consent for donation after registration.

Several papers assess factors that relate to the attrition at the CT stage. Correlates of attrition are ethnic background (Switzer et al., 2004; Myaskovsky et al., 2004; Onitilo et al., 2004; Lown et al., 2014), which can be due to religious and medical objections to donation, less trust that stem cells would be allocated equitably, and a greater likelihood of having been discouraged from donating (e.g., Onitilo et al., 2004; Switzer et al., 2013). Other correlates include time in the registry (Switzer et al., 1999; Monaghan et al., 2021), whether registration was patient-centered (Switzer et al., 2004), age (Switzer et al., 2013), sex (Lown et al., 2014; Fingrut et al., 2018), and communication of being the only known donor match to the patient (Switzer et al., 2018). Anthias et al. (2020) find that donor's mental and physical health, as well as interaction with the registry correlate with CT availability. Also, Switzer et al. (2004) find that intrinsic motivation to donate, realistic expectations about donation, and more contact with the registry is associated with being available. Ambivalence about donation in the form of doubt and uncertainty, or wishing someone else would donate instead, is a strong driver of attrition across all ethnic groups (Switzer et al., 2013). Other papers assess the donors' motivations for registering with a stem cell donor registry (Switzer et al., 2003; Aurelio et al., 2011; Mclaren et al., 2012; Bart et al., 2014) and find that these motivations also affect CT availability. As a result, intrinsic registration motives predict much higher donation availability than extrinsic motives (e.g., social pressure or incentives) (La Casta et al., 2019).

Importantly, more time from diagnosis to transplant may adversely affect patient outcomes. Reasons for not being able to proceed to stem cell donation though a matching donor is registered include, notably, low and intermediate resolution or incomplete HLA typing (Sauter et al., 2016) and, more importantly, donor attrition at the CT stage (Lown and Shaw, 2013). Thus, it is interesting whether the retyping initiatives, which should increase the possibility of finding a matching donor, are related to the attrition at the CT stage. We are one of the first to empirically assess the impact of initiatives requiring a status update, which are aimed at identifying a select group of donors to report unavailability in advance, so they can be removed from searches during this time.

The remainder of the paper is organized as follows. Section 5.2 provides further background on the process from enrolling to becoming a stem cell donor and on the retyping and status update initiatives. In Section 5.3, we describe our data set and present initial descriptive analyses. Section 5.4 presents our estimation results. Finally, section 5.6 describes implications of our findings, discusses limitations, and concludes.

5.2 Background

5.2.1 The process from registering with a stem cell donor registry to donating

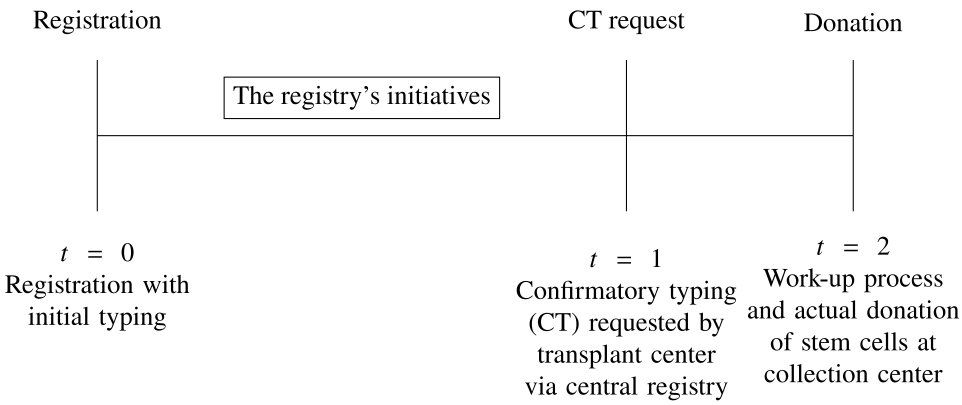
Typically, stem cell donations take two forms. They are either autologous donations, which means that stem cells are extracted from the patient before a cancer treatment and re-transplanted after the treatment, or allogeneic, where stem cells are collected from another person (related or unrelated). To put this into perspective, in the United States, for example, about 40% of stem cell transplants between 2015 and 2019 were allogeneic.[6] We consider the stem cell donor registry DKMS Germany, that provides unrelated donor data for stem cell donor searches for patients in need of an unrelated allogeneic transplant. DKMS is one of the world's largest registries[7] with more than eleven million registered donors in seven countries: Chile, Germany, India, Poland, South Africa, United Kingdom, and the United States (more than 65% of registered donors are registered in Germany). Since its foundation in 1991, DKMS has facilitated over 95,000 stem cell collections, and its share of all unrelated stem cell donations worldwide

[6] See https://bloodstemcell.hrsa.gov/data/donation-and-transplantation-statistics.
[7] See https://statistics.wmda.info for an overview.

actually amounts to about 40% (Schmidt et al., 2020a).

Figure 5.1 illustrates the multi-stage process from registering with a stem cell donor registry to actually becoming a stem cell donor. First, potential donors join the registry by providing oral mucosa cells via buccal swabs or small blood samples, which are then typed for human leukocyte antigens (HLA) by a genotyping service provider (Schöfl et al., 2017). Donors' typing results are then listed on a computerized registry ($t = 0$). Potential occasions to register include public community drives, which might evolve around a specific patient requiring a donation (patient-centered), company drives, donor drives targeted at specific populations such as students, visitors of sports events, police, or military staff (special projects), or online registration.[8]

Fig. 5.1: Overview of key events for registered stem cell donors



The quality of a match between a potential donor and a patient is mainly determined by the fit in their genetic information (HLA type).[9] Besides the HLA type, age and genetic parameters such as the ABO blood group and markers as the cytomegalovirus (CMV) antibody status[10] are secondary criteria to select a suitable donor. DKMS has been recruiting potential donors for about 30 years. During this period, the medically desirable and technically and financially feasible methods of typing prospective donors have evolved considerably (Schmidt et al., 2020b). It therefore continues to be essential to invest in the quality of prospective donors' genetic information (high-resolution typing) in the database by retyping donors with low-resolution profiles (some time after the initial registration at $t = 0$) using up-to-date laboratory methods.

Second, whenever a prospective donor turns out to be a suitable match for a patient based on HLA types and other characteristics, the transplant center responsible for the patient's treatment requests that the donor undergoes a CT ($t = 1$). The objective is to confirm that a donor is suitable, still committed, and medically eligible for a stem cell collection. During the CT stage, prospective donors participate in a phone consultation and a questionnaire-based medical clearing. Fresh blood samples are requested from the donor to confirm the donor-recipient HLA match, and to test for infectious disease markers, including the CMV antibody status. The CT stage is thus *the* important milestone in the process of becoming a stem cell donor. Finally, if a suitable donor for HSCT is identified during the CT stage, the "work-up" process, which includes the scheduling of the donation date, a medical examination at

[8] For example, DKMS' figures from 2018 indicate that 139,146 potential stem cell donors were recruited at 427 public drives, of which 109 drives were patient-centered. The average recruitment number at patient-centered drives was 948, while at drives without focus on a specific patient the average number was 109. 115,388 potential donors were recruited at 1,117 special drives with an average recruitment number of 103. However, more than half of actual donor recruits come from online registration.

[9] The six most relevant genes are HLA-A, HLA-B, HLA-C, HLA-DRB1, HLA-DQB1, and HLA-DPB1 (Dehn et al., 2019).

[10] The donor's CMV antibody status is an important transplantation-relevant parameter. For transplant patients with a weakened immune system, a CMV infection can have life-threatening consequences.

the collection center, and the organization of travel and accommodation, begins. Subsequently, the actual donation process may be initiated ($t = 2$).

The process from registration to actual donation carries several uncertainties for the stem cell donor registries that might hinder the efficient search for a prospective donor. First, it is uncertain whether a prospective donor, who is potentially a good genetic match, is (still) readily available in case of a CT request. Typical reasons for unavailability include pregnancy and long-term stays abroad. Second, the quality of the initial typing might be insufficiently precise, leaving it uncertain whether the initially determined HLA type is confirmed by high-resolution typing methods. This problem occurs predominantly with donors whose original HLA typing dates back many years. In donors typed in recent years, unconfirmed results are very rare (Baier et al., 2019). Finally, it is uncertain whether prospective donors are still committed to donate stem cells once they are asked for a CT, in particular, if there is a long time span between registration and CT stage. The initiatives implemented by DKMS and outlined in the next section are meant to address these issues.

5.2.2 The registry's initiatives

Over the years, DKMS Germany has launched several initiatives aimed at mitigating the above-mentioned problems of low-resolution typing, missing information about periods of donor unavailability, and lacking donor commitment at the CT stage. A group of potential donors is asked to participate in an initiative some time after signing up with the stem cell donor registry, but before the call for a CT.[11] Table 5.1 outlines the initiatives.

First, the initiative PROSPECTIVE, which was conducted between 2013 and 2018, requests retyping of potential donors usually via buccal swab or blood draw through a physician. The initiative also involves a short health questionnaire. Typically, donors with low resolution HLA profiles are asked to participate. Further criteria for potential donors to be invited to the initiative PROSPECTIVE are age, sex, HLA genotype frequency, BMI, and missing CMV antibody status. A request for retyping can be issued directly by the registry (the focus of our paper) or a transplant center (Schmidt et al., 2011). Updating the typing profile of a donor can ultimately accelerate the donor search process, as the search coordinator has more detailed information on the donor's suitability. Besides, it increases the chance that the donor will be requested in future donor searches. In fact, incompletely typed donors may remain unidentified in donor searches, even if they are full matches for the patient (Sauter et al., 2016). Potential donors invited to this initiative were reminded only once to send the sample in and were not excluded from the registry if they did not participate.

Second, since high-resolution typing upon recruitment became more commonplace, DKMS introduced an initiative in 2015, which we label STATUS UPDATE. This initiative emphasizes in the invitation letter that the potential donor has, based on certain biological parameters, a higher likelihood of actually being matched to a patient in need of a stem cell donation. The motivation for this initiative was to minimize delays at the future CT stage by asking donors to complete a short health questionnaire and to report any future unavailability that lasts longer than three weeks, for example, due to a longer stay abroad or a pregnancy. Based on reported periods of future unavailability of participating donors, the registry is able to block the donor for the corresponding time period and the donor cannot be requested by a search coordinator, which should increase CT availability. Furthermore, as the commitment to report future unavailability dates requires a particularly high donor motivation, the initiative was also intended to create a pool of high-availability donors.

Third, similarly to STATUS UPDATE, the initiative STATUS UPDATE-BLOOD aims to identify a subgroup of potential donors who have, by virtue of their common HLA genotype, age, sex, or other parameters a higher probability of actually being asked to donate at the CT stage. Again, the motivation was to minimize delays at the CT stage by

[11] DKMS selects these donors mainly based on exogenous, biological factors. These factors are predetermined prior to the invitations and the donors are highly unlikely to be aware of these in advance to the invitation.

bringing forward specific parts of the CT. Since a considerable number of donors were typed at an intermediate or a low resolution during their registration, this initiative included, in addition to a health questionnaire, high-resolution retyping through blood draw (which is different as compared to STATUS UPDATE). Another key aim of this initiative is again an increase in CT availability, as donors were also requested to inform DKMS in cases of temporal future unavailability. In contrast to STATUS UPDATE, the invitation letter to the initiative STATUS UPDATE-BLOOD also contains an explicit team framing, as DKMS asks invitees in the invitation letter to become part of a "team of quickly available donors". Invitees were also told that participants in the initiative are more genetically likely to be requested to donate stem cells, compared to non-participants. Between 2015 and 2017, the initiatives STATUS UPDATE and STATUS UPDATE-BLOOD were run simultaneously, and from 2018 onwards the STATUS UPDATE-BLOOD initiative was discontinued.

The invitation to each of the initiatives is done on a rolling basis, where blocks of prospective donors are created for each genotype frequency rank. For each rank, a pre-specified number of donors is selected who simultaneously meet additional selection criteria.

Table 5.1: Overview on the registry's initiatives

	Initiatives		
	Prospective	**Status update**	**Status update−Blood**
A. Initiative properties			
Years run (within our sample period)	2013-2018	2015 - 2018	2013 - 2017
Motivation for initiatives	Typing profile incomplete or low resolution	Increase in CT availability for the group of participating donors, in particular by allowing them to report periods of unavailability for CT request; bringing forward parts of the CT process	Increase in CT availability for the group of participating donors, in particular by allowing them to report periods of unavailability for CT request; bringing forward parts of the CT process
Selection criteria	Selection criteria include age, low or intermediate resolution (i.e., incomplete) HLA typing profile, BMI, no medical bans, sex, in DKMS database for a specific time span, e-mail address, missing information on CMV antibody status, genotype frequency rank.	Selection criteria include age, genotype frequency rank, sex (in the beginning of the initiative), maximum number of participants per rank, high resolution HLA typing profile, e-mail address, BMI, no medical bans. Blood type, rhesus factor, and cytomegalovirus immunity status are known.	Selection criteria are analogous to STATUS UPDATE. But blood type, rhesus factor, and cytomegalovirus immunity status are unknown.
Invitation letter	Focus on technological advancement in retyping (e.g., CMV antibody status, additional locus)	They are asked since they have higher chance of being matched	They are asked since they higher chance of match; letter includes team framing
Our sample: potential donors with CT request	5,193	2,144	10,905
B. Participation requirement in initiative			
Provide unavailability dates	No	Yes	Yes
Health questions	Yes	Yes	Yes
Re-typing required	Yes	No	Yes
Sample	New/Frozen	No	New

This table provides an overview of the DKMS initiatives that we study in this paper. Notice that this is a subset of a broader set of initiatives run by DKMS. In this study, we only consider potential donors with a CT request.

