Gutenberg School of Management and Economics

& Research Unit "Interdisciplinary Public Policy"

Discussion Paper Series

A critical assessment of the two-way fixed-effects model for firm-level dependent variables

Johannes Carow

March 18, 2024

Discussion paper number 2405

Johannes Gutenberg University Mainz
Gutenberg School of Management and Economics
Jakob-Welder-Weg 9
55128 Mainz
Germany
https://wiwi.uni-mainz.de/

Contact details

Johannes Carow
Chair of Applied Statistics and Econometrics and Ministry for Wirtschaft, Verkehr, Landwirtschaft und Weinbau, Rheinland-Pfalz
Johannes Gutenberg University Mainz
Jakob-Welder-Weg 4
55128 Mainz
Germany
johannes.carow@uni-mainz.de

A critical assessment of the two-way fixed-effects model for firm-level dependent variables*

Johannes Carow§
Johannes Gutenberg-University Mainz and MWVLW RLP

March 18, 2024

Abstract

Since the seminal work from Bertrand & Schoar (2003), the separate estimation of person effects and firm effects remains a widely used method in the analysis of firm-level dependent variables. Recently, this class of models has experienced serious methodological criticism, stating that person effects only reflect spurious variation. Rather than rejecting this estimation technique *per se*, I recommend a strategy based on simulation analysis to test for the presence of person effects. This strategy takes limitations of a previous test for idiosyncratic person effects into account. Further, I show that the estimation of person effects is subject to attenuation bias and that the size of this bias increases in the number of persons per firm-year. I also demonstrate that the use of Unconditional Quantile Regressions for estimated person effects can produce statistical artefacts at the upper and lower tail of the distribution. Additionally, attenuation bias impairs the analysis of the correlation of person effects pertaining to different dependent variables.

JEL classification C15, C18, C21, L25

Keywords: Two-way fixed-effects, simulations, managers, spurious variation, attenuation bias

*I gratefully acknowledge comments by Thorsten Schank, Christopher Koch, Daniel Schunk, Reyn van Ewijk, Sandra Kronenberger, Manuel Denzer, Vivien Voigt, Di Lu, Tiago Pereira, Stefan Schwarz, Ying Liang, Alexander Moog and Carl Hase as well as seminar participants at JGU Mainz. Excellent research assistance in the collection of data was provided by Christiane Buschinger and Benedikt Brandt. I declare that I am not exposed to any conflict of interest. The views expressed in this paper are those of the author and do not necessarily reflect those of the MWVLW RLP.

§Gutenberg School of Management and Economics, Jakob-Welder-Weg 4, 55128 Mainz,
✉ johannes.carow@uni-mainz.de

1 Introduction

The influence of managers and corporate directors on firm-level outcomes has been a central research topic during the past decades. While many studies consider management teams or corporate boards as collective bodies, the seminal work by Bertrand & Schoar (2003) has intensified the discussion on the role of individuals for corporate decision-making, initiating a new branch of literature on manager effects.[1] The main focus of this research area is to uncover the role of manager heterogeneity for the implementation of certain types of firm policies, unveiling trade-offs and complementarities between different policies. A key feature of the studies from this branch is the estimation of two-way fixed-effects models, allowing the separate identification of firm effects and person effects. This permits to analyse the impact of managers or directors without confounding their effect with time-constant firm characteristics. Consequently, this literature aims to provide insights into specific channels how the management can exert influence on firm-level outcomes. Two-way fixed-effects models for firm-level outcomes have been applied to a range of research contexts, for instance in the fields of accounting (Bamber et al., 2010; Dyreng et al., 2010; Gul et al., 2013; Dejong & Ling, 2013; Wells, 2020) or finance (Huang & Wang, 2015; Cavaco et al., 2017; Francis et al., 2020; Hagendorff et al., 2021; Baltrunaite et al., 2023).

Recently, the controversial nature of two-way fixed-effects models for firm-level dependent variables has become apparent: While Schoar et al. (2023) build upon the seminal contribution by Bertrand & Schoar (2003) to analyse the effect of managers on firms' exposure to systematic risk, Jarosiewicz & Ross (2023) re-investigate the findings from Bertrand & Schoar (2003) and argue that manager effects only reflect spurious variation. The key argument of the criticism by Jarosiewicz & Ross (2023) posits that managers should only have an effect at those firms where they are actually observed. Yet, they show that the explanatory power of manager effects persists and is sometimes even stronger once manager-spells are randomly assigned to other firms in the sample, contradicting this hypothesis. While the approach by Jarosiewicz & Ross (2023) of creating random counterfactuals and assessing the explanatory power of the model for the actual data against these counterfactuals is not entirely new,[2] it should still be considered a serious criticism.

Addressing these different methodological views on the validity of separate person and firm effects estimates, I provide three contributions: First, I use simulations to show the advantages and disadvantages of the scrambling procedure suggested by Jarosiewicz & Ross (2023). On the one hand, this procedure can be a powerful tool to test for the existence of idiosyncratic person effects. Researchers do not have to reject the estimation of two-way fixed-effects model for firm-level outcomes *per se*. I demonstrate that the plausibility of idiosyncratic person effects

[1] While the majority of these studies analyses managers, there is also related research on the effect of individual directors or auditors. Hence, the present paper uses terms such as manager style, manager (fixed-)effects or director (fixed-)effects depending on the context of the specific study to be discussed, while the term person (fixed-)effects is used as the more general expression.

[2] Bertrand & Schoar (2003), Dyreng et al. (2010), Fee et al. (2013) and Herpfer (2021) apply similar checks to break the link between persons and firms.

has to be tested for each dataset and every dependent variable of interest individually. On the other hand, I show that the scrambling procedure suggested by Jarosiewicz & Ross (2023) can perform poorly under certain scenarios. In datasets with low mobility, randomly scrambling the person-firm-spells tends to maintain the explanatory power of person effects even if person heterogeneity is part of the underlying data-generating process (DGP). Applied researchers might misinterpret the unchanged explanatory power of person effects, inferring that genuine person effects only reflect random noise. Therefore, I recommend a refinement of the scrambling test from Jarosiewicz & Ross (2023) as a strategy to test for idiosyncratic person effects: In a first step, one computes the change in explanatory power of person effects when scrambling the actual data. In a second step, this change should be benchmarked against a change in explanatory power for a simulated dependent variable. This takes the expected sensitivity of the scrambling test for a given structure of the dataset into account. I demonstrate this procedure by using a real-world dataset of publicly listed German firms.

Second, I show that even in the presence of genuine person effects, the estimated person effects are subject to attenuation bias when the dependent variable varies on the firm-level. I demonstrate that this bias increases in the number of persons per firm-year.[3] As a consequence, studies that relate the person effect to demographic or other characteristics might understate the measured correlations in absolute terms. Relatedly, I demonstrate that the application of Unconditional Quantile Regressions, a commonly used estimation technique introduced by Firpo et al. (2009), likely produces statistical artefacts at the upper and lower tails of the distribution when applied to person effects from two-way fixed-effects models.