5.3 Methods

5.3.1 Data and variables

The initial DKMS data set contains the universe of 104,116 observations from *first* CT requests, and associated donor information, from 1 November 2013 to 31 October 2018. From the initial data set, we excluded observations with CT cancellations from the patient side prior to the registered donor making a decision on the CT request (5,931 observations), since it cannot be determined whether these donors would have followed through with the CT request. We also excluded observations, where the latest retyping request before the CT request originated from a transplant center in the context of a search for a specific patient (7,198 observations). Some individuals had both a CT cancellation from the patient side and a patient-related retyping request, hence 12,637 observations were removed from the analysis. This leaves us finally with 91,479 observations for analysis.

The sample collected only contains first CT requests, since availability here is of primary interest to stem cell donor registries. If a potential donor does not take part in the first CT, the likelihood that they will be available for further CT requests is lower (in our data, it amounts to an average of 47%).

The outcome we are interested in is prospective donors' completion of the CT request. This is a binary variable called "availability": Prospective donors are regarded as available (coded 1) if they could be contacted and successfully completed the CT process, or as unavailable (coded 0) if they did not complete the CT process. More specifically, available donors declared they were willing to donate, provided the requested blood samples, filled out the health questionnaire, and were medically eligible. Reasons for unavailability are, for instance, no longer being interested in donating, medical ineligibility due to illness, or temporary unavailability (e.g. being overseas or having an important reason limiting time available to donate). Importantly, DKMS is able to track persons who have changed address and not informed DKMS about this through the local municipality the person was last registered at.[12] As a result, there are only 779 cases in our sample where potential donors could not be contacted.

In case donors received invitations to multiple initiatives, we consider the initiative closest to the CT request as the decisive initiative that can affect the CT availability for reasons of salience.[13] In these cases, we include a dummy variable (called "Multiple invitations") as a control. We ignore all CT requests where the most recent initiative was explicitly part of a donor search for a specific patient initiated by a transplant center, since non-compliers with this initiative are excluded from receiving a further CT request, which introduces a large sample-selection bias.[14]

Table 5.2 provides an overview of the rich set of covariates, which can be categorized into registry-related and donor-related characteristics. All summary statistics refer to non-missing covariate values from the regression sample.[15] The former category comprises the type of registration and the sample collection method. The type of registration includes the categories: public drive centered/not centered around a specific patient, company drive, special projects (such as donor drives at a school, university, sports event, and among police, fire fighters, and armed forces), and online registration. Sample collection was performed either by blood draw or by buccal swab. We also account for seasonality by controlling for the month of the CT request.

[12] In Germany, there is mandatory registration of address to the local municipality, so one can track where people have moved to within Germany.

[13] For example, if a donor is only invited to STATUS UPDATE, which is a dummy variable, they receive a 1. A person is also coded as 1 for STATUS UPDATE if he or she had earlier been invited to the PROSPECTIVE initiative, but later received a STATUS UPDATE invitation.

[14] Also, for participation variables, we have 34 cases from PROSPECTIVE, where the participation decision was unclear, and these are re-coded into the comparison group.

[15] We have 47 missing observations for the registration method and 3,431 missing observations of ancestry (since this was not mandatory information), and otherwise full covariate information for all CT requests. These missing values have been included in all analyses using an extra dummy for each variable (for ethnicity, a 0-1 dummy), or a separate category for the type of registration.

Table 5.2: Description of control variables included in our regression analyses

Variable	Description
A. Registry-related characteristics	
Registration method	Categorical variable indicating whether prospective donor had registered at public drive centered/not centered around a specific patient, company drive, special project drive (e.g., at schools, universities, sports events, and among police, fire fighters, and armed forces), or online via the DKMS website.
Mode of sample collection	Categorical variable measured at time of prospective donor's registration for collection through blood draw or buccal swab.
Month of CT request	Categorical variable for month of CT request.
B. Donor-related characteristics	
Sex	Dummy variable for female (zero for male).
Age	Categorical variable (eight categories) for donors' age at time of CT request: 17 to 25 years, 26 to 30 years, 31 to 35 years, 36 to 40 years, 41 to 45 years, 46 to 50 years, 51 to 55 years, and 56 to 61 years.
Ancestry	Dummy variable for (self-reported) ancestry being either German or from other countries.
State of residence	Categorical variable for prospective donors' state of residence during registration: Baden-Württemberg, Bavaria, Berlin, Brandenburg, Bremen, Hamburg, Hesse, Lower Saxony, Mecklenburg-Western Pomerania, North Rhine-Westphalia, Saarland, Saxony, Saxony-Anhalt, Schleswig-Holstein, and Thuringia.
Population size at residence	Categorical variable for prospective donors' state of residence: $<50{,}000$, 50,000 to 99,999, 100,000 to 199,000, 200,000 to 499,999, and $\geq 500{,}000$.
BMI	Continuous variable comprising prospective donors' body mass index (BMI).
Date of registration	Categorical variable for time when registration took place: before 2007, from 2007 to 2010, 2011 to 2014, and 2015 to 2018.
Information letter	Dummy indicating whether prospective donors received an information letter that their frozen sample was used for retyping.
Previous initiative	Dummy indicating that prospective donor had been previously invited to another initiative.

Donor-related variables include prospective donors' sex, age at CT request, self-reported ancestry,[16] the federal state of residence, the population size at their place of residence, prospective donors' body mass index (BMI), and the date of registration (which we sort into the following categories: registration before 2007, from 2007 to 2010, from 2011 to 2014, and from 2015 to 2018). Most of the potential donors' characteristics are mandatory information, which is collected at the recruitment stage, such as sex, address, weight and height, and the date of birth.[17]

5.3.2 Descriptive statistics

Table 5.3 provides an overview on donors' availability by initiative they had been invited to (Panel A) and conditional on whether prospective donors participate in an initiative (Panel B). Out of 91,479 potential donors with CT requests in our sample, 5,193 were invited to the initiative Prospective closest to the CT request, while 10,905 potential donors were invited to Status update-Blood, and 2,144 to Status update closest to the CT request. That implies that around 80% of all donors who received a CT request were not invited to any of the initiatives (73,237 out of 91,479).

We observe the highest CT availability (82.4%) among potential donors who had been invited to the initiative Status update-Blood. For initiatives Status update and Prospective, CT availability is slightly lower with 80.7% and 79.5%, respectively. Among the donors who did not receive an invitation, 77.1% were available for a CT.

As participation in any of these initiatives was voluntary, it is instructive to analyze differences in participation levels between the initiatives. The participation rate for Prospective is 58.6% (3,021 participants, 2,138 non-participants). For the initiatives Status update and Status update-Blood, we observe participation rates of 31.3% and 34.7%, respectively.

We are also able to compare the participation rates in the initiatives within our sample among potential donors without a CT request. Analyzing data for potential donors from the registry unconditional on having a CT request, we observe 48% participation rate in the Prospective initiative. For Status update-Blood, the participation rate was 24.4%, and for Status update is was 25%. Since participation rates in the initiatives are higher in the CT request sample, the initiatives themselves may lead to a positive selection of donors that are invited to CT. To address this issue, we later correct our estimated intention-to-treat effects by taking participation rates unconditional on CT requests into account.

Table 5.3 shows that CT availability varies with the prospective donors' participation in the initiatives. As a general pattern across all initiatives, we observe that participants of the initiatives show a considerably higher availability for CT that non-participants. For the initiative Prospective, the difference in availability is greatest with 15.2pp. The availability difference is 10.9pp and 12.2pp between participants and non-participants for the initiatives Status update and Status update-Blood, respectively. Put differently, considering the *un*availability of potential donors, participation in an initiative yields substantial decreases in potential donors' CT unavailability. For example, for the initiative Prospective the unavailability is more than 50% smaller for participants in the initiative (14.2% not available) compared to non-participants (29.4% not available). These descriptive results already indicate the potential sorting effect of the initiatives, and also some differences between the initiatives with respect to their implied burden to the participants (i.e., registered donors who agree to participate in a high-burden

[16] We categorize this variable in German and non-German, as there is a large number of minority backgrounds in the sample. Overall, there were 143 reported different nationalities, of which the ten most common were German (87.8%), Turkish (4.7%), Polish (0.9%), Russian (0.9%), Italian (0.7%), Greek (0.3%), Kazakh (0.3%), Romanian (0.2%), Austrian (0.2%), and French (0.1%).

[17] Potential donor BMI ranges from 13 to 40, and is 24 at median. Potential donors' age upon CT request ranges from 17 to 60 years with an average of 30 (median 28). Country of origin is self-reported (as additional information on a consent form) and was available for 94.8% of the studied donors.

initiative are signaling a relatively high level of commitment to go forward with the process should they be asked to donate).[18]

Table 5.3: CT (un)availability by initiative and participation

	(1) Obs.	(2) Average CT availability (%)
A. Initiatives		
PROSPECTIVE	5,193	79.5
STATUS UPDATE	2,144	80.7
STATUS UPDATE-BLOOD	10,905	82.4
No invitation to initiative (baseline)	73,237	77.1
B. Initiatives and participation		
PROSPECTIVE; Participation	3,021	85.8
PROSPECTIVE; No participation	2,138	70.6
STATUS UPDATE; Participation	671	88.2
STATUS UPDATE; No participation	1,473	77.3
STATUS UPDATE-BLOOD; Participation	3,783	90.4
STATUS UPDATE-BLOOD; No particip.	7,122	78.2

This table describes availability rates across initiatives for N=91,479 first CT requests from 1 Nov 2013 to 31 Oct 2018 at DKMS Germany. Column 1 shows the number of observations by initiative invitation category, column (2) the average CT availability. For 34 invitations to the PROSPECTIVE initiative, it is unclear whether the prospective donor participated.

5.3.3 Empirical strategy

5.3.3.1 Intention-to-treat regression specification

We apply a logistic regression approach to investigate the impact of the different initiatives on the potential donors' availability at CT. The unit of observation is at the CT request level. Our goal is to measure the overall effectiveness of these initiatives to increase CT availability, unconditional on actual participation of donors in these initiatives, and controlling for registry-related and donor-related characteristics.

The dependent variable, AVAILABLE_i, measures the availability for CT upon the first request, which equals 1 if a donor is available and 0, otherwise. Hence, we estimate the following regression specification:

$$P(\text{AVAILABLE}_i|\mathbf{x}) = \Lambda(\beta_0 + \beta_1 \text{PROSPECTIVE}_i + \beta_2 \text{STATUS UPDATE-BLOOD}_i + \beta_3 \text{STATUS UPDATE}_i + \delta \mathbf{X}_i), \quad (5.1)$$

where $\Lambda(x) = exp(x)/(1+exp(x))$. Here, the independent variables are indicators of the invitation to the different initiatives. For each category, we include one dummy, apart from when the donor is not invited to any of the

[18] Table 5.7 also reports the number of observations and univariate differences in availability rates across registry-related and donor-related characteristics.

initiatives, which forms the reference group. In each regression specification, we include a wide range of donor- and registry-related control variables as described in Table 5.2. We report average marginal effects for all independent variables of interest, and report robust standard errors in all specifications.

The identifying assumptions of a causal effect of an invitation into a treatment (i.e., an initiative) would imply that the intention to treat is as good as randomly assigned, conditional on selection criteria **X**, that are applied to select participants for initiatives. Potentially, characteristics of a prospective donor can be correlated with the outcome variable of interest, availability, and the initiatives, causing an omitted variables bias problem. But in our setting, the selection into initiatives is primarily based on *exogenous*, biological factors that are predetermined prior to the treatment and the donor is unlikely to be aware of these in advance, for example, upon registration or before receiving an invitation letter to an initiative. Hence, potential donors are unable to behaviorally influence the probability to be invited to an initiative, once they are enrolled in the registry, which should mitigate endogeneity concerns.

5.3.3.2 Initiative participation specification

We also observe whether registered donors who have been asked to participate in any of these initiatives accept the invitation and participate, or whether they decline, or do not respond to the invitation.[19] This motivates a second analysis, where we estimate differences in complying to CT requests for participants ("yes") and non-participants ("no") in our initiatives. We estimate the following regression specification:

$$\begin{aligned} \text{P}(\text{AVAILABLE}_i|\mathbf{x}) = \Lambda(&\beta_0 + \beta_4\text{PROSPECTIVE\_YES}_i + \beta_5\text{PROSPECTIVE\_NO}_i + \beta_6\text{STATUS UPDATE-BLOOD\_YES}_i \\ &+ \beta_7\text{STATUS UPDATE-BLOOD\_NO}_i + \beta_8\text{STATUS UPDATE\_YES}_i + \beta_9\text{STATUS UPDATE\_NO}_i + \delta\mathbf{X}_i). \end{aligned} \tag{5.2}$$

The analysis intentionally includes a selection effect, as a potential donor chooses whether or not to participate. This is particularly important from the perspective of a manager running a stem cell donor registry, since one aim of these initiatives is to identify pools of readily available donors with high commitment, and also to collect additional medical information (e.g., high-resolution HLA typing profile, CMV status, or blood group). From a practical perspective, the analysis may indicate whether these initiatives can effectively sort for more available donors on factors that are otherwise unobserved, as well as collecting more information about these donors that is match relevant. If donors who participate are not only better typed by the registry, but also show higher motivation to follow through, this can, ceteris paribus, provide grounds for prioritization of a participating donor over one that did not participate in these initiatives, given this information is available to the search coordinator.

5.3.3.3 Local average treatment effects

In the next specification, we assess the impact of donor invitations on the subgroup that actually participate, net of self-selection into the initiatives, by estimating a local average treatment effect (LATE) (Imbens and Angrist, 1994; Abadie and Cattaneo, 2018). Here, the endogenous variable of interest takes the value 1 if a person participated in the relevant initiative, and 0 otherwise, and exists for each of the three initiatives where there is an endogenous participation decision. The endogenous variables are each instrumented by a dummy for being invited to the initiative. The second stage is

$$\text{AVAILABLE}_i = \beta_0 + \beta_1\text{PROSPECTIVE YES}_i + \beta_2\text{STATUS UPDATE-BLOOD YES}_i + \beta_3\text{STATUS UPDATE YES}_i + \delta\mathbf{X}_i + u_i, \tag{5.3}$$

[19] One mechanical reason for the non-responses is that individuals cannot be contacted by DKMS (e.g., after someone has moved away). This can be considered very unlikely, as DKMS Germany has access to resident data through the local residents' registration offices.

and the first stage is a system of three equations, one for each initiative:

$$\begin{aligned} \textsc{Prospective yes}_i &= \gamma_0 + \gamma_1 \textsc{Prospective}_i + \gamma \mathbf{X}_i + r_i, \\ \textsc{Status update yes}_i &= \rho_0 + \rho_1 \textsc{Status update}_i + \rho \mathbf{X}_i + s_i, \\ \textsc{Status update-blood yes}_i &= \tau_0 + \tau_1 \textsc{Status update-blood}_i + \tau \mathbf{X}_i + t_i, \end{aligned} \tag{5.4}$$

which we estimate with robust standard errors. This analysis assumes that conditional on observable characteristics, the initiative assignment is exogenous. We control for all observable criteria for being selected into initiatives. This also assumes that the instrument (i.e., having received an invitation) can only affect CT availability through the endogenous variable (i.e., participating in an initiative). The last assumption likely holds, since participation is otherwise impossible. Also, there is one-sided non-compliance, which means the comparison group never participates, but the group invited to participate does not fully participate.

5.4 Results

Table 5.4 reports average marginal effects from intention-to-treat (logit) regressions using donor availability at first CT request as the dependent variable. The first two columns show intention-to-treat effects excluding (Model 1) and including (Model 2) our main set of controls.[20] These regression results are unconditional on potential donors' participation status in DKMS' initiatives. Focusing on the regression including our base controls in Model (2), the PROSPECTIVE initiative, also statistically significant, shows a 2.52pp greater predicted probability of being available at CT for donors invited to this initiative, compared to receiving no invitation (significant at the 1% level).

Focusing next on the status update initiatives, being requested to participate in the initiative STATUS UPDATE-BLOOD is associated with an average increase in predicted CT availability of 3.20pp, compared to receiving no request (significant at the 1% level). Considering that participation in the initiative is 27.8% unconditional on getting a CT request, this result is quite large. We also find that the coefficients on STATUS UPDATE-BLOOD and PROSPECTIVE are not statistically different from each other ($p = 0.3768$, Wald test). Being invited to the STATUS UPDATE initiative is associated with a 1.33pp increase, but this is not statistically significant. STATUS UPDATE-BLOOD has a 1.87pp larger impact on availability than STATUS UPDATE ($p = 0.083$, Wald test). STATUS UPDATE is not different from the initiative PROSPECTIVE ($p = 0.2994$, Wald test).

Model (3) of Table 5.4 shows the predicted CT availability of potential donors by donor participation status in DKMS' initiatives. Participation is defined as the active, voluntary choice by a donor to take part in an initiative, after having received an invitation letter, under the condition that she is medically eligible. Conditional on donors' participation, we observe that if donors actively decided to participate in an initiative, their availability in a future CT request is significantly higher than for donors who do not receive any invitation (base group). Second, if donors did not agree to participate in an initiative, their future availability is significantly lower than that of the base group. This applies to all initiatives we study, and all coefficient tests between participants and non-participants within an initiative are significant at the 1% level.

In detail, donors participating in STATUS UPDATE-BLOOD show a 13.9pp higher predicted CT availability (significant at the 1% level) and STATUS UPDATE participants show a 10.7pp higher availability rate, compared to the average donor that does not receive any invitation. Furthermore, donors participating in the initiative PROSPECTIVE show a 10.1pp higher CT availability. Testing the equality of status update initiatives' coefficients, we find that the coefficient belonging to STATUS UPDATE-BLOOD, YES is not significantly larger than STATUS UPDATE, YES (Wald

[20] Table 5.8 in the Appendix shows all control variables.

test: $p = 0.1536$). The coefficient of PROSPECTIVE, YES is about 3.8pp smaller than STATUS UPDATE-BLOOD, YES ($p = 0.0028$), but not different to STATUS UPDATE, YES ($p = 0.7727$).

Table 5.4: Effects of individual initiatives on potential donors' CT availability

Dependent variable:	CT availability			
Method:	Logit	Logit	Logit	2SLS
Model:	(1)	(2)	(3)	(4)
PROSPECTIVE	0.0236***	0.0252***		
	(0.0057)	(0.0063)		
STATUS UPDATE	0.0354***	0.0133		
	(0.0084)	(0.0097)		
STATUS UPDATE-BLOOD	0.0530***	0.0320***		
	(0.0039)	(0.0050)		
PROSPECTIVE, NO			-0.0524***	
			(0.0083)	
PROSPECTIVE, YES			0.1005***	0.0427***
			(0.0090)	(0.0105)
STATUS UPDATE, NO			-0.0217**	
			(0.0109)	
STATUS UPDATE, YES			0.1069***	0.0383
			(0.0204)	(0.0286)
STATUS UPDATE-BLOOD, NO			-0.0100*	
			(0.0055)	
STATUS UPDATE-BLOOD, YES			0.1389***	0.0817***
			(0.0094)	(0.0123)
Controls	No	Yes	Yes	Yes
Observations	91,479	91,479	91,479	91,479
(Pseudo) R^2	0.0019	0.0379	0.0425	0.0445
Wald tests (p-values):				
Prospective = Status update		0.2994		
Prospective = Status update-blood		0.3768		
Status update = Status update-blood		0.0830		
Prospective, yes = Prospective, no			0.0000	
Status update-blood, yes = Status update-blood, no			0.0000	
Status update, yes = Status update, no			0.0000	
Prospective, yes = Status update, yes			0.7727	0.8843
Prospective, yes = Status update-blood, yes			0.0028	0.0120
Status update, yes = Status update-blood, yes			0.1536	0.1594
Prospective, no = Status update, no			0.0235	
Prospective, no = Status update-blood, no			0.0000	
Status update, no = Status update-blood, no			0.3298	

This table shows average marginal effects from an intention-to-treat logistic regression (models 1 and 2), logistic regressions by participation status (model 3) and LATE using 2SLS (model 4), using a potential donor's availability for first CT request as the dependent variable. The control variables comprise of: registration method, mode of sample collection, sex, ancestry, population size of municipality of residence, body mass index (squared), age categories, year of registration, federal state of residence, information letter dummy, previous invitation to initiative dummy, and month of request. We include dummies (or separate categories) for missing continuous (categorical) covariates. Base category in all specifications: no invitation. Robust standard errors are reported in parentheses. The symbols *, **, and *** represent significance levels of 10%, 5%, and 1%, respectively. Constant not shown.

Turning to non-participants, we find a sizable and statistically significant negative correlation with CT availability for all initiatives. The largest negative effect size can be found for non-participation in the PROSPECTIVE initiative with a 5.2pp lower average availability. For status update initiatives, predicted availability of non-participating donors are between 1.0pp (STATUS UPDATE BLOOD, NO) and 2.2pp (STATUS UPDATE, NO) lower than the reference group. The difference between PROSPECTIVE, NO and STATUS UPDATE-BLOOD, NO is significant ($p = 0.0000$), as is the difference between PROSPECTIVE, NO and STATUS UPDATE, NO ($p = 0.0235$), but not significant between STATUS UPDATE-BLOOD, NO and STATUS UPDATE, NO ($p = 0.3298$). Overall, these results show that participants in the registry's initiatives are more likely to be available than non-participants.

Model (4) of Table 5.4 shows the estimated local average treatment effect for each initiative. STATUS UPDATE-BLOOD has the highest impact on CT availability with an 8.2pp effect size, followed by PROSPECTIVE with an effect size of 4.3pp. However, the coefficient for STATUS UPDATE is not significantly different from zero, which might be partly explained by the lower number of observations for this initiative. These results are largely in line with the intention-to-treat estimates. However, they show a larger impact of initiatives.

These results show that the initiatives do not just sort donors on availability, but actually have an impact on their participation decision. The magnitude of the LATE associated with STATUS UPDATE-BLOOD is about 59% the size of the associated coefficient in the *by participation* specification, suggesting that only 41% of the coefficient from the *by participation* specification is due to a self-selection of potential donors into the initiative, and the rest is an impact on donor availability through the initiative. Similar computations can be made for the PROSPECTIVE initiative, where 43% of the "yes" coefficient in the *by participation* specification is due to a positive impact of the initiative on the participants, and the rest is due to sorting.

5.5 Robustness checks

We now go over various robustness checks. First, we replicate the main intention-to-treat results using a regression adjustment model. Below, we test the proportion of unobserved selection to observed selection that would generate a null result, using the methodology of Oster (2019). We also calculate a bias-correction of point estimates using a re-weighting technique.

5.5.1 Average treatment effects on the treated

In a first test, we use a regression adjustment approach with a logistic outcome model to estimate average treatment effects on the treated, where "treated" denotes all individuals invited to an initiative (not just participants). This asks the question, how the CT availability of those persons in the registry who were invited to an initiative would have changed, had they not been invited.

This approach has multiple advantages. First, we only need to assume that the covariates for the individuals invited to the initiative have a positive probability of not being invited, and not a positive probability of invitation for the covariate values for all of those who were not invited, except for those with similar individual characteristics. For example, it would not make sense to only look at the average treatment effect, if there were many individuals in the group not invited that had covariates making them very unlikely to be invited (e.g. age over 45, or too high BMI) (Heckman, 1997). The ATET regression adjustment estimator thus compares potential outcomes for the group of treated individuals, conditional on their covariates, weighted by the probability of being invited (Abadie and Cattaneo, 2018), i.e.

$$\tau_{ATET} = E\left[E[Y|X=x, W=1] - E[Y|X=x, W=0]\right] Pr(X=x|W=1),$$

where Y is the estimated probability of being available for CT, X are covariates, and W is an indicator for being invited to an initiative of interest. We model the probability of being available for both treated and untreated with a logistic model using all controls. Since only a sub-population of potential invitees actually receives an invitation, we can be relatively sure that the reduced overlap assumption is fulfilled. We have a large number of individuals with heterogeneous BMI and age categories in the group not invited to any initiative.

The results of average marginal effects from logistic ATET estimates are shown in Table 5.5. We use robust standard errors in all estimates. Column 1 of Table 5.5 shows the ATET of PROSPECTIVE, where we find a 2.77 pp increase on CT availability for the invited. The STATUS UPDATE-BLOOD initiative has a slightly larger effect size of 3.22 pp. Again, in column 3, the STATUS UPDATE shows a 1.29 pp effect size, which is not significantly different from zero.

Regarding the plausibility of results, note that the potential outcome mean of availability for not being treated is about 2.5pp higher for the status update initiatives than for retyping initiatives (0.77 for PROSPECTIVE in column 1, 0.79 for STATUS UPDATE-BLOOD and STATUS UPDATE in Table 5.5. This is likely because those invited to status update initiatives are younger and more likely to be male, who have a higher CT availability. Also, overall, the ATET results show consistent estimates, suggesting that no individual model is significantly biased, and that the logistic regression models can be interpreted similarly to the ATET estimates.

Table 5.5: Regression adjustment: ATET

Dep. variable:	CT availability		
Model:	(1)	(2)	(3)
Prospective	0.0277***		
	(0.0061)		
Status update blood		0.0322***	
		(0.0041)	
Status update			0.0129
			(0.0091)
Potential outcome mean	0.7674***	0.7922***	0.7938***
(no invitation)	(0.0030)	(0.0021)	(0.0035)
Observations	78,430	84,142	75,381

This table shows average treatment effects on the treated (ATET), estimated using regression adjustment with a logistic outcome model and robust standard errors. The dependent variable is a potential donor's availability for first CT request. The control variables comprise of: registration method, mode of sample collection, gender, ancestry, population size of municipality of residence, body mass index (squared), age categories, year of registration, federal state of residence, multiple requests, information letter, and month of request. Robust standard errors in parentheses. The symbols *, **, and *** represent significance levels of 10%, 5%, and 1%, respectively.