Third, I analyse the correlation of person effects of two distinct dependent variables, an exercise frequently carried out by empirical studies. It serves to uncover trade-offs and complementarities between different corporate policy variables. I document an attenuation bias in this analysis as well. Hence, I find that also this type of analysis understates correlational patterns between person effects of different firm-level variables.

The remainder of the paper is structured as follows: Section 2 presents studies that apply two-way fixed-effects methods to settings with firm-level dependent variables. Section 3 sketches which data, either simulated or real-world, is used for the analysis. Afterwards, Section 4 presents the estimation procedure and details all its steps based on a simplified example. Section 5 presents the existing methodological criticism and investigates it by means of simulations, yielding a strategy to test for the presence of genuine person effects. It further derives different types of biases associated with the two-way fixed-effects model for firm-level dependent variables. Section 6 concludes.

[3]Hence, it might be particularly prevalent when applying the two-way fixed-effects model to analyse directors on large corporate boards.

2 Literature review

This section presents studies that estimate individual contributions to firm-level outcomes - a strand of literature that has attracted the attention of researchers during the past two decades.[4] The seminal study stems from Bertrand & Schoar (2003). They consider a sample of managers and analyse their individual impact on corporate policy variables such as investment, cash holdings or R&D expenditure. To rule out a confounding of person effects with firm effects, they include both kinds of fixed-effects in their model. Consequently, they can only identify the person effect for moving managers, motivating the name *Mover Dummy Variable (MDV)* approach. The study by Graham et al. (2012) constitutes an important extension to the approach by Bertrand & Schoar (2003). By borrowing from the labour economics literature, Graham et al. (2012) apply the *AKM* approach developed by Abowd, Kramaz and Margolis (Abowd et al., 1999) for separate identification of person and firm effects.[5] It allows to identify person effects for all firms connected by movers. Thus, it remedies the concern of endogenous sample selection that might be raised regarding Bertrand & Schoar (2003). These seminal contributions have spawned a plethora of follow-up studies that separately estimate manager and firm effects. While many of these studies consider person-level dependent variables such as executive compensation in Graham et al. (2012)[6], I focus on firm-level dependent variables in the present study to discuss statistical properties that arise from multiple persons that are assigned to one firm-year.[7]

A selection of studies that carry out a two-way fixed-effects estimation (via the AKM approach from Abowd et al. (1999) or the MDV approach from Bertrand & Schoar (2003)) for firm-level dependent variables are presented in Table 1. Those comprise the research fields of accounting (Bamber et al. (2010), Dyreng et al. (2010), Gul et al. (2013), Dejong & Ling (2013), Wells (2020)), finance (Bertrand & Schoar (2003), Huang & Wang (2015), Cavaco et al. (2017), Francis et al. (2020), Hagendorff et al. (2021), Baltrunaite et al. (2023), Schoar et al. (2023)) or innovation (Cho et al. (2016)). All of these studies have in common that multiple persons i are observed in firm j in year t such that they can jointly influence the dependent variable on the firm j-year t-level. Hence, these studies posit that managers, auditors or directors, respectively, have idiosyncratic time-invariant effects on firm-level outcomes even when controlling for firm fixed-effects. Table 1 shall illustrate both the variety of questions addressed with these methods as well as recurring types of analyses.

[4] Existing methodological criticism is provided in Section 5.1.

[5] The estimation procedure is described in detail in Section 4.

[6] While the main part of Graham et al. (2012) focuses on the person-level dependent variable of executive compensation, Table 7 from Graham et al. (2012) also analyses manager effects for different firm-level dependent variables like corporate policies.

[7] If only one person is observed per firm-year and each person belongs to only one firm at a time, the structure of the dataset is identical to person-level dependent variables, i.e. the setup which Abowd et al. (1999) originally consider. In case of only one person per firm-year, it is only a matter of definition whether the dependent variable is considered to vary on the firm-level or on the person-level while the structure of the dependent variable allows both interpretations. Hence, I do not focus on this stream of research. Examples of studies with only one person per firm-year that apply the AKM approach include Foerster et al. (2017), Davidson et al. (2019), Minaker (2021) and Fenizia (2022).

Table 1: List of studies that separately estimate person effects and firm effects, analyse a firm-level dependent variable and use data with multiple person observations per firm-year.

	Study	Dependent variable $y_{j(i)t}$	Decomposition of R^2 into components attributable to person effects and firm effects (based on or similar to Graham et al. (2012))	Relate person effects to person or firm characteristics	Correlation of person effects of two different dependent variables (or regression)	Data structure
1	Bertrand & Schoar (2003)	Corporate investment and financial policy variables			yes	i: managers j: firms
2	Bamber et al. (2010)	Financial disclosure		yes		i: managers j: firms
3	Dyreng et al. (2010)	Effective tax rate		yes	yes	i: managers j: firms
4	Gul et al. (2013)	Audit quality		yes		i: auditors j: firms
5	Dejong & Ling (2013)	Accruals			yes	i: managers j: firms
6	Huang & Wang (2015)	Fund performance	yes observable characteristics: 0.72 manager effects: 0.08 fund effects: 0.04 year effects: 0.06			i: managers j: funds
7	Cho et al. (2016)	Innovation measures	yes observable characteristics: 0.15 manager effects: 0.28 firm effects: 0.48			i: managers j: firms
8	Cavaco et al. (2017)	Firm performance		yes (quantile regressions)		i: directors j: firms
9	Francis et al. (2020)	Price terms of loan contracts	yes observable characteristics: 0.38 manager effects: 0.22 firm effects: 0.07 year effects: 0.14	yes	yes	i: managers j: firms
10	Wells (2020)	Accrual quality	yes observable characteristics and year effects: 0.33 manager effects: 0.19 firm effects: 0.22		yes (untabulated)	i: managers j: firms
11	Hagendorff et al. (2021)	Bank risk	yes manager effects: 0.23 other components untabulated	yes	yes	i: managers j: banks
12	Baltrunaite et al. (2023)	Total factor productivity		yes		i: directors j: firms
13	Schoar et al. (2023)	Systematic risk		yes	yes	i: managers j: firms

Note: The table contains a selection of studies that fulfil three criteria: First, these studies separately estimate person and firm effects, either based on the approach by Bertrand & Schoar (2003) or on the method from Abowd et al. (1999). Second, they use a dependent variable that varies on the firm-level (more general: on the j-level as specified in the last column). Third, there are firm-years where multiple persons i are observed. In the decomposition of R^2, the components add up to R^2 so that the remaining difference to 1.00 reflects the unexplained variation of the dependent variable. If multiple R^2 decompositions are presented in the studies, the decomposition for the baseline dependent variable based on the connectedness sample is reported.

As these methods allow to estimate person effects rather than only controlling for it, it is possible to relate the obtained person effects to other magnitudes in the data. Hence, Table 1 reports analyses frequently carried out in this literature: First, many of the aforementioned studies assess the explanatory power of person effects by computing the component of R^2 attributable to person effects, firm effects and observable components of the model. For instance, for an R^2 of 0.90 in a regression of fund performance, Huang & Wang (2015) attribute a component of 0.72 to observable manager and fund characteristics, 0.08 to manager effects, 0.04 to fund effects and another 0.06 to year effects. Wells (2020) finds that person effects explain 19 percent of the variation in the dependent variable accrual quality. Hence, this decomposition can underline the explanatory power of person effects, motivating their inclusion in addition to firm effects.