5.5.2 Unobserved selection and coefficient stability

Next we use the methodology of Oster (2019) to assess the potential extent of unobserved confounding factors that could lead to a null result. Although we control for all selection criteria used by DKMS, the controls may be proxies of variables that do not perfectly measure the underlying variable. To do this, we use the approximate formula for the bias-corrected coefficient β^*, assuming proportionate selection of observables and unobservables according to δ (Oster, 2019, p. 193)[21]

$$\beta^* \approx \tilde{\beta} - \delta \left[\mathring{\beta} - \tilde{\beta}\right] \frac{R_{max} - \tilde{R}}{\tilde{R} - \mathring{R}} \tag{5.5}$$

where $\mathring{\beta}$ and $\mathring{R}$ are the coefficient estimate and R^2 from a regression of the dependent on the explanatory variable of interest without controls, $\tilde{\beta}$ and $\tilde{R}$ with controls, and $R_{max} = \Pi\tilde{R}$ is the maximum R^2 from a hypothetical regression with observed controls and unobserved factors.

We do two assessments of coefficient stability using this formula. First, how large would δ, the ratio of selection on unobservables to observables have to be, for the true β to move to zero. This is done assuming a value of $\Pi = 1.3$, as suggested by Oster, since about 90% of findings from randomized data would survive this. In studies with non-random data, less than 40% of studies assessed in Oster (2019) survive this threshold. Second, we ask how much larger the explanatory power of a regression including the unobserved elements could become for the coefficient to move to zero, i.e. we estimate the Π s.t. $\beta = 0$ assuming $\delta = 1$.

We use results from Table 5.4 and untabulated results from LATE estimations without controls, to calculate the results. Results are shown in Table 5.6. For PROSPECTIVE, the coefficient becomes larger when taking controls into account, which means that the δ is negative. This means that if observables are positively correlated with the initiative, the unobservables would have to be very strongly negatively correlated with the initiative for $\beta = 0$. Here it seems that the results are very robust. The STATUS UPDATE-BLOOD initiative shows at least a three (LATE) to four times (ITT) larger amount of selection on unobservables, relative to observables, that would be needed to drive to coefficients to zero, and $\Pi > 2$, indicating the R^2 would need to be at least twice as large to drive the coefficients to zero. This strongly suggests that the positive impact of status update initiative with retyping is not entirely the result of unobserved selection. This exercise, all in all, validates the findings from ITT and LATE specifications. We do not assess the *by participation* specification, as this is intentionally analyzing self-selection on unobserved factors.

Table 5.6: Selection on unobservable parameters

Acitivity		ITT	LATE
PROSPECTIVE	δ	-49.87	-104.48
	Π	-13.96	-30.34
STATUS UPDATE-BLOOD	δ	4.82	3.32
	Π	2.44	2.00

This table shows the results from tests for selection on unobservables based on Oster (2019). We report $\delta(\beta^* = 0, R_{max} = 1.3\tilde{R})$ and $\Pi(\beta^* = 0, \delta = 1)$ to test whether the result is robust to unobserved selection. We only include ITT and LATE estimates that were significant in Table 5.4.

[21] This is deemed by the author as an ex post valid test of robustness of results. Although we estimate an average marginal effects logit model, we see very similar results when testing the results of IV regressions, which are linear. This suggests that our test results are not driven by the non-linearity of the coefficient estimation.

5.5.3 Bias correction

Further, to account for the fact that there are relatively more participants to non-participants in initiatives who are invited to CT, we calculate a back of the envelope bias-adjustment. We do this by calculating the difference in the descriptive availability by initiative that would result from re-weighting the participants and non-participants, so that their relative amounts are the same as in the total registry. Then, we subtract this difference from the effect size. For example, for status update blood there are 24.4% participants in the entire registry, and in the sample of CT requests there are 34.7% participants. This suggests the point estimate is over-estimated. Re-weighting observations gives weights 0.7 to the participants and 1.15 to the non-participants. This would lead to a re-weighted descriptive average availability of 81.2%, which is 1.2pp lower. Correcting our point estimate of 3.2pp by this amount would lead to a point estimate of about 2.0pp, which would still reach statistical significance, given the estimated standard error is 0.50pp. Similarly, for the PROSPECTIVE initiative, we see descriptively a 1.1pp lower availability rate in aggregate when re-weighting the participants and non-participants uniformly. This would lead to an effect size of 1.4pp, which would still reach standard levels of statistical significance, given the estimated standard error is 0.63pp. Thus, even with a substantial reduction in effect size due to the potential endogenous sample selection, we still expect to see a significantly large impact of both initiatives.

5.6 Discussion and conclusions

We studied the impact of a set of initiatives implemented by DKMS Germany, a large stem cell donor registry, on availability to move forward with the donation process by registered donors who are a match for a patient in need of a transplant. Our results indicate that the initiatives were beneficial for two reasons. First, based on our ITT results, we found that the invitation to participate had a direct positive effect on completing the CT request (and hence, a relatively large relative reduction in attrition). Second, based on the predictive effect in which potential donors sorted themselves into participants and non-participants, the initiatives improved DKMS's ability to identify potential donors who were more likely to follow through with a CT request. Further, we also document a large causal impact on the availability of participants, for whom the reduction in attrition reduces by between 17% (3.8/22.9) and 36% (8.2/22.9).

Regarding the underlying mechanisms behind our results, we observe a much larger relative impact for the initiative STATUS UPDATE-BLOOD than for the initiative PROSPECTIVE on the participants in the LATE specification. One plausible reason for the higher availability among the STATUS UPDATE-BLOOD than the PROSPECTIVE initiative could be due to the higher participation burden (i.e., higher costs due to a blood draw) that could have served as a stronger commitment device for follow through with the donation. In the initiative STATUS UPDATE-BLOOD, donors must first give a blood draw, then participate in a health questionnaire, and are further asked to update the registry about their future unavailability periods, whereas in PROSPECTIVE, they only do either a blood draw or hand in a buccal swab.

The letters to the initiative PROSPECTIVE ask donors to send in buccal swabs or a blood sample to either improve the resolution of the factors relevant for a transplantation or to include further parameters (e.g., CMV status) that might be helpful in potential future donor searches. At the same time, the action-requiring PROSPECTIVE letters can be seen as costly by the registered donors, because the expected benefits from fulfilling the required task might be small from the donor's point of view, while the upfront costs of retyping might be perceived as substantially larger. The costs could be even larger if donors have strong time preferences.

There are a few channels through which the status update initiatives could have increased CT availability. First, the request for unavailability dates worked as intended. Specifically, to the extent that registered donors participating in these initiatives provided their unavailability dates, DKMS was able to successfully reduce the likelihood to ask these individuals when they were unavailable. Second, the additional requirement of reporting periods of absence, which might at a first glance be perceived as too much of a burden for potential donors, turned out to increase availability, on average, for the STATUS UPDATE-BLOOD initiative. This interpretation is supported by the fact that participation rates for status update initiatives are far lower than for retyping initiatives. Here, a foot-in-the-door mechanism might be effective, since STATUS UPDATE-BLOOD donors are required to provide blood samples, which costs some time and effort. However, this may not be the only reason for differences. For instance, the wording of letters substantially changed from STATUS UPDATE-BLOOD to STATUS UPDATE. Letters of the initiative STATUS UPDATE-BLOOD are emphasizing the recruitment of "quickly available team members". This could potentially nudge STATUS UPDATE-BLOOD invitees to participate, whereas in STATUS UPDATE, the letter is more neutrally framed. Thus, the overall effect of informing donors about their higher likelihood of being asked to donate, without nudging potential donors to participate is likely to be small. Overall, the results support the intuition that bringing forward part of the costs of the CT request can improve availability by acting as a commitment device.

One mechanism potentially driving the CT availability, due to the initiatives, is the time between invitation to an initiative and the CT request. Here, a few channels may play a role. First, the memory of participating or being invited may fade over time. Second, participants in STATUS UPDATE-BLOOD and STATUS UPDATE may initially report unavailability, thus leading to less CT requests for those unavailable initially, but over time, if participants forget to report unavailability in advance, they may be more likely to be unavailable when called to CT.

Another potential mechanism driving the effects is that, participants in PROSPECTIVE and STATUS UPDATE-BLOOD are more likely to be matched with a patient, since they have, per definition better HLA resolution and are thus more likely to be discovered in donor searches by doctors as a matching donor. Thus, we might see more donors that participated in retyping in the sample of CT requests. If participants are more motivated, as we have shown, then retyping can also positively behaviorally influence CT availability on top of providing better match information to search coordinators.

The results provide some practical implications for donor registries. We show that information on CT availability is contained in the participation choice of potential donors in an initiative. This information from participation decisions could be used to identify committed donors, since participants are more available than non-participants. Thus, given two biologically matched donors for a patient, all other things equal, one would prefer to always ask a potential donor participating in an initiative for a CT if faced with the constraint of only being able to invite one donor. This strategy can potentially lead to a faster transplantation and a reduced risk of the CT not being carried out for a matching donor, which would mean having to go back to other matches and start the work-up process over again. From a health policy perspective, these results have important implications for the information exchange between donor registries and transplant centers. In particular, to improve the efficiency of the donation process, donor registries should have the possibility to make such information available to search coordinators, which is currently not feasible.

This study has strengths and limitations. First, we were able to analyze a rich data set of 91,479 CT requests with initiatives that either retyped patients or asked them about their future unavailability. This is fairly unique given that the probability of getting a CT request is, on average, rather low, and thus a large number of invitations were sent out to registered donors beforehand. Also, the data is unique in that there is a potentially long time frame from registration, and the initiatives, to CT request. This is helpful, since it can potentially offer insights applicable to other areas of volunteering, where people are called to receive training in between initially signing up to be, for example, an emergency first aid helper, or a volunteer firefighter or rescuer, and the call to duty. This shows that similar settings that ask people to once again refresh their status as a member, or continually provide updates, can help increase the endline commitment toward the cause.

Another strength of the data is that we can control for the selection criteria employed by DKMS to invite donors into the initiatives. This enables us to get closer to a causal impact of the initiatives than would have been the case without knowledge of assignment. Also, since the selection criteria are predetermined and cannot easily be affected by the donors themselves, there is unlikely to be self-selection to being invited. We also have information on participation in the initiatives, which enables us to go beyond the ITT effect and test how large the initiative impact on participants is, relative to the self-selection effect. This is useful, as it shows whether initiatives sort donors, and whether there is an *impact* on donor availability for participants.

While we have been able to identify a causal link between invitation to participate in the initiatives and subsequent CT availability, we are unable to go further to identify the precise channel(s) through which the initiatives had their effects. For example, for the initiative STATUS UPDATE-BLOOD, we cannot exactly identify the channels, since we have no simultaneous random assignment of a similar letter. This is why a randomized controlled field experiment could help identify the behavioral channels, by appealing to different preferences or psychological constructs that can crowd in donors more likely to donate, and lead them to commit to donation. Here, we have some first evidence that the STATUS UPDATE-BLOOD initiative can serve as a commitment device, hence crowding-in potential donors with appeals to identity-related preferences may tend more to participate. Sorting could have been driven by an appeal indicated in the letter to the initiative to join a team. We are unable, however, to exactly identify this potential channel.

In sum, our paper shows that the retyping and status update initiatives can be meaningful to enhance donor availability at the CT stage and may help to sort potential donors. Further, the evaluation of retyping and status update schemes and potential donor characteristics may help to optimize the donor recruitment process and thereby increase the chances that donors requested will be available. This should help to reduce delays in the donation process and, ultimately, improve patient outcomes. Insights from our results shed light on how to optimally design the recruitment and status update initiatives of a registry for two-stage volunteer markets in order to maintain a high commitment rate, so that donors positively respond to donation requests. This will help to enable a fast and efficient donor search, which is essential for many patients in need of a transplant. Finally, our results highlight the importance of carefully designing initiatives and to leverage the potential of insights from the behavioral sciences.

5.7 Additional tables

Table 5.7: Descriptive statistics of categorical variables

A. Registry-related characteristics			
Variable name	**Variable value**	**Number of cases**	**Rate of CT availability (%)**
Recruitment method	Non-patient-centered public community drive	21,425	77.2
	Patient-centered public community drive	20,042	77.2
	Company drive	4,442	73.8
	Special projects	14,264	74.6
	Online registration	31,259	82.5
	Missing	47	80.6
Sample collection	Blood draw	34,123	76.3
	Buccal swab	48,038	79.6
	Unknown	9,318	75.9
Month of Request	January	7,546	78.0
	February	7,231	78.0
	March	7,851	78.8
	April	7,420	78.1
	May	7,457	77.9
	June	7,929	77.7
	July	8,042	77.7
	August	8,100	78.0
	September	7,501	78.4
	October	8,001	77.6
	November	7,235	77.7
	December	7,166	77.7
B. Donor-related characteristics			
Gender	Female	34,853	73.8
	Male	56,626	80.5
Age categories	17 to 25 (years old)	34,636	79.7
	26 to 30	20,330	77.7
	31 to 35	12,749	75.8
	36 to 40	8,542	77.9
	41 to 45	6,347	79.4
	46 to 50	5,238	76.8
	51 to 55	2,809	72.2
	56 to 61	828	61.5
Ancestry	German	77,801	80.2
	Non-German	10,947	62.5
	No response	3,431	75.8
	Baden-Württemberg	14,358	78.4
	Bavaria	15,118	79.3
	Berlin	3,009	76.2
	Brandenburg	1,424	78.7
	Bremen	766	75.7
	Hamburg	2,052	77.4
	Hesse	6,663	77.1

	Lower Saxony	21,028	77.1
Federal state of residence	Mecklenburg-Western Pomerania	1,180	76.0
	North Rhine-Westphalia	10,841	78.9
	Rhineland-Palatinate	4,708	78.2
	Saarland	1,014	77.0
	Saxony	2,964	77.0
	Saxony-Anhalt	960	77.8
	Schleswig-Holstein	3,880	78.1
	Thuringia	1,456	77.6
	Missing	58	72.4
	< 50,000	54,470	79.3
	50,000-99,999	8,777	77.5
Population of place of residence	100,000-199,000	6,048	76.1
	200,000-499,000	8,259	76.0
	> 500,000	13,925	77.5
	until end 2006	8,279	74.3
Year registered	2007-2010	16,355	74.6
	2011-2014	38,960	79.2
	2015-Nov. 2018	27,855	79.3
Information letter	No	83,097	78.3
	Yes	8,382	74.4
Previous request	No	89,966	77.9
	Yes	1,513	81.22

Table 5.8: Effects of individual initiatives on potential donors' CT availability reporting all covariates

Dependent variable:	CT availability			
Method:	Logit	Logit	Logit	2SLS
Model:	(1)	(2)	(3)	(4)
PROSPECTIVE	0.0236***	0.0252***		
	(0.0057)	(0.0063)		
STATUS UPDATE	0.0354***	0.0133		
	(0.0084)	(0.0097)		
STATUS UPDATE-BLOOD	0.0530***	0.0320***		
	(0.0039)	(0.0050)		
PROSPECTIVE, NO			-0.0524***	
			(0.0083)	
PROSPECTIVE, YES			0.1005***	0.0427***
			(0.0090)	(0.0105)
STATUS UPDATE, NO			-0.0217**	
			(0.0109)	
STATUS UPDATE, YES			0.1069***	0.0383
			(0.0204)	(0.0286)
STATUS UPDATE-BLOOD, NO			-0.0100*	
			(0.0055)	
STATUS UPDATE-BLOOD, YES			0.1389***	0.0817***
			(0.0094)	(0.0123)
Information letter		-0.0128**	-0.0074	-0.0121*
		(0.0064)	(0.0065)	(0.0071)
Multiple invitations		0.0253**	0.0336***	0.0342***
		(0.0126)	(0.0126)	(0.0114)
Recruitment method:				
Missing		0.0839	0.0891	0.0850
		(0.0619)	(0.0615)	(0.0552)
Special projects		-0.0171***	-0.0163***	-0.0174***
		(0.0054)	(0.0054)	(0.0059)
Company drive		-0.0071	-0.0071	-0.0091
		(0.0070)	(0.0069)	(0.0077)
Online		0.0648***	0.0621***	0.0609***
		(0.0053)	(0.0053)	(0.0055)
Patient-rel. public drive		0.0224***	0.0229***	0.0231***
		(0.0047)	(0.0047)	(0.0049)
Sex: Female dummy		-0.0541***	-0.0558***	-0.0561***
		(0.0029)	(0.0029)	(0.0030)
Federal state of residence:				
Baden-Württemberg		0.0853*	0.0873*	0.0911
		(0.0504)	(0.0501)	(0.0583)
Bavaria		0.0827	0.0849*	0.0881
		(0.0504)	(0.0501)	(0.0583)
Berlin		0.0614	0.0641	0.0668
		(0.0510)	(0.0508)	(0.0589)
Brandenburg		0.0602	0.0628	0.0668
		(0.0514)	(0.0512)	(0.0592)
Bremen		0.0664	0.0695	0.0712
		(0.0524)	(0.0522)	(0.0603)
Hamburg		0.0634	0.0654	0.0685
		(0.0513)	(0.0510)	(0.0591)
Hesse		0.0718	0.0740	0.0774
		(0.0505)	(0.0503)	(0.0584)
Mecklenburg-Western Pomerania		0.0393	0.0411	0.0444

	(0.0515)	(0.0513)	(0.0595)
North Rhine-Westphalia	0.0782	0.0796	0.0835
	(0.0504)	(0.0502)	(0.0583)
Lower Saxony	0.0694	0.0711	0.0747
	(0.0503)	(0.0501)	(0.0583)
Rhineland-Palatinate	0.0764	0.0792	0.0820
	(0.0506)	(0.0504)	(0.0585)
Saarland	0.0573	0.0624	0.0644
	(0.0518)	(0.0516)	(0.0597)
Saxony	0.0424	0.0447	0.0480
	(0.0508)	(0.0506)	(0.0588)
Saxony-Anhalt	0.0489	0.0508	0.0551
	(0.0519)	(0.0516)	(0.0597)
Schleswig-Holstein	0.0651	0.0669	0.0709
	(0.0507)	(0.0504)	(0.0586)
Thuringia	0.0520	0.0535	0.0576
	(0.0513)	(0.0511)	(0.0592)
Ancestry: German	0.1426***	0.1396***	0.1684***
	(0.0037)	(0.0037)	(0.0049)
Missing ancestry	0.1382***	0.1335***	0.1648***
	(0.0087)	(0.0087)	(0.0099)
Population of place of residence:			
50,000-99,999 inhabitants	-0.0046	-0.0040	-0.0042
	(0.0048)	(0.0047)	(0.0048)
100,000-199,000 inhabitants	-0.0150***	-0.0149***	-0.0152***
	(0.0054)	(0.0054)	(0.0057)
200,000-499,999 inhabitants	-0.0145***	-0.0144***	-0.0150***
	(0.0049)	(0.0049)	(0.0051)
Above 500,000 inhabitants	0.0032	0.0034	0.0036
	(0.0050)	(0.0050)	(0.0051)
Body mass index:			
BMI	0.0584***	0.0574***	0.0634***
	(0.0043)	(0.0043)	(0.0049)
BMI squared	-0.0010***	-0.0010***	-0.0011***
	(0.0001)	(0.0001)	(0.0001)
Age:			
Under 26 years old	0.0560***	0.0567***	0.0563***
	(0.0044)	(0.0044)	(0.0045)
26-30 years old	0.0223***	0.0232***	0.0233***
	(0.0045)	(0.0045)	(0.0047)
36-40 years old	0.0261***	0.0261***	0.0272***
	(0.0056)	(0.0056)	(0.0057)
41-45 years old	0.0455***	0.0456***	0.0464***
	(0.0063)	(0.0063)	(0.0063)
46-50 years years old	0.0208***	0.0225***	0.0226***
	(0.0066)	(0.0066)	(0.0070)
51-55 years years old	-0.0197**	-0.0169**	-0.0226**
	(0.0081)	(0.0080)	(0.0093)
56-61 years old	-0.0961***	-0.0918***	-0.1230***
	(0.0128)	(0.0127)	(0.0174)
Year registered			
Registered from 2007 to 2010	0.0001	0.0059	0.0023
	(0.0071)	(0.0071)	(0.0077)
Registered from 2011 to 2014	0.0380***	0.0449***	0.0418***
	(0.0077)	(0.0077)	(0.0084)
Registered from 2015 to 2018	0.0381***	0.0460***	0.0423***

		(0.0081)	(0.0081)	(0.0087)
Sample collection:				
Buccal swab		-0.0079	-0.0068	-0.0073
		(0.0050)	(0.0050)	(0.0053)
Typing method unclear		0.0249***	0.0231***	0.0240***
		(0.0063)	(0.0063)	(0.0066)
Month of CT request:				
CT request in February		-0.0003	-0.0012	-0.0006
		(0.0067)	(0.0067)	(0.0067)
CT request in March		0.0057	0.0056	0.0056
		(0.0066)	(0.0066)	(0.0065)
CT request in April		0.0002	0.0008	0.0006
		(0.0067)	(0.0067)	(0.0066)
CT request in May		-0.0033	-0.0026	-0.0033
		(0.0066)	(0.0066)	(0.0066)
CT request in June		-0.0045	-0.0035	-0.0042
		(0.0065)	(0.0065)	(0.0065)
CT request in July		-0.0045	-0.0041	-0.0047
		(0.0065)	(0.0065)	(0.0065)
CT request in August		-0.0042	-0.0042	-0.0044
		(0.0065)	(0.0065)	(0.0065)
CT request in September		0.0004	0.0010	0.0005
		(0.0067)	(0.0067)	(0.0066)
CT request in October		-0.0084	-0.0076	-0.0083
		(0.0065)	(0.0065)	(0.0065)
CT request in November		-0.0020	-0.0021	-0.0020
		(0.0067)	(0.0066)	(0.0067)
CT request in December		-0.0021	-0.0020	-0.0019
		(0.0067)	(0.0067)	(0.0067)
Constant				-0.3896***
				(0.0861)
Observations	91,479	91,479	91,479	91,479
(Pseudo) R^2	0.0019	0.0379	0.0425	0.0445
Wald tests (p-values):				
Prospective = Status update		0.2994		
Prospective = Status update-blood		0.3768		
Status update = Status update-blood		0.0830		
Prospective, yes = Prospective, no			0.0000	
Status update-blood, yes = Status update-blood, no			0.0000	
Status update, yes = Status update, no			0.0000	
Prospective, yes = Status update, yes			0.7727	0.8843
Prospective, yes = Status update-blood, yes			0.0028	0.0120
Status update, yes = Status update-blood, yes			0.1536	0.1594
Prospective, no = Status update, no			0.0235	
Prospective, no = Status update-blood, no			0.0000	
Status update, no = Status update-blood, no			0.3298	

This table shows average marginal effects from an intention-to-treat logistic regression (models 1 and 2), logistic regressions by participation status (model 3) and LATE using 2SLS (model 4), using a potential donor's availability for first CT request as the dependent variable. All controls reported. Base categories comprise: no invitation (initiatives), non-patient-centered public community drive (recruitment method), missing (federal state), non-German (ancestry), smaller than 50,000 inhabitants (population place of residence, 31-35 years (age), until end of 2006 (year of registration), blood draw (typing method), January (month of CT request). Robust standard errors are reported in parentheses. The symbols *, **, and *** represent significance levels of 10%, 5%, and 1%, respectively. Constant not shown.

Chapter 6
Conclusion

This dissertation has focused on how to implement desired actions within organizational structures. Chapter 2 focused on incentives for executives, and found that pay is heterogeneously correlated with long-term firm performance across the conditional distribution. Importantly, this chapter found a reversal in the performance—pay relationship. Yearly pay has varies more strongly with short-term performance at the bottom of the distribution, and total wealth varies less strongly with long-term firm performance at the top of the distribution. Chapters 3 and 4 explored sorting of workers on behavioral outcomes, and the drivers of helping and antisocial behavior across firms. I found that preferences, trust, and leadership drive helping, while personality, trust and leadership drives antisocial behavior. Chapter 5 explored the role of initiatives run by a major stem cell donor center, and how these affect donor availability, and help to predict motivated donors

Altogether, the dissertation thus points to many open questions, some of which are being assessed in further work. First, what are the preferences, attitudes, and motivations of stem cell donors? This is being assessed in a donor survey. Second, are the results of Chapter 5 causal, and what are the mechanisms? This is being assessed by a field experiment in cooperation with the donor center. Regarding Chapter 2, more detailed information about contracts used to compensate executives could be useful to dig deeper into the mechanisms at work driving the distributional differences in compensation. Regarding Chapters 3 and 4, it would be optimal to see a quasi-random allocation of workers to firms regarding their preferences. In this case, it would be unlikely to observe an effect if all workers were similar. Other research focusing on heterogeneous effects of management practices by worker preferences is more promising here.

References

Abadie, Alberto and Matias D Cattaneo, "Econometric Methods for Program Evaluation," *Annual Review of Economics*, 2018, *10*, 465–503.

Abowd, John M, Francis Kramarz, and David N Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, 1999, *67* (2), 251–333.

Alan, Sule, Gozde Corekcioglu, and Matthias Sutter, "Improving Workplace Climate in Large Corporations: A Clustered Randomized Intervention," IZA Discussion Papers 14726, Institute of Labor Economics (IZA) 2021.

Alchian, Armen A and Harold Demsetz, "Production, Information Costs, and Economic Organization," *American Economic Review*, 1972, *62* (5), 777–795.

Almlund, Mathilde, Angela Lee Duckworth, James Heckman, and Tim Kautz, "Personality Psychology and Economics," in "Handbook of the Economics of Education," Vol. 4, Elsevier, 2011, pp. 1–181.

Andreoni, James, "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving," *The Economic Journal*, 1990, *100* (401), 464–477.

Angrist, Joshua D and Jörn-Steffen Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, 2008.

Anthias, Chloe, Bronwen E Shaw, Jessica G Bruce, Dennis L Confer, Linda K Abress, Mary Amanda Dew, Annelies Billen, Ann O'Leary, Henny Braund, and Galen E Switzer, "Role of Race/Ethnicity in Donor Decisions about Unrelated Hematopoietic Progenitor Cell Donation: Exploring Reasons for Higher Attrition among Racial/Ethnic Minorities," *Biology of Blood and Marrow Transplantation*, 2020, *26* (3), 593–599.