Second, previous studies estimate person effects and relate them to person or firm characteristics. For instance, Schoar et al. (2023) regress manager effects on a gender dummy as well as on an indicator whether the manager's career started during a recession. Further, Gul et al. (2013) link auditor fixed-effects to career paths and other individual properties. They show that holding graduate degrees coincides with more auditor aggressiveness while work experience in Big N firms correlates with a more conservative style of audit reporting. As a modified version of this sort of analysis, Cavaco et al. (2017) relate person effects to directors' characteristics by applying quantile regressions. In the selection process of independent directors, they argue that CEOs have a preference for a right-truncated ability distribution of independent directors, i.e. that they prefer independent directors with a lower ability. In contrast, shareholders are assumed to prefer high-ability directors, i.e. directors from a left-truncated distribution. Hence, Cavaco et al. (2017) investigate the correlation between director independence and the estimated director effect along the entire distribution of director effects. This allows to characterise the distribution from which independent directors are recruited, relating it to CEO- or shareholder-friendly selection patterns.

Third, another well-established analysis targets the correlation of person effects derived from different dependent variables. Bertrand & Schoar (2003) apply this analysis to unveil trade-offs and complementarities between different corporate policy variables. They show, for instance, that managers who are more inclined to foster external acquisitions coincide with lower R&D expenditure, sketching a trade-off between different expansion strategies. As another example, Hagendorff et al. (2021) analyse the role of individual managers for bank risk as measured by both market risk variables and policy variables. They find that person effects based on market risk variables are highly correlated with each other whereas these are only weakly linked to person effects from balance-sheet-based risk variables such as non-interest income, allowing for a comprehensive definition of managerial risk-taking.

While these three analyses appear to be most frequently applied, some studies suggest further usages of the estimated person effects: Schoar et al. (2023) and Wells (2020) consider the effects of executive transitions. Both estimate manager effects for given manager-firm-spells of a person. When the manager moves to a new firm, they document significant changes in the dependent variable of the new firm, highlighting the economic importance of person effects. Similarly, Huang & Wang (2015) investigate fund performance by focusing on the role of manager heterogeneity. They apply a rolling estimation window and show that the estimated manager effects are predictors for future fund performance.

Multiple studies posit that person effects from firm-level outcomes are important determinants of managers' careers and earnings. Huang & Wang (2015) show that high-ability fund managers are more likely to be promoted whereas low-ability managers correlate with a higher likelihood of demotion. In a similar manner, Francis et al. (2020) investigate the influence of managers for price terms for loan contracts. They argue that high-ability managers are able to reduce the cost of debt, which, in turn, should improve firm performance. However, they find that these beneficial outcomes for the firm are not reflected in manager compensation. Baltrunaite et al. (2023) construct a measure for director's talent, based on the director effects from total factor productivity and the individual rank of this director effect in comparison to co-directors. They analyse director replacements and characterise these replacements based on talent of leaving and new directors.

3 Data and descriptive statistics

Assessing properties of two-way fixed-effects estimates requires underlying datasets where multiple persons are observed per firm-year. To investigate this method for different structures of the dataset, most analyses of this study are based on two sorts of simulated data:

First, I use datasets that are entirely generated by simulations (Sections 5.2, 5.4 and 5.5). I simulate datasets with 100 firms observed over 15 years. When generating the sample, I start with boardsizes of 2, 4, ..., 20. However, those boardsizes are reduced by implementing delayed entry and attrition, enabling the analysis of unbalanced panels.[8] I assume that only 30 percent of all persons are observed in 2005 so that 70 percent of persons join the simulated sample during the remaining time span 2006 – 2019. Equivalently, 30 percent of all persons are observed in 2019 while the remaining 70 percent leave the sample in an earlier year. The years of entering or exiting the sample are drawn from independent uniform distributions. Considering delayed entry and random attrition reduces the average boardsize compared to the initial simulated sample with fixed board composition. Overall, the 10 resulting average boardsizes cover a range from approximately 1.3 to about 8. Using the data structure with delayed entry and attrition, I proceed by implementing mobility as follows: I allow for one move per person. Hence, movers are subsequently observed in two firms whereas non-movers are observed in only one firm. By restricting the number of moves to one, I avoid that the number of moves per person confounds the simulation results. I introduce mobility by assigning random numbers to all persons in the

[8] The simulation analysis in Section 5 tests how the explanatory power of the two-way fixed-effects model changes when scrambling the dataset by merging person-firm spells to other randomly selected firms. Hence, it is crucial to introduce some source of unbalancedness to the simulation samples. If all persons were observed during the same estimation window and had the same spell structure, the test of scrambling spells would be meaningless.

sample and by defining all persons with the random number under a given threshold as movers. The year of changing the firm is drawn from a uniform distribution over all the years where a person is observed. Similarly, the new firm of a mover is randomly drawn from the remaining firms in the sample. I consider 10 different shares of movers (0.05, 0.15, ..., 0.95). Hence, all simulation iterations can be stratified along 100 boardsize - mobility combinations. They are visualised in Figure A2. Focusing on different structures of the dataset allows the analysis of different scenarios where the two-way fixed-effects model might be applied. While the number of managers per firm-year is typically small, studies that focus on corporate boards might feature large boards (Cavaco et al., 2017). Additionally, since the work of Andrews et al. (2008), it is known that a low amount of mobility can lead to biases in certain types of analyses based on estimated person effects.[9] Therefore, I also systematically vary this magnitude in the simulated datasets.

Second, I use the structure of a real-world dataset (used in Sections 5.3 and 5.4 while the real-world dataset is described in the next paragraph) and replace the dependent variable by simulated data with different DGPs. This approach is common practice, for instance in studies that analyse the CEO effect on firm outcomes (Fee et al., 2013; Fitza, 2014; Quigley & Graffin, 2017; Fitza, 2017). Using the structure of real-world datasets overcomes limitations of entirely simulated datasets. The latter might not be able to capture all structural features of real-world dataset such as the distribution of spell durations or boardsizes. Hence, replacing the dependent variable from a real-world dataset with a specific DGP offers the advantage of maintaining all the structural features of a real-world dataset while imposing a certain dependent variable.

To assess whether the methodological concerns for the separate identification of person effects and firm effects[10] apply to real-world data, Section 5.3 of the present study considers a sample of supervisory board directors of German firms that were publicly listed in 2009. When assembling the sample, I start with a total of 178 firms. I exclude 18 firms from the financial sector, consistent with previous studies (Cavaco et al., 2017). Further, I take into account that the empirical methodology (as described in Section 4) yields results that are comparable only within a set of firms connected by persons. Hence, I exclude four further firms without any director connections to other supervisory boards in the sample, restricting the final sample to 156 firms. The estimation window covers 15 years, ranging from 2005 to 2019. Data on the supervisory board composition is hand-collected and complemented by data from *BoardEx*.[11] As dependent variables, I consider two accounting-based performance measures (ROA, ROE) and one market-based performance measure (stock return). Data on ROA and ROE is manually collected and complemented by information from *Orbis*. Further, data on annual stock return (adjusted for dividend payments) is computed based on data from *boerse.de*. To rule out the impact of outliers, I winsorise all three performance variables on the 1 % and 99 % percentiles.