Arrow, Kenneth J, *The Limits of Organization*, WW Norton & Company, 1974.

Aurelio, Teresa M, Alberto Aniasi, Simone E Haworth, Maria B Colombo, T Dimonopoli, MC Mocellin, Francesca Poli, Rosanna Torelli, Loretta Crespiatico, Marta Serafini et al., "Analysis of the Motivation for Hematopoietic Stem Cell Donation," *Transplantation Proceedings*, 2011, *43* (4), 981–984.

Baier, Daniel M, Jan A Hofmann, Heike Fischer, Gabi Rall, Julia Stolze, Katja Ruhner, Vinzenz Lange, Jürgen Sauter, and Alexander H Schmidt, "Very Low Error Rates of NGS-based HLA Typing at Stem Cell Donor Recruitment Question the Need for a Standard Confirmatory Typing Step Before Donor Work-up," *Bone Marrow Transplantation*, 2019, *54* (6), 928–930.

Bart, Thomas, Thomas Volken, Yvonne Fischer, and Behrouz Mansouri Taleghani, "Giving Blood and Rnrolling on the Stem Cell Donor Registry: Ranking of Obstacles and Motives in Switzerland," *Transfusion Medicine and Hemotherapy*, 2014, *41* (4), 264–272.

Bartling, Bjorn, Ernst Fehr, Michel André Maréchal, and Daniel Schunk, "Egalitarianism and Competitiveness," *American Economic Review*, 2009, *99* (2), 93–98.

Bebchuk, Lucian A and Jesse M Fried, *Pay Without Performance*, Vol. 29, Cambridge, MA: Harvard University Press, 2004.

— **and Lars A Stole**, "Do Short-Term Objectives Lead to Under or Overinvestment in Long-Term Projects?," *Journal of Finance*, 1993, *48* (2), 719–729.

— **, Yaniv Grinstein, and Urs Peyer**, "Lucky CEOs and Lucky Directors," *Journal of Finance*, 2010, *65* (6), 2363–2401.

Becker, Anke, Thomas Deckers, Thomas Dohmen, Armin Falk, and Fabian Kosse, "The Relationship Between Economic Preferences and Psychological Personality Measures," *Annual Review of Economics*, 2012, *4* (1), 453–478.

Beckmann, Michael and Matthias Kräkel, "Empowerment, Task Commitment, and Performance Pay," *Journal of Labor Economics*, 2022.

Beierlein, Constanze, Christoph J Kemper, Anastassiya Kovaleva, and Beatrice Rammstedt, "Short Scale for Measuring General Self-Efficacy Beliefs (ASKU)," *Methods, Data, Analyses*, 2013, *7* (2), 28.

Bellemare, Charles and Bruce Shearer, "Sorting, Incentives and Risk Preferences: Evidence From a Field Experiment," *Economics Letters*, 2010, *108* (3), 345–348.

Bellmann, Lisa, Benjamin Lochner, Stefan Seth, and Stefanie Wolter, "AKM Effects for German Labour Market Data," Technical Report, FDZ-Methodenreport, 01/2020, Institut für Arbeitsmarkt-und Berufsforschung (IAB), Nürnberg 2020.

Bellmann, Lutz, Stefan Bender, Mario Bossler, Sandra Broszeit, Christian Dickmann, Miriam Gensicke, Reiner Gilberg, Philipp Grunau, Patrick Kampkötter, Katharina Laske, Jens Mohrenweiser, Helmut Schröder, Holger Schütz, Dirk Sliwka, Susanne Steffes, Jens Stephani, Nikolai Tschersich, and Stefanie Wolter, "Linked Personnel Panel. Quality of work and economic success: Longitudinal study in German establishments," 2015.

Bender, Stefan, Nicholas Bloom, David Card, John Van Reenen, and Stefanie Wolter, "Management Practices, Workforce Selection, and Productivity," *Journal of Labor Economics*, 2018, *36* (S1), 371–409.

Berger, Johannes, Claus Herbertz, and Dirk Sliwka, "Incentives and Cooperation in Firms: Field Evidence," IZA Discussion Papers 5618, Institute of Labor Economics (IZA) 2011.

Bergstrom, Theodore C, Rodney J Garratt, and Damien Sheehan-Connor, "One Chance in a Million: Altruism and the Bone Marrow Registry," *American Economic Review*, 2009, *99* (4), 1309–34.

Bernile, Gennaro, Vineet Bhagwat, and P Raghavendra Rau, "What Doesn't Kill you Will Only Make You More Risk-Loving: Early-Life Disasters and CEO Behavior," *Journal of Finance*, 2017, *72* (1), 167–206.

Bertrand, Marianne and Sendhil Mullainathan, "Are CEOs Rewarded for Luck? The Ones Without Principals Are," *Quarterly Journal of Economics*, 2001, *116* (3), 901–932.

Bizjak, John M, James A Brickley, and Jeffrey L Coles, "Stock-Based Incentive Compensation and Investment Behavior," *Journal of Accounting and Economics*, 1993, *16* (1-3), 349–372.

__ , Michael L Lemmon, and Lalitha Naveen, "Does the Use of Peer Groups Contribute to Higher Pay and Less Efficient Compensation?," *Journal of Financial Economics*, 2008, *90* (2), 152–168.

Bloom, Nicholas and John Van Reenen, "Measuring and Explaining Management Practices Across Firms and Countries," *Quarterly Journal of Economics*, 2007, *122* (4), 1351–1408.

__ and __ , "Why do Management Practices Differ Across Firms and Countries?," *Journal of Economic Perspectives*, 2010, *24* (1), 203–24.

__ , Christos Genakos, Ralf Martin, and Raffaella Sadun, "Modern Management: Good for the Environment or Just Hot Air?," *The Economic Journal*, 2010, *120* (544), 551–572.

__ , Erik Brynjolfsson, Lucia Foster, Ron Jarmin, Megha Patnaik, Itay Saporta-Eksten, and John Van Reenen, "What Drives Differences in Management Practices?," *American Economic Review*, 2019, *109* (5), 1648–83.

__ , __ , __ , Ron S Jarmin, Itay Saporta Eksten, and John Van Reenen, "Management in America," *US Census Bureau Center for Economic Studies Paper No. CES-WP-13-01*, 2013.

Borghans, Lex, Angela Lee Duckworth, James J Heckman, and Bas Ter Weel, "The Economics and Psychology of Personality Traits," *Journal of Human Resources*, 2008, *43* (4), 972–1059.

__ , Bas Ter Weel, and Bruce A Weinberg, "People Skills and the Labor-Market Outcomes of Underrepresented Groups," *ILR Review*, 2014, *67* (2), 287–334.

Brekke, Kjell Arne, Karen Evelyn Hauge, Jo Thori Lind, and Karine Nyborg, "Playing With the Good Guys. A Public Good Game With Endogenous Group Formation," *Journal of Public Economics*, 2011, *95* (9-10), 1111–1118.

Burns, Arthur F and Wesley C Mitchell, *Measuring Business Cycles* 1947.

Burr, Hermann, Hanne Berthelsen, Salvador Moncada, Matthias Nübling, Emilie Dupret, Yucel Demiral, John Oudyk, Tage S Kristensen, Clara Llorens, Albert Navarro et al., "The Third Version of the Copenhagen Psychosocial Questionnaire," *Safety and Health at Work*, 2019, *10* (4), 482–503.

Cadman, Brian D, Tjomme O Rusticus, and Jayanthi Sunder, "Stock Option Grant Vesting Terms: Economic and Financial Reporting Determinants," *Review of Accounting Studies*, 2013, *18* (4), 1159–1190.

Campbell, T Colin and Mary Elizabeth Thompson, "Why Are CEOs Paid for Good Luck? An Empirical Comparison of Explanations for Pay-for-Luck Asymmetry," *Journal of Corporate Finance*, 2015, *35*, 247–264.

Canay, Ivan A, "A Simple Approach to Quantile Regression for Panel Data," *The Econometrics Journal*, 2011, *14* (3), 368–386.

Card, David, Jörg Heining, and Patrick Kline, "Workplace Heterogeneity and the Rise of West German Wage Inequality," *Quarterly Journal of Economics*, 2013, *128* (3), 967–1015.

Casta, Ana Hallgarten La, Bronwen E Shaw, Chloe Anthias, Jessica G Bruce, Gabrielle Pastorek, Annelies Billen, Ann O'Leary, and Galen E Switzer, "Motives for Joining an Unrelated HSC Donor Registry: Description, Categorization, and Association With Donor Availability," *Bone Marrow Transplantation*, 2019, *54* (3), 425–431.

Christiano, Lawrence J and Terry J Fitzgerald, "The Band Pass Filter," *International Economic Review*, 2003, *44* (2), 435–465.

Coase, Ronald Harry, "The Nature of the Firm," *Economica*, 1937, *4* (16), 386–405.

Cobb-Clark, Deborah A and Stefanie Schurer, "The stability of big-five personality traits," *Economics Letters*, 2012, *115* (1), 11–15.

Coffman, Lucas C, Clayton R Featherstone, and Judd B Kessler, "Can Social Information Affect What Job You Choose and Keep?," *American Economic Journal: Applied Economics*, 2017, *9* (1), 96–117.

Costa, Paul T and Robert R McCrae, "Domains and Facets: Hierarchical Personality Assessment Using the Revised NEO Personality Inventory," *Journal of Personality Assessment*, 1995, *64* (1), 21–50.

Craig, Ashley C, Ellen Garbarino, Stephanie A Heger, and Robert Slonim, "Waiting to Give: Stated and Revealed Preferences," *Management Science*, 2017, *63* (11), 3672–3690.

Daniel, Naveen D, Yuanzhi Li, and Lalitha Naveen, "Symmetry in Pay for Luck," *The Review of Financial Studies*, 2019.

Danilov, Anastasia and Dirk Sliwka, "Can Contracts Signal Social Norms? Experimental Evidence," *Management Science*, 2017, *63* (2), 459–476.

Dasgupta, Kushan, "Generosity and Compliance: Recruitment-Work and the Pathways to Participation in Bone Marrow Donation," *Social Science & Medicine*, 2018, *206*, 86–92.

Dehn, Jason, Stephen Spellman, Carolyn K Hurley, Bronwen E Shaw, Juliet N Barker, Linda J Burns, Dennis L Confer, Mary Eapen, Marcelo Fernandez-Vina, Robert Hartzman et al., "Selection of Unrelated Donors and Cord Blood Units for Hematopoietic Cell Transplantation: Guidelines From the NMDP/CIBMTR," *Blood*, 2019, *134* (12), 924–934.

Delfgaauw, Josse, Robert Dur, Oke Onemu, and Joeri Sol, "Team Incentives, Social Cohesion, and Performance: A Natural Field Experiment," *Management Science*, 2022, *68* (1), 230–256.

Deming, David J., "The Growing Importance of Social Skills in the Labor Market," *Quarterly Journal of Economics*, 06 2017, *132* (4), 1593–1640.

DeVaro, Jed, Jin-Hyuk Kim, and Nick Vikander, "Non-performance Pay and Relational Contracting: Evidence from CEO Compensation," *The Economic Journal*, 2017, *128* (613), 1923–1951.

Deversi, Marvin, Martin G Kocher, and Christiane Schwieren, "Cooperation in a Company: A Large-Scale Experiment," *CESifo Working Paper No. 8190*, 2020.

Dhaene, Geert and Koen Jochmans, "Split-Panel Jackknife Estimation of Fixed-Effect Models," *Review of Economic Studies*, 2015, *82* (3), 991–1030.

Dohmen, Thomas and Armin Falk, "Performance Pay and Multidimensional Sorting: Productivity, Preferences, and Gender," *American Economic Review*, 2011, *101* (2), 556–90.

—, —, David Huffman, and Uwe Sunde, "Representative Trust and Reciprocity: Prevalence and Determinants," *Economic Inquiry*, 2008, *46* (1), 84–90.

—, —, —, and —, "Homo reciprocans: Survey evidence on behavioural outcomes," *The Economic Journal*, 2009, *119* (536), 592–612.

—, —, —, and —, "Are Risk Aversion and Impatience Related to Cognitive Ability?," *American Economic Review*, 2010, *100* (3), 1238–60.

—, —, —, —, Jürgen Schupp, and Gert G Wagner, "Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences," *Journal of the European Economic Association*, 2011, *9* (3), 522–550.

Drago, Robert and Geoffrey K Turnbull, "Individual Versus Group Piece Rates Under Team Technologies," *Journal of the Japanese and International Economies*, 1988, *2* (1), 1–10.

— and —, "Competition and Cooperation in the Workplace," *Journal of Economic Behavior & Organization*, 1991, *15* (3), 347–364.

— and Gerald T Garvey, "Incentives for Helping on the Job: Theory and Evidence," *Journal of Labor Economics*, 1998, *16* (1), 1–25.

Dur, Robert and Joeri Sol, "Social Interaction, Co-Worker Altruism, and Incentives," *Games and Economic Behavior*, 2010, *69* (2), 293–301.

Eckel, Catherine C and Philip J Grossman, "Altruism in Anonymous Dictator Games," *Games and Economic Behavior*, 1996, *16* (2), 181–191.