[9] Andrews et al. (2008) consider the correlation of person effects and firm effects in a two-way fixed-effects model for workers' wages. They report a downward bias in the estimation of this correlation for true positive correlations. Further, they highlight that this downard bias is particularly pronounced if the share of movers is low. Analysing the correlation of person effects and firm effects is beyond the scope of the present study.

[10] The studies that criticise this strand of literature are presented in Section 5.1.

[11] Data on board composition and firm performance was collected at the Chair of Applied Statistics and Econometrics (JGU Mainz), inter alia by Viktor Bozhinov, see Bozhinov (2019).

Due to the methodological nature of this paper, I concentrate the analysis on the boardsize[12] and the share of movers, i.e. on magnitudes that concern the structure of the data. The consideration of further demographic characteristics would increase the complexity of the analysis without providing a specific value-added.

Table 2 presents descriptive statistics for the three firm performance variables (Panel A) and additional figures on the structure of the dataset (Panel B). While each of the three dependent variables covers a range over positive and negative values, their sample averages take on positive values. The market-based performance measure stock return exhibits a higher volatility than ROA and ROE. In sum, 4,076 directors are observed in 156 firms during the years 2005 – 2019. Each person-firm-spell is, on average, observed for 8.11 years. As the sample focuses on large publicly-listed firms, the average boardsize amounts to 14 seats. While the share of movers is 11 percent, these directors account for 25 percent of all firm-person-year observations.

Table 2: Descriptive Statistics

Panel A: Dependent variables

	Mean	SD	Min	Max
ROA	5.09	8.76	-67.89	36.18
ROE	11.63	30.05	-187.64	81.03
Stock Return	13.19	40.55	-86.43	239.00

Panel B: Structure of the data

Firms	156	Firm-years	2,034
Persons	4,076	Person-firm combinations (spells)	4,732
Firm-person years	24,036	Average duration of a spell (in years)	8.11
Average boardsize	14.07	Average board appointments per person	1.17
Share of movers	0.11	Share of mover observations	0.25

Note: The data consists of supervisory board directors of German firms that were publicly listed in 2009. It covers the years 2005 – 2019. Panel A presents the three firm performance measures that are used as dependent variables for the two-way fixed-effects estimation explained in Section 4.1. All three variables are measured in percent and winsorised on the 1 % and 99 % percentiles to reduce the impact of outliers. Panel B depicts the structure of the dataset. Shares are represented as numbers between 0 and 1.

The analyses in Section 5 rely on a scrambling procedure suggested by Jarosiewicz & Ross (2023) which randomly shuffles the person-firm-spells to other firms. The structure of the original dataset as presented in Table 2 remains unchanged in the sense that the number of firms, persons etc. remains constant. However, by assigning directors to new firms, the director composition of each firm is changed. Thus, in the scrambled sample, the directors face new co-directors as compared to the original sample. Further, by breaking the director-firm-connections, each director is matched to firm-year-level outcomes of other firms. This analysis will help to investigate whether director effects estimated in the original sample reflect variation that is due to the influence of directors on firm-level outcomes or whether these effects can also be measured in scrambled samples. Due to the random assignment of directors to firms in the scrambling process, no genuine effect of directors on firm outcomes can be expected. Hence, Jarosiewicz & Ross (2023) use this procedure to investigate whether person effects only reflect random noise.

[12] In the institutional sense, the term *boardsize* is mainly used in the context of corporate boards. The present study adopts the term *boardsize* to refer to the number of persons per firm-year, regardless of whether it describes directors on a corporate board or managers in a firm.

4 Empirical specification: Two-way fixed-effects model

This study discusses the application of the two-way fixed-effects model for firm-level dependent variables. Section 4.1 presents the estimation procedure for this model. For illustration purposes, Section 4.2 details all the steps of the estimation procedure based on a simplified example.

4.1 Setup of the two-way fixed-effects model

A growing number of studies attempt to disentangle outcome variables into separate firm and person components.[13] Those two-way fixed-effects models are frequently estimated by the *AKM* approach which was developed by Abowd et al. (1999) and subsequently named after the authors Abowd, Kramarz and Margolis. This class of models is based on the separate identification of person effects and firm effects. They are included as additively separable terms.[14]

In the initial version by Abowd et al. (1999), this model is used to decompose worker i's wages. However, a number of studies have employed this technique to analyse effects on firm-year-level outcome variables. Hence, I consider the following specification for firm j of director i in year t:

$$y_{j(i)t} = \gamma_1 X_{it} + \gamma_2 W_{jt} + \delta_t + \theta_i + \psi_j + u_{ijt} \tag{1}$$

where $y_{j(i)t}$ is an outcome variable on the firm-year-level, such as firm performance. X_{it} and W_{jt} represent vectors of individual-level and firm-level characteristics. Further, δ_t captures year effects and u_{ijt} is the idiosyncratic error term. The model separately specifies person fixed-effects θ_i and firm fixed-effects ψ_j. Hence, θ_i captures ability, i.e. a time-invariant component of $y_{j(i)t}$ that is attributable to a specific person.

The estimation procedure runs as follows: In a first step, the larger set of fixed-effects θ_i is eliminated by subtracting the person i-level averages of each variable (Abowd et al., 1999):

$$y_{j(i)t} - \overline{y_i} = \gamma_1(X_{it} - \overline{X_i}) + \gamma_2(W_{jt} - \overline{W_{ij}}) + (\delta_t - \overline{\delta_t}) + (\psi_j - \overline{\psi_{ij}}) + (u_{ijt} - \overline{u_i}) \tag{2}$$

After this within-transformation, the firm effects are no longer binary indicators. Instead, the expression $(\psi_j - \overline{\psi_{ji}})$ reflects a within-person mean-deviated term. Let F_{ijt} denote a dummy that equals one if person i was employed at firm j in year t. Equation (2) becomes:

$$y_{j(i)t} - \overline{y_i} = \gamma_1(X_{it} - \overline{X_i}) + \gamma_2(W_{jt} - \overline{W_{ij}}) + (\delta_t - \overline{\delta_t}) + \sum_{j=1}^{J}(F_{ijt} - \overline{F_{ij}})\psi_j + (u_{ijt} - \overline{u_i}) \tag{3}$$

In a second step, this de-meaned equation allows to obtain estimates of ψ_j as well as γ_1 and γ_2.

[13] I closely follow the notation and the presentation of the two-way fixed-effects model from Graham et al. (2012), Ewens & Rhodes-Kropf (2015) and Ma (2018).

[14] Whenever the researcher is purely interested in controlling for time-invariant person-level and firm-level heterogeneity rather than estimating the fixed-effects, she might include interacted person-firm effects, i.e. spell fixed-effects, as described in Andrews et al. (2006).

Plugging these estimates to the within-person deviated version of (1) and rearranging yields the estimates for person effects θ_i:

$$\hat{\theta}_i = \overline{y_i} - \hat{\gamma}_1 \overline{X_i} - \hat{\gamma}_2 \overline{W_{ji}} - \overline{\delta_t} - \sum_{j=1}^{J} \overline{F_{ij}} \hat{\psi}_j \tag{4}$$

This procedure illustrates the identification requirements for person effects and firm effects. The expression $(F_{ijt} - \overline{F_{ij}})$ in (3) is zero for all persons i who never change the firm in the sample. Hence, firm effects and, subsequently, person effects, are identified by movers. However, Abowd et al. (1999) and Abowd et al. (2002) highlight that the estimation sample is not limited to directors switching between firms. Instead, the estimated person effects $\hat{\theta}_i$ are comparable across firms that are connected by movers.[15] For this reason, empirical studies applying the AKM approach typically reduce the available sample to the largest connected set, including the present study (see Section 3).[16]

Card et al. (2013) emphasise the assumptions about the assignment process between units i and j. More specifically, they underline the assumption of exogenous mobility needed for identification of person and firm effects.[17] In the present setting, exogenous mobility means that the error component u_{ijt} in equation (1) is not correlated with the mobility patterns. Stated differently, there shall be no selection on time-variant unobservables of person i to firm j. As the model includes firm and person effects separately, it does not account for a match component. Therefore, the idiosyncratic match component, i.e. an indicator for each firm-person-combination, is absorbed by the error term (Card et al., 2013).

Previous studies, inter alia Graham et al. (2012) or Francis et al. (2020), decompose the explanatory power of the two-way fixed-effects model into components attributable to each component of equation (1):

$$R^2 = \frac{Cov(y_{j(i)t}, X_{ijt}\hat{\gamma}_1 + W_{jt}\hat{\gamma}_2)}{Var(y_{j(i)t})} + \frac{Cov(y_{j(i)t}, \hat{\delta}_t)}{Var(y_{j(i)t})} + \frac{Cov(y_{j(i)t}, \hat{\theta}_i)}{Var(y_{j(i)t})} + \frac{Cov(y_{j(i)t}, \hat{\psi}_j)}{Var(y_{j(i)t})} \tag{5}$$

4.2 Illustrative example

This section illustrates the derivation of person effects and firm effects as outlined in Section 4.1. For simplicity, I abstract from covariate vectors X_{it} and W_{jt} as well as from time effects δ_t. Consider the following simplified example of two firms and two years as presented in Panel A of Table 3. Within each firm-year, five persons are observed, making this setting comparable to corporate boards and directors. The two firms are connected via one mover (person 5) which

[15] Abowd et al. (2002) specify an algorithm which allows to decompose the sample into mutually exclusive sets of connected firm-person-year observations.

[16] Bertrand & Schoar (2003) separate manager effects and firm effects by demeaning over each manager i. Hence, manager effects are only identified for movers in their study. Thus, their mover dummy variable (MDV) approach coincides with stronger sample restrictions than the AKM approach.

[17] In Section A.1, I discuss the assumption of exogenous mobility for the real-world data used in this study.

provides the basis for identification of person effects and firm effects. The outcome of interest y_{jt} varies on the firm-year-level.

As outlined above, the two-way fixed-effects model is estimated by demeaning on the person i-level first in order to remove the person effects θ_i. Panel B of Table 3 illustrates the intermediate estimation steps. For the dependent variable, the demeaning yields the difference $y_{j(i)t} - \overline{y_i}$. Averaging each firm indicator F_{ij} on the person-level yields the information which share of person-year observations person i spends at firm j. Hence, the non-movers are assigned values of zero or one for F_{ij} while the movers receive shares such that $\sum_{j=1}^{J} F_{ij} = 1$. As pointed out, the movers identify the firm effects. Only for the movers, the differences $F_{i1t} - \overline{F_{i1}}$ and $F_{i2t} - \overline{F_{i2}}$ are unequal to zero. Hence, relating the demeaned dependent variable to the demeaned firm indicators for the only mover in the sample identifies a firm effect of firm 2 of -0.27 with firm 1 being the reference category. While the firm effect $\hat{\psi}_j$ is estimated via the only mover in the sample, its estimates are used to obtain person effects for all the other directors in the connected set. Given $\hat{\psi}_j$, the person effects $\hat{\theta}_i$ are recovered as in (4), i.e. by computing the difference between $\overline{y_i}$ and the estimated firm effect. To normalise the resulting values of $\widetilde{\theta_i}$, its mean of 0.975 is subtracted, yielding the final estimates $\hat{\theta}_i$ for the person effects.

The interpretation of the estimated person effects $\hat{\theta}_i$ should take the structure of the data set into account: It is apparent that persons with the same pattern of spells are assigned the same person effects. For instance, persons 1, 2, 3, and 4 are all observed during years 1 and 2 in firm 1, making their specific contribution to $y_{j(i)t}$ indistinguishable from each other. Hence, their person effects amount to -0.165 each. The only mover that is part of firm 1, gets a higher person effect of 0.155. This is due to the higher within-person mean of the dependent variable $\overline{y_i}$ of 0.995 as compared to 0.81 of most other persons at firm 1. This discrepancy in $\overline{y_i}$ results from person 5's affiliation to firm 2 in the second period where a higher dependent variable is observed than in firm 1 (0.86 vs. 0.49). Another component that leads to the difference in person effects between person 5 and, for instance, person 1 consists in the different person-mean of firm effects. Thus, based on a rearranged and simplified version of (4):

$$
\begin{aligned}
\hat{\theta}_5 - \hat{\theta}_1 &= (\overline{y_5} - \overline{\hat{\psi}_{j5}}) - (\overline{y_1} - \overline{\hat{\psi}_{j1}}) \\
0.155 - (-0.165) &= (0.995 - (-0.135)) - (0.81 - 0) = 0.32
\end{aligned} \tag{6}
$$

Put differently, person 5's estimated person effects exceeds $\hat{\theta}_1$ for two reasons: First, moving to firm 2 in the second period coincides with a larger $y_{j(i)t}$ and, consequently, with a higher $\overline{y_i}$ as compared to the counterfactual scenario of remaining in firm 1.[18] Second, firm 2 coincides with a lower firm effect, reducing the firm effect averaged over all observations of person 5.

[18] In the latter case, there would be no mover in the sample, creating a collinearity between person effects and firm effects and, thus, making it impossible to separately identify both.