Edmans, Alex and Xavier Gabaix, "Executive Compensation: A Modern Primer," *Journal of Economic Literature*, 2016, *54* (4), 1232–87.

—, Vivian W Fang, and Allen Huang, "The Long-Term Consequences of Short-Term Incentives," *European Corporate Governance Institute (ECGI)-Finance Working Paper*, 2019, (527).

—, —, and Katharina A Lewellen, "Equity Vesting and Investment," *The Review of Financial Studies*, 2017, *30* (7), 2229–2271.

—, Xavier Gabaix, and Dirk Jenter, "Executive Compensation: A Survey of Theory and Evidence," in "The Handbook of the Economics of Corporate Governance," Vol. 1, Elsevier, 2017, pp. 383–539.

—, —, Tomasz Sadzik, and Yuliy Sannikov, "Dynamic CEO Compensation," *Journal of Finance*, 2012, *67* (5), 1603–1647.

Eeckhout, Jan, "Sorting in the Labor Market," *Annual Review of Economics*, 2018, *10*, 1–29.

Efing, Matthias, Harald Hau, Patrick Kampkötter, and Jean-Charles Rochet, "Bank bonus pay as a risk sharing contract," *HEC Paris Research Paper No. FIN-2018-1285*, 2018.

Encinosa, William E, Martin Gaynor, and James B Rebitzer, "The Sociology of Groups and the Economics of Incentives: Theory and Evidence on Compensation Systems," *Journal of Economic Behavior & Organization*, 2007, *62* (2), 187–214.

Englmaier, Florian, Stefan Grimm, Dominik Grothe, David Schindler, and Simeon Schudy, "The Value of Leadership: Evidence from a Large-Scale Field Experiment," 2021.

Exley, Christine L and Ragan Petrie, "The Impact of a Surprise Donation Ask," *Journal of Public Economics*, 2018, *158*, 152–167.

Falk, Armin and Urs Fischbacher, "A Theory of Reciprocity," *Games and Economic Behavior*, 2006, *54* (2), 293–315.

—, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde, "Global Evidence on Economic Preferences," *Quarterly Journal of Economics*, 2018, *133* (4), 1645–1692.

—, —, —, David Huffman, and Uwe Sunde, "The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences," *IZA DP No. 9674.*, 2016.

Fehr, Ernst, "On the Economics and Biology of Trust," *Journal of the European Economic Association*, 2009, *7* (2-3), 235–266.

— and Urs Fischbacher, "The Nature of Human Altruism," *Nature*, 2003, *425* (6960), 785–791.

Fernandes, Nuno, Miguel A Ferreira, Pedro Matos, and Kevin J Murphy, "Are US CEOs Paid More? New International Evidence," *The Review of Financial Studies*, 2013, *26* (2), 323–367.

Fingrut, Warren, Kiran Rikhraj, and David Allan, "Targeted Recruitment of Male Donors for Allogeneic Haematopoietic Cell Transplantation: A Review of the Evidence," *Vox Sanguinis*, 2018, *113* (4), 307–316.

Fischer, Gabriele, Florian Janik, Dana Müller, and Alexandra Schmucker, "The IAB Establishment Panel-Things Users Should Know," *Journal of Contextual Economics*, 2009, *129* (1), 133.

Fitzenberger, Bernd, Gary Mena, Jan Nimczik, and Uwe Sunde, "Personality Traits Across the Life Cycle: Disentangling Age, Period, and Cohort Effects*," *The Economic Journal*, 11 2021.

FitzRoy, Felix R and Kornelius Kraft, "Profitability and Profit-Sharing," *The Journal of Industrial Economics*, 1986, pp. 113–130.

Friebel, Guido, Matthias Heinz, Miriam Krueger, and Nikolay Zubanov, "Team Incentives and Performance: Evidence From a Retail Chain," *American Economic Review*, 2017, *107* (8), 2168–2203.

Frydman, Carola and Dirk Jenter, "CEO Compensation," *Annual Review of Financial Economics*, 2010, *2* (1), 75–102.

Gabaix, Xavier and Augustin Landier, "Why has CEO pay increased so much?," *Quarterly Journal of Economics*, 2008, *123* (1), 49–100.

Gächter, Simon and Christian Thöni, "Social Learning and Voluntary Cooperation Among Like-Minded People," *Journal of the European Economic Association*, 2005, *3* (2-3), 303–314.

—, Daniele Nosenzo, Elke Renner, and Martin Sefton, "Who Makes a Good Leader? Cooperativeness, Optimism, and Leading-by-Example," *Economic Inquiry*, 2012, *50* (4), 953–967.

Gangadharan, Lata, Philip J. Grossman, and Joe Vecci, *Antisocial Behavior in the Workplace*, Cham: Springer International Publishing,

Garvey, Gerald T and Todd T Milbourn, "Asymmetric Benchmarking in Compensation: Executives are Rewarded for Good Luck but not Penalized for Bad," *Journal of Financial Economics*, 2006, *82* (1), 197–225.

Gerlitz, Jean-Yves and Jürgen Schupp, "Zur Erhebung der Big-Five-Basierten Persoenlichkeitsmerkmale im SOEP," *DIW Research Notes*, 2005, *4*, 2005.

Gibbons, Robert and John Roberts, *The Handbook of Organizational Economics*, Princeton University Press, 2013.

Gittell, Jody Hoffer, Andrew Von Nordenflycht, and Thomas A Kochan, "Mutual Gains or Zero Sum? Labor Relations and Firm Performance in the Airline Industry," *ILR Review*, 2004, *57* (2), 163–180.

Godfrey, Leslie G, "Testing for Multiplicative Heteroskedasticity," *Journal of Econometrics*, 1978, *8* (2), 227–236.

Goette, Lorenz and Alois Stutzer, "Blood Donations and Incentives: Evidence From a Field Experiment," *Journal of Economic Behavior & Organization*, 2020, *170*, 52–74.

Gopalan, Radhakrishnan, Todd Milbourn, Fenghua Song, and Anjan V Thakor, "Duration of Executive Compensation," *Journal of Finance*, 2014, *69* (6), 2777–2817.

Gragert, Loren, Mary Eapen, Eric Williams, John Freeman, Stephen Spellman, Robert Baitty, Robert Hartzman, J Douglas Rizzo, Mary Horowitz, Dennis Confer et al., "HLA Match Likelihoods for Hematopoietic Stem-Cell Grafts in the US Registry," *New England Journal of Medicine*, 2014, *371* (4), 339–348.

Graham, John R, Campbell R Harvey, and Shiva Rajgopal, "The Economic Implications of Corporate Financial Reporting," *Journal of Accounting and Economics*, 2005, *40* (1-3), 3–73.

Gratwohl, Alois, Marcelo C Pasquini, Mahmoud Aljurf, Yoshiko Atsuta, Helen Baldomero, Lydia Foeken, Michael Gratwohl, Luis Fernando Bouzas, Dennis Confer, Karl Frauendorfer et al., "One Million Haemopoietic Stem-Cell Transplants: A Retrospective Observational Study," *Lancet Haematology*, 2015, *2* (3), e91–e100.

Grieco, Daniela, Nicola Lacetera, Mario Macis, and Daniela Di Martino, "Motivating Cord Blood Donation With Information and Behavioral Nudges," *Scientific reports*, 2018, *8* (1), 1–12.

Grund, Christian and Dirk Sliwka, "Evidence on Performance Pay and Risk Aversion," *Economics Letters*, 2010, *106* (1), 8–11.

— and Tanja Hofmann, "The Dispersion of Bonus Payments Within and Between Firms," *Journal of Business Economics*, 2019, *89* (4), 417–445.

Gürerk, Özgür, Bernd Irlenbusch, and Bettina Rockenbach, "On Cooperation in Open Communities," *Journal of Public Economics*, 2014, *120*, 220–230.

Hallock, Kevin F, Regina Madalozzo, and Clayton G Reck, "CEO Pay-for-Performance Heterogeneity Using Quantile Regression," *Financial Review*, 2010, *45* (1), 1–19.

Hamilton, Barton H, Jack A Nickerson, and Hideo Owan, "Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation," *Journal of Political Economy*, 2003, *111* (3), 465–497.

Haylock, Michael, "Distributional Differences in the Time Horizon of Executive Compensation," *Empirical Economics*, 2022, *62*, 157–186.

— and Patrick Kampkötter, "Dataset: The Linked Personnel Panel (LPP)," *Data in Brief*, 2019, *27*, 104824.

— and **Patrick Kampkötter**, "The Role of Preferences, Attitudes, and Personality Traits in Labor Market Matching," *Economics Letters*, 2019, *185*, 108718.

Heckman, James, "Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations," *Journal of Human Resources*, 1997, *32* (3), 441–462.

Heckman, James J, Jora Stixrud, and Sergio Urzua, "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics*, 2006, *24* (3), 411–482.

Heger, Stephanie A, Robert Slonim, Ellen Garbarino, Carmen Wang, and Daniel Waller, "Redesigning the Market for Volunteers: A Donor Registry," *Management Science*, 2020, *66* (8), 3528–3541.

Hermalin, Benjamin E, "Toward an Economic Theory of Leadership: Leading by Example," *American Economic Review*, 1998, *88* (5), 1188–1206.

Hoffman, Mitchell and Steven Tadelis, "People Management Skills, Employee Attrition, and Manager Rewards: An Empirical Analysis," *Journal of Political Economy*, 2021, *129* (1), 243–285.

Holmstrom, Bengt, "Moral hazard and Observability," *Bell Journal of Economics*, 1979, *10* (1), 74–91.

— , "Moral Hazard in Teams," *The Bell Journal of Economics*, 1982, *13* (2), 324–340.

Imbens, Guido W. and Joshua D. Angrist, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 1994, *62* (2), 467–475.

Ishihara, Akifumi, "Relational Contracting and Endogenous Formation of Teamwork," *The RAND Journal of Economics*, 2017, *48* (2), 335–357.

Itoh, Hideshi, "Incentives to Help in Multi-Agent Situations," *Econometrica*, 1991, *59* (3), 611–636.

— , "Cooperation in Hierarchical Organizations: An Incentive Perspective," *Journal of Law, Economics and Organization*, 1992, *8*, 321.

Jensen, Michael C, "Value Maximization, Stakeholder Theory, and the Corporate Objective Function," *Journal of Applied Corporate Finance*, 2001, *14* (3), 8–21.

— **and Kevin J Murphy**, "Performance Pay and Top-Management Incentives," *Journal of Political Economy*, 1990, *98* (2), 225–264.

Kagel, John and Peter McGee, "Personality and Cooperation in Finitely Repeated Prisoner's Dilemma Games," *Economics Letters*, 2014, *124* (2), 274–277.

Kampkötter, Patrick, "Non-Executive Compensation in German and Swiss Banks Before and After the Financial Crisis," *The European Journal of Finance*, 2015, *21* (15), 1297–1316.

— **, Jens Mohrenweiser, Dirk Sliwka, Susanne Steffes, and Stefanie Wolter**, "Measuring the Use of Human Resources Practices and Employee Attitudes: The Linked Personnel Panel," *Evidence-based HRM: A Global Forum for Empirical Scholarship*, 2016, *4* (2), 94–115.

Kandel, Eugene and Edward P Lazear, "Peer Pressure and Partnerships," *Journal of Political Economy*, 1992, *100* (4), 801–817.

Kim, Tae-Yeol and Kwok Leung, "Forming and Reacting to Overall Fairness: A Cross-Cultural Comparison," *Organizational Behavior and Human Decision Processes*, 2007, *104* (1), 83–95.

Knez, Marc and Duncan Simester, "Firm-Wide Incentives and Mutual Monitoring at Continental Airlines," *Journal of Labor Economics*, 2001, *19* (4), 743–772.

Koenker, Roger, "Quantile Regression for Longitudinal Data," *Journal of Multivariate Analysis*, 2004, *91* (1), 74–89.

— , *Quantile Regression* Econometric Society Monographs, Cambridge University Press, 2005.

— **and Gilbert Bassett**, "Regression Quantiles," *Econometrica*, 1978, *46* (1), 33–50.

— **and** — , "Robust Tests for Heteroscedasticity Based on Regression Quantiles," *Econometrica*, 1982, *50* (1), 43–61.

Kosfeld, Michael, "The Role of Leaders in Inducing and Maintaining Cooperation: The CC Strategy," *The Leadership Quarterly*, 2020, *31* (3), 101292.

— **and Devesh Rustagi**, "Leader Punishment and Cooperation in Groups: Experimental Field Evidence From Commons Management in Ethiopia," *American Economic Review*, 2015, *105* (2), 747–83.

— **and Ferdinand A von Siemens**, "Worker Self-Selection and the Profits From Cooperation," *Journal of the European Economic Association*, 2009, *7* (2-3), 573–582.

— **and** — , "Competition, Cooperation, and Corporate Culture," *The RAND Journal of Economics*, 2011, *42* (1), 23–43.

Krueger, Alan B and David Schkade, "Sorting in the Labor Market Do Gregarious Workers Flock to Interactive Jobs?," *Journal of Human Resources*, 2008, *43* (4), 859–883.