Table 3: Illustrative Example

	Panel A				Panel B								
	firm j	year t	person i	$y_{j(i)t}$	$\overline{y_i}$	$y_{j(i)t}-\overline{y_i}$	$\overline{F_{i1}}$	$F_{i1t}-\overline{F_{i1}}$	$\overline{F_{i2}}$	$F_{i2t}-\overline{F_{i2}}$	$\hat{\psi}_j$	$\widetilde{\theta_i}$	$\hat{\theta}_i$
1	1	1	1	1.13	0.81	0.32	1	0	0	0	0	0.81	-0.165
2	1	1	2	1.13	0.81	0.32	1	0	0	0	0	0.81	-0.165
3	1	1	3	1.13	0.81	0.32	1	0	0	0	0	0.81	-0.165
4	1	1	4	1.13	0.81	0.32	1	0	0	0	0	0.81	-0.165
5	1	1	5	1.13	0.995	0.135	0.5	0.5	0.5	-0.5	0	1.13	0.155
6	1	2	1	0.49	0.81	-0.32	1	0	0	0	0	0.81	-0.165
7	1	2	2	0.49	0.81	-0.32	1	0	0	0	0	0.81	-0.165
8	1	2	3	0.49	0.81	-0.32	1	0	0	0	0	0.81	-0.165
9	1	2	4	0.49	0.81	-0.32	1	0	0	0	0	0.81	-0.165
10	1	2	6	0.49	0.49	0	1	0	0	0	0	0.49	-0.485
11	2	1	7	0.88	0.87	0.01	0	0	1	0	-0.27	1.14	0.165
12	2	1	8	0.88	0.87	0.01	0	0	1	0	-0.27	1.14	0.165
13	2	1	9	0.88	0.87	0.01	0	0	1	0	-0.27	1.14	0.165
14	2	1	10	0.88	0.87	0.01	0	0	1	0	-0.27	1.14	0.165
15	2	1	11	0.88	0.88	0	0	0	1	0	-0.27	1.15	0.175
16	2	2	5	0.86	0.995	-0.135	0.5	-0.5	0.5	0.5	-0.27	1.13	0.155
17	2	2	7	0.86	0.87	-0.01	0	0	1	0	-0.27	1.14	0.165
18	2	2	8	0.86	0.87	-0.01	0	0	1	0	-0.27	1.14	0.165
19	2	2	9	0.86	0.87	-0.01	0	0	1	0	-0.27	1.14	0.165
20	2	2	10	0.86	0.87	-0.01	0	0	1	0	-0.27	1.14	0.165

Note: Simplified example dataset with two firms observed over two years, connected by one mover (person 5). The firm-year outcome is given by $y_{j(i)t}$ where j indexes firms, i indexes persons and t indexes years. The within-person deviated person identifiers are omitted as they consist only of zeros by definition. The notation is based on the presentation of the two-way fixed-effects model in Graham et al. (2012), Ewens & Rhodes-Kropf (2015) or Ma (2018). The mean of $\widetilde{\theta_i}$ amounts to 0.975. Hence, $\widetilde{\theta_i}$ is normalised by subtracting 0.975 from each of its values, yielding the final estimates $\hat{\theta}_i$ for the person effects.

5 Testing the use of the two-way fixed-effects model for firm-level outcomes

The large set of studies that apply methods for separate identification of person effects and firm effects (see Section 2 for an overview) has spawned multiple studies that criticise these approaches. Subsection 5.1 gives an overview on the methodological criticism. Afterwards, Section 5.2 contains simulations to check whether the existing criticism is justified under different scenarios. In doing so, it focuses on the scrambling check from Jarosiewicz & Ross (2023), the most recently suggested test for the presence of time-invariant person effects, and analyses the sensitivity of this test. Section 5.3 applies this scrambling to the present dataset and suggests a strategy for applied researchers to test for idiosyncratic person effects. Further, Section 5.4 uses simulations to compare the person effects specified in the DGP with the corresponding estimated person effects. This section documents attenuation bias in the estimation of person effects. It further shows that attenuation bias increases in boardsize. Additionally, it demonstrates a statistical artefact at the upper and lower tail when applying Unconditional Quantile Regressions to estimated person effects. Finally, Section 5.5 studies the correlation between person effects of two different dependent variables, an analysis frequently carried out by studies on manager styles. Again, it shows that these correlations are subject to attenuation bias.

5.1 Existing methodological criticism

This section assembles contributions that criticise the usage of two-way fixed-effects models in the research of manager styles (see Abernethy & Wallis (2019) for an overview of empirical challenges). A central criticism to the research on manager styles, primarily addressed to the MDV approach from Bertrand & Schoar (2003), but also applicable to AKM (Abernethy & Wallis, 2019), stems from Fee et al. (2013). They question the basic assumption of manager styles, namely the existence of time-invariant and firm-invariant idiosyncratic manager styles. Fee et al. (2013) find that that manager styles are no portable component that is the same for all employers. Hence, they do not confirm the finding of Bertrand & Schoar (2003) that manager effects at the old and the new employer are strongly correlated. Moreover, they randomly scramble moving CEOs to other firms and show that this procedure does not remove the significance of the person effects, thus questioning the methodology. Further, they hypothesise that their spurious results might stem from autocorrelation of their firm-level dependent variables. Accordingly, they simulate an autoregressive process of dependent variables and show that the significance of their results persists when using this randomly simulated firm-level variable, underlining the spurious nature of person effects. Additionally, they question the interpretation of person effects previously applied in the literature. They compare changes in firm policy variables after exogenous manager turnover, as induced by death or health reasons, and endogenous manager turnover. They report policy changes only after endogenous manager turnover, suggesting that the turnover results from corporate boards' existing preferences for policy shifts. Hence, they argue that if the concept of manager style exists in practice, it reflects a style selected by the board rather than active influence of managers.

Similar to Fee et al. (2013), Blettner et al. (2012) also discuss estimation issues regarding the CEO effect. They point out that the setup by Bertrand & Schoar (2003) implies the assumption that the CEO effect is not only invariant per firm but also independent of previous CEO experience. Further, they state that manager effects are confounded by time effects. Following their argumentation, many studies try to control for time effects by including year dummies in regressions. However, those yearly averages are computed over all firms in the sample and do not consider that firms might be differently well adapted to the current economic situation. Thus, they highlight that it is difficult to separate a manager effect from the timing effect when a manager is employed in a given company. Further, Blettner et al. (2012) note that a manager effect can be influenced by previous managers' decisions. Framed differently, they argue that a current manager might benefit from performance-enhancing actions of previous managers, thus confounding the effect of both managers. Note that this criticism potentially applies to managers with all tenures and that the recommendation of Bertrand & Schoar (2003) to restrict the sample to managers observed for at least three years in a firm does not remedy this concern.

Another strand of literature investigates whether manager effects represent statistical artefacts and occur only due to chance (Fitza, 2014; Quigley & Graffin, 2017; Fitza, 2017). These studies carry out ANOVA analyses to decompose the variance of firm performance into components attributable to different determinants, including a manager effect. Fitza (2014) reports an increment of approximately 18 percent in the explanatory power of the model when adding CEO effects. However, Fitza (2014) argues that this effect has to be compared to a baseline of a randomly generated variable with the same mean and the same variance to develop an appropriate comparison benchmark. In this case, CEO effects would explain about 13 percent of the variation in the dependent variable. This points to a component of random noise embodied in the estimated CEO effect that has to be extracted when measuring the CEO effect. Quigley & Graffin (2017) respond to the study by Fitza (2014). They argue that the randomness component that Fitza (2014) identifies as part of the CEO effect is overstated. Their methodological improvement consists in the application of multilevel modelling due to restrictive assumptions of the ANOVA model employed by Fitza (2014). This methodological change reduces the randomness component in the CEO effect from 13 percent to 0.1 percent. Still, Fitza (2017) defends the results from Fitza (2014). He shows that the results are also robust to multilevel modelling. Further, he argues that the results are not driven by the order in which variables are incrementally added to the model when applying ANOVA.