Lacetera, Nicola, Mario Macis, and Robert Slonim, "Will There be Blood? Incentives and Displacement Effects in Pro-social Behavior," *American Economic Journal: Economic Policy*, 2012, *4* (1), 186–223.

— **,** — **, and Sarah S Stith**, "Removing Financial Barriers to Organ and Bone Marrow Donation: The Effect of Leave and Tax Legislation in the US," *Journal of Health Economics*, 2014, *33*, 43–56.

Lamarche, Carlos, "Robust Penalized Quantile Regression Estimation for Panel Data," *Journal of Econometrics*, 2010, *157* (2), 396–408.

Lang, Frieder R, Dennis John, Oliver Lüdtke, Jürgen Schupp, and Gert G Wagner, "Short Assessment of the Big Five: Robust Across Survey Methods Except Telephone Interviewing," *Behavior research methods*, 2011, *43* (2), 548–567.

Lazear, Edward P, "Pay Equality and Industrial Politics," *Journal of Political Economy*, 1989, *97* (3), 561–580.

Lazear, Edward P., Ulrike Malmendier, and Roberto A. Weber, "Sorting in Experiments With Application to Social Preferences," *American The Economic Journal: Applied Economics*, January 2012, *4* (1), 136–63.

Levitt, Steven D and John A List, "What do Laboratory Experiments Measuring Social Preferences Reveal About the Real World?," *Journal of Economic Perspectives*, 2007, *21* (2), 153–174.

Lindenlaub, Ilse, "Sorting Multidimensional Types: Theory and Application," *Review of Economic Studies*, 2017, *84* (2), 718–789.

Lown, RN, SGE Marsh, GE Switzer, KA Latham, JA Madrigal, and BE Shaw, "Ethnicity, Length of Time on the Register and Sex Predict Donor Availability at the Confirmatory Typing Stage," *Bone Marrow Transplantation*, 2014, *49* (4), 525–531.

Lown, Robert N and Bronwen E Shaw, "Beating the Odds: Factors Implicated in the Speed and Availability of Unrelated Haematopoietic Cell Donor Provision," *Bone Marrow Transplantation*, 2013, *48* (2), 210–219.

Machado, José AF and JMC Santos Silva, "Quantiles via Moments," *Journal of Econometrics*, 2019, *213* (1).

Malmendier, Ulrike and Geoffrey Tate, "CEO Overconfidence and Corporate Investment," *Journal of Finance*, 2005, *60* (6), 2661–2700.

__ and __ , "Who Makes Acquisitions? CEO Overconfidence and the Market's Reaction," *Journal of Financial Economics*, 2008, *89* (1), 20–43.

Marinovic, Iván and Felipe Varas, "CEO Horizon, Optimal Pay Duration, and the Escalation of Short-Termism," *Journal of Finance*, 2019, *74* (4), 2011–2053.

Mclaren, Patrick J, Melissa K Hyde, and Katherine M White, "Exploring the Role of Gender and Risk Perceptions in People's Decisions to Register as a Bone Marrow Donor," *Health Education Research*, 2012, *27* (3), 513–522.

Miettinen, Topi, Michael Kosfeld, Ernst Fehr, and Jörgen Weibull, "Revealed Preferences in a Sequential Prisoners' Dilemma: A Horse-Race Between Six Utility Functions," *Journal of Economic Behavior & Organization*, 2020, *173*, 1–25.

Monaghan, Madeline, Qi-Long Yi, Meagan Green, Todd Campbell, Jason T Weiss, Nicholas Dibdin, Dena Mercer, Heidi Elmoazzen, and David S Allan, "Factors Associated With Registrant Availability for Unrelated Adult Donor Hematopoietic Stem Cell Donation: Analysis of the Stem Cell Registry at Canadian Blood Services," *Transfusion*, 2021, *61* (1), 24–28.

Müller, Dana and Stefanie Wolter, "German Labour Market Data-Data Provision and Access for the International Scientific Community," *German Economic Review*, 2020, *21* (3), 313–333.

Murphy, Kevin J, "Executive Compensation," *Handbook of Labor Economics*, 1999, *3*, 2485–2563.

Myaskovsky, Larissa, Galen E Switzer, Mary A Dew, Jean M Goycoolea, Dennis L Confer, and Linda Abress, "The Association of Donor Center Characteristics With Attrition From the National Marrow Donor Registry," *Transplantation*, 2004, *77* (6), 874–880.

Naef, Michael and Jürgen Schupp, "Measuring Trust: Experiments and Surveys in Contrast and Combination," *SOEP Papers*, 2009.

Narayanan, MP, "Managerial Incentives for Short-Term Results," *Journal of Finance*, 1985, *40* (5), 1469–1484.

Onitilo, AA, YH Lin, EC Okonofua, Lawrence B Afrin, J Ariail, and BC Tilley, "Race, Education, and Knowledge of Bone Marrow Registry: Indicators of Willingness to Donate Bone Marrow Among African Americans and Caucasians," *Transplantation Proceedings*, 2004, *36* (10), 3212–3219.

Oster, Emily, "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business & Economic Statistics*, 2019, *37* (2), 187–204.

Parente, Paulo MDC and João MC Santos Silva, "Quantile Regression With Clustered Data," *Journal of Econometric Methods*, 2016, *5* (1), 1–15.

Patterson, Malcolm G, Michael A West, Viv J Shackleton, Jeremy F Dawson, Rebecca Lawthom, Sally Maitlis, David L Robinson, and Alison M Wallace, "Validating the Organizational Climate Measure: Links to Managerial Practices, Productivity and Innovation," *Journal of Organizational Behavior*, 2005, *26* (4), 379–408.

Peng, Lin and Ailsa Röell, "Managerial Incentives and Stock Price Manipulation," *Journal of Finance*, 2014, *69* (2), 487–526.

Porta, Rafael La, Florencio Lopez de Silanes, Andrei Shleifer, and Robert W Vishny, "Trust in Large Organizations," *American Economic Review*, 1997, *87* (2), 333–338.

Potters, Jan, Martin Sefton, and Lise Vesterlund, "Leading-by-Example and Signaling in Voluntary Contribution Games: An Experimental Study," *Economic Theory*, 2007, *33* (1), 169–182.

Proto, Eugenio, Aldo Rustichini, and Andis Sofianos, "Intelligence, Personality, and Gains From Cooperation in Repeated Interactions," *Journal of Political Economy*, 2019, *127* (3), 1351–1390.

Rand, Thomas M and Kenneth N Wexley, "Demonstration of the Effect,"Similar to Me," in Simulated Employment Interviews," *Psychological Reports*, 1975, *36* (2), 535–544.

Richter, David, Maria Metzing, Michael Weinhardt, and Jürgen Schupp, "SOEP Scales Manual," Technical Report, SOEP Survey Papers 2013.

Rotemberg, Julio J, "Human Relations in the Workplace," *Journal of Political Economy*, 1994, *102* (4), 684–717.

Roth, Alvin E, Tayfun Sönmez, and M Utku Ünver, "Kidney Exchange," *Quarterly Journal of Economics*, 2004, *119* (2), 457–488.

Rustagi, Devesh, Stefanie Engel, and Michael Kosfeld, "Conditional Cooperation and Costly Monitoring Explain Success in Forest Commons Management," *Science*, 2010, *330* (6006), 961–965.

Sacerdote, Bruce, "Peer Effects With Random Assignment: Results for Dartmouth Roommates," *Quarterly Journal of Economics*, 2001, *116* (2), 681–704.

Sauter, Jürgen, Ute V Solloch, Anette S Giani, Jan A Hofmann, and Alexander H Schmidt, "Simulation Shows That HLA-Matched Stem Cell Donors Can Remain Unidentified in Donor Searches," *Scientific Reports*, 2016, *6* (1), 1–9.

Schildberg-Hörisch, Hannah, "Are Risk Preferences Stable?," *Journal of Economic Perspectives*, 2018, *32* (2), 135–54.

Schmidt, Alexander H, Jürgen Sauter, Daniel M. Baier, Jessica Daiss, Andreas Keller, Anja Klussmeier, Thilo Mengling, Gabi Rall, Tobias Riethmüller, Gerhard Schöfl et al., "Immunogenetics in Stem Cell Donor Registry Work: The DKMS Example (Part 1)," *International Journal of Immunogenetics*, 2020, *47* (1), 13–23.

__, Jürgen Sauter, Daniel M Baier, Jessica Daiss, Andreas Keller, Anja Klussmeier, Thilo Mengling, Gabi Rall, Tobias Riethmüller, Gerhard Schöfl et al., "Immunogenetics in Stem Cell Donor Registry Work: The DKMS Example (Part 2)," *International Journal of Immunogenetics*, 2020, *47* (2), 139–148.

__, Thilo Mengling, Camila J Hernández-Frederick, Gabi Rall, Julia Pingel, Johannes Schetelig, and Gerhard Ehninger, "Retrospective Analysis of 37,287 Observation Years After Peripheral Blood Stem Cell Donation," *Biology of Blood and Marrow Transplantation*, 2017, *23* (6), 1011–1020.

__, Ute V Solloch, Daniel Baier, Alois Grathwohl, Jan Hofmann, Julia Pingel, Andrea Stahr, and Gerhard Ehninger, "Support of Unrelated Stem Cell Donor Searches by Donor Center-Initiated HLA Typing of Potentially Matching Donors," *PloS One*, 2011, *6* (5), e20268.

Schmitt, Manfred, Anna Baumert, Mario Gollwitzer, and Jürgen Maes, "The Justice Sensitivity Inventory: Factorial Validity, Location in the Personality Facet Space, Demographic Pattern, and Normative Data," *Social Justice Research*, 2010, *23* (2-3), 211–238.

Schöfl, Gerhard, Kathrin Lang, Philipp Quenzel, Irina Böhme, Jürgen Sauter, Jan A Hofmann, Julia Pingel, Alexander H Schmidt, and Vinzenz Lange, "2.7 Million Samples Genotyped for HLA by Next Generation Sequencing: Lessons Learned," *BMC genomics*, 2017, *18* (1), 1–16.

Schwaba, Ted and Wiebke Bleidorn, "Individual Differences in Personality Change Across the Adult Life Span," *Journal of Personality*, 2018, *86* (3), 450–464.

Sivasankaran, Adarsh, Eric Williams, Mark Albrecht, Galen E Switzer, Vladimir Cherkassky, and Martin Maiers, "Machine Learning Approach to Predicting Stem Cell Donor Availability," *Biology of Blood and Marrow Transplantation*, 2018, *24* (12), 2425–2432.

Sliwka, Dirk, "Trust as a Signal of a Social Norm and the Hidden Costs of Incentive Schemes," *American Economic Review*, 2007, *97* (3), 999–1012.

Slonim, Robert, Carmen Wang, and Ellen Garbarino, "The Market for Blood," *Journal of Economic Perspectives*, 2014, *28* (2), 177–96.

Stutzer, Alois, Lorenz Goette, and Michael Zehnder, "Active Decisions and Prosocial Behaviour: A Field Experiment on Blood Donation," *Economic Journal*, 2011, *121* (556), F476–F493.

Swanquist, Quinn Thomas and Robert Lowell Whited, "Out of Control: The (Over)Use of Controls in Accounting Research," *The Accounting Review*, Forthcoming.

Switzer, Galen E, Jessica G Bruce, Larissa Myaskovsky, Andrea DiMartini, Diana Shellmer, Dennis L Confer, Linda K Abress, Roberta J King, Allyson G Harnaha, Sibylle Ohngemach et al., "Race and ethnicity in decisions about unrelated hematopoietic stem cell donation," *Blood*, 2013, *121* (8), 1469–1476.

__, Larissa Myaskovsky, Jean M Goycoolea, Mary A Dew, Dennis L Confer, and Roberta King, "Factors Associated With Ambivalence About Bone Marrow Donation Among Newly Recruited Unrelated Potential Donors," *Transplantation*, 2003, *75* (9), 1517–1523.

__, Mario Macis, Rachel Fabi, Linda Abress, Dennis Confer, Jessica Bruce, Katie Howe, Sarah Fowler, Mary McNulty, Gabrielle Pastorek et al., "Providing Level-of-Match Information to Perfectly Matched Unrelated Stem Cell Donors: Evaluating Acceptability and Potential Changes in Donor Availability," *Biology of Blood and Marrow Transplantation*, 2018, *24* (10), 2110–2118.

__, Mary A Dew, Arthur A Stukas, Jean M Goycoolea, Janet Hegland, and Roberta G Simmons, "Factors Associated With Attrition From a National Bone Marrow Registry," *Bone Marrow Transplantation*, 1999, *24* (3), 313–319.

__, Mary Amanda Dew, Jean M Goycoolea, Larissa Myaskovsky, Linda Abress, and Dennis L Confer, "Attrition of Potential Bone Marrow Donors at Two Key Decision Points Leading to Donation," *Transplantation*, 2004, *77* (10), 1529–1534.

Tervio, Marko, "The Difference That CEOs Make: An Assignment Model Approach," *American Economic Review*, 2008, *98* (3), 642–68.

Wildman, John and Bruce Hollingsworth, "Blood Donation and the Nature of Altruism," *Journal of Health Economics*, 2009, *28* (2), 492–503.

Yermack, David, "Good Timing: CEO Stock Option Awards and Company News Announcements," *Journal of Finance*, 1997, *52* (2), 449–476.