In a recent contribution, Jarosiewicz & Ross (2023) argue that manager styles computed by either the MDV or the AKM approach reflect random noise. They employ the same dataset as Bertrand & Schoar (2003) and reconsider the actual explanatory power of manager fixed-effects. For this purpose, they randomly scramble manager spells to other firms to investigate whether results persist. They find that this random allocation of managers to other firms does not reduce the amount of variation explained by manager effects. Additionally, they apply another placebo test already suggested by Bertrand & Schoar (2003): To test the assumption of constant person effects across firms, Bertrand & Schoar (2003) compute firm-year-level residuals from regressions of a firm policy variable on firm effects, year effects and covariates. They compute the person-firm-level means of those residuals and compare those between the first and the second firm where each manager is observed. The placebo test consists in placing the manager spells three years prior to their actual observations to the firms where they are observed. Bertrand & Schoar (2003) find that this placebo test substantially weakens the link between person-firm-average residuals, supporting the notion of invariant person effects across firms. Jarosiewicz & Ross (2023) question this analysis and apply a random scrambling of manager-spells to other firms also to this placebo test. Jarosiewicz & Ross (2023) find that the scrambled placebo data still shows similar correlational patterns between person-level average residuals between different firms where a person is observed. Thus, Jarosiewicz & Ross (2023) interpret the findings from Bertrand & Schoar (2003) such that person effects from the MDV or AKM approach reflect random noise. Consequently, they argue that the concept of managerial style is more complex than person fixed-effects derived from policy variables as the latter reflect statistical artefacts.

In Section, 5.2, I carry out simulations to assess which outcomes of the scrambling procedure by Jarosiewicz & Ross (2023) can be expected under different scenarios. Further, in Section 5.3, I turn to a real-world dataset of 156 publicly listed German firms to evaluate for three commonly used dependent variables whether they are consistent with idiosyncratic person effects. In this context, I provide guidelines for applied researchers regarding which analyses should be carried out before estimating a two-way fixed-effects model for a firm-level dependent variable.

5.2 Testing for the presence of person effects: Insights from simulations

Jarosiewicz & Ross (2023) question the findings of Bertrand & Schoar (2003). They argue that the estimated manager effects in equations with firm-level dependent variables represent only random noise rather than managerial styles. Their criticism is based on scrambling manager spells to other firms. They show that this random allocation of managers to firms does not reduce the explanatory power of the person effects, suggesting that the latter represent only random noise and lead to spurious results.

To address this concern, I run simulations to assess whether the two-way fixed-effects model is appropriate to analyse variables on the firm-year-level. I use the following DGPs where the person effect θ_i, the firm effect ψ_j and the error term u_{jt}[19] are all drawn from independent standard normal distributions and n_{jt} denotes the number of persons that are observed in firm j in year t:

1) The firm-level dependent variable follows a two-way fixed-effects model where the sum of person fixed-effects θ_i per firm-year enters the DGP:

$$y_{j(i)t} = (\sum_{i=1}^{n_{jt}} \theta_{i \in (jt)}) + \psi_j + u_{jt} \tag{7}$$

2) The firm-level dependent variable follows a two-way fixed-effects model where the mean of person fixed-effects θ_i per firm-year enters the DGP:

$$y_{j(i)t} = (\frac{1}{n_{jt}} * \sum_{i=1}^{n_{jt}} \theta_{i \in (jt)}) + \psi_j + u_{jt} \tag{8}$$

3) The firm-level dependent variable follows a one-way fixed-effects model:

$$y_{j(i)t} = \psi_j + u_{jt} \tag{9}$$

4) The firm-level dependent variable is entirely driven by random shocks:

$$y_{j(i)t} = u_{jt} \tag{10}$$

The models (7) and (8) assume that both person effects θ_i and firm effects ψ_j are part of the DGP for the firm-level variable. This would imply that each person has an idiosyncratic effect on the firm outcome. Put differently, person heterogeneity would matter in addition to time-constant firm heterogeneity. Model (7) includes the sum of person effects per firm-year. The effect of a given person on the outcome is assumed to be independent of boardsize as all individual person effects enter linearly with a coefficient of 1. This is motivated by equation (1) which disentangles each firm-level outcome by a model where person effects also enter linearly with a coefficient of 1. In contrast, equation (8) features the inclusion of the average of person effects. Hence, the

[19] While the estimated error term varies on the person-firm-year level, the error term as used in the DGP varies on the firm-year level to ensure that there is no within-firm-year variation in the dependent variable.

impact of one person on the dependent variable is assumed to depend on boardsize. While I consider equation (7) as the main DGP specification consistent with the two-way fixed-effects model, I maintain the DGP from equation (8) for illustration purposes to show whether results differ when the researcher assumes person effects to enter the DGP linearly with coefficient $\frac{1}{n_{jt}}$.[20]

Dropping person-level heterogeneity, the DGP specified in equation (9) is based on the notion that only firm-level heterogeneity together with random shocks matter to determine firm-level outcomes. If this represents the underlying DGP for a given firm-level dependent variable, the inclusion of person effects should add no explanatory power beyond firm effects and, thus, reflect only random noise according to Jarosiewicz & Ross (2023). As an additional variant of equation (9), I further analyse a DGP where the firm-level dependent variables is entirely driven by random shocks as given by equation (10). This provides a comparison benchmark for the DGP from equation (9) as it rules out that estimated person effects absorb the variation that is due to firm effects.

For these four DGPs, I analyse the explanatory power of person-level heterogeneity. To do so, I follow Jarosiewicz & Ross (2023) by testing whether the explanatory power of person effects survives the random scrambling of person spells to new firms. If the respective person-firm-spell is arbitrarily assigned to another firm, the explanatory power should reduce substantially. Conversely, Jarosiewicz & Ross (2023) posit that unchanged or rising explanatory power of person effects in the case of scrambling questions the existence of idiosyncratic person effects.

To scramble the person-firm spells, I proceed as follows: Similar to Jarosiewicz & Ross (2023), I assign consecutive index numbers to each spell, denoted as *old spell*. Afterwards, I draw random numbers from a uniform distribution and order them by magnitude, generating *new spell* numbers for all person-firm combinations based on their ranking from the random number. Next, I replace the firm identifier of the *old spell* by the firm identifier of the *new spell* with the same number. Additionally, I take into account that not all spells are compatible with all firms. For instance, merging a spell which is observed from 2005 – 2009 to a firm which is observed in 2010 and thereafter would generate missing values. In cases where data might be lost, I randomly select a different firm based on a uniform distribution over all the firm identifiers in the sample. In Figure A4, I plot the distribution of correlation coefficients between the dependent variables for one person-firm-year in the original sample and the dependent variable from the scrambled sample. Those are strongly concentrated around zero, fostering the random re-allocation of spells to firms.

To quantify the explanatory power of person effects, I focus on two magnitudes: First, I follow Jarosiewicz & Ross (2023) and compute the p-values associated with F-statistics as given by:

$$F(q, N-k) = \frac{(R^2_{unrestricted} - R^2_{restricted})/q}{(1 - R^2_{unrestricted})/(N-k)} \tag{11}$$

[20] In later chapters, Figures 3 and 5 show that the documented attenuation biases are even more pronounced when applying the DGP from equation (8). Panel B from Table 6 provides an example that the DGP with the mean of person effects might be inadequate for testing for the presence of genuine person effects. In this case, a DGP featuring the sum of person effects might be a better choice to investigate the existence of person effects.

where q reflects the number of restrictions, N is the number of observations and k represents the number of parameters in the unrestricted model. The unrestricted model includes person effects and firm effects whereas the restricted model is limited to firm effects. Second, based on the decomposition of R^2 as presented in equation (5), I consider the component of R^2 that is attributable to person effects as given by $\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$. Hence, I compare this term between the original and the scrambled data for the DGPs described above. While critical values for the F-statistic from (11) are available to assess the joint significance of person effects, the test of scrambling the data targets the *change* in p-values associated with the F-statistic or the *change* in $\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$. While these changes cannot be compared against existing critical values, the subsequent simulation analysis shall provide insights how often the person effects are jointly significant in the original and in a scrambled sample.

Table 4 shows how the explanatory power of person effects, indicated by different expressions in the first column, changes when scrambling the person-firm spells. The analysis is carried out for four DGPs defined by equations (7), (8), (9) and (10). The simulation iterations can be stratified along 10 different boardsizes, ranging from an average of 1.3 persons per firm-year to an average boardsize of about 8, and 10 different mobility shares within each boardsize (0.05, 0.15, ..., 0.95).[21] Each boardsize-mobility combination comprises 100 different datasets from 100 iterations, yielding a total of $10 * 10 * 100 = 10,000$ iterations. Each iteration is based on an underlying dataset of 100 firms, observed over 15 years.

Column (1) presents results regarding the first DGP which is based on the sum of person effects θ_i. Person effects are significant on the 5% - level in all 10,000 iterations. Hence, the F-test identifies the presence of significant person effects in all iterations, leading to no occurrence of Type II error. The scrambling procedure, however, still delivers a p-value below 5% regarding the joint significance of person effects in 8,774 of 10,000 cases. This is a remarkable finding as it indicates significance on the 5% - level in considerably more than 5% of cases, the frequency which would be expected in case of a non-existent effect of person-level heterogeneity in the determination of the outcome variable. In the literature on manager effects for firm-level dependent variables, it is common practice to compare the number of significant results to the benchmark of how many significant results would be expected due to the significance level. Figure 1 from Schoar et al. (2023) reports significant results considerably more often than implied by the benchmark. In contrast, Jarosiewicz & Ross (2023) consider 14 different outcome variables for 3 different executive types and document that significant person effects are more frequently found in scrambled datasets than in the original data. Jarosiewicz & Ross (2023) infer that person effects reflect spurious variation. However, relating these divergent findings to the first column of Table 4 shows that many scrambling iterations where person effects remain significant are also consistent with an underlying DGP with two-way fixed-effects. Hence, considering only the frequency of significant results indicates that the scrambling procedure by Jarosiewicz & Ross (2023) would falsely characterise person effects as random noise in those 8,774 of 10,000 iterations while they are, in fact, part of the underlying DGP.

[21] Figure A2 illustrates the 100 boardsize-mobility combinations used for the simulation.

Table 4: Testing the scrambling procedure for different DGPs.

	(1) DGP 1: Sum of θ_i	(2) DGP 2: Mean of θ_i	(3) DGP 3: No θ_i	(4) DGP 4: Error term
All iterations	10,000	10,000	10,000	10,000
p-value for person effects (original) < 0.05	10,000	4,174	2,489	24
p-value for person effects (scrambled) < 0.05	8,774	2,655	2,305	23
p-value (orig.) $= 0$, p-value (scram.) $= 0$	6,837	2,252	1,988	0
p-value (orig.) < 0.05, p-value (orig.) - p-value (scram.) > 0	0	34	53	4
p-value (orig.) < 0.05, p-value (orig.) - p-value (scram.) < 0	3,163	1,887	447	20
p-value (orig.) < 0.05, p-value (orig.) - p-value (scram.) < -0.05	1,226	1,607	325	14
p-value (orig.) < 0.05, p-value (orig.) - p-value (scram.) < -0.10	958	1,532	301	11
$\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (orig.) - $\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (scram.) < 0	936	2,521	4,847	5,014
$\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (orig.) - $\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (scram.) > 0	9,064	7,479	5,153	4,986
$\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (orig.) - $\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (scram.) > 0.05	5,865	2,307	1,026	258
$\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (orig.) - $\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (scram.) > 0.10	2,054	1,130	442	34
$\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (orig.) - $\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (scram.) > 0.15	871	615	217	7

Note: The table presents simulated data based on 10,000 iterations for four different DGPs. The iterations stem from 10 different boardsizes (number of persons per firm-year), 10 different mobility shares within each boardsize (0.05, 0.15, ..., 0.95) and 100 iterations per boardsize-mobility combination, yielding a total of $10 * 10 * 100 = 10,000$ iterations. Each iteration is based on an underlying dataset of 100 firms, observed over 15 years. All p-values are based on F-tests, as computed by equation (11), on the joint significance of person effects estimated via the AKM approach. The expression $\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ describes the component of R^2 attributable to estimated person effects based on equation (5). The DGPs are given by
(1) $y_{j(i)t} = (\sum_{i=1}^{n_{jt}} \theta_{i \in (jt)}) + \psi_j + u_{jt}$, equation (7)
(2) $y_{j(i)t} = (\frac{1}{n_{jt}} * \sum_{i=1}^{n_{jt}} \theta_{i \in (jt)}) + \psi_j + u_{jt}$, equation (8)
(3) $y_{j(i)t} = \psi_j + u_{jt}$, equation (9)
(4) $y_{j(i)t} = u_{jt}$, equation (10)
where person effects θ_i, firm effects ψ_j and error terms u_{jt} are drawn from independent standard normal distributions.

A related indicator for the presence of person effects consists in the comparison between p-values for the significance of person effects between original and scrambled data. Due to the high frequency of iterations with significant person effects, I report the number of iterations where p-values equal zero, rounded to four digits, in both the original and in the scrambled sample. For the DGP that includes the sum of person effects, this applies to 6,837 iterations. For the DGP based on equation (7), the scrambling test produces misleading results in more than two thirds of the iterations under considerations. Further, in less than one third of iterations (3,163 of 10,000 iterations), the p-value increases (indicating a lower level of significance) when scrambling.

Given that the joint significance of person effects is not always removed by scrambling, I consider a further sensitivity check of the scrambling test. In this case, I compare the components of R^2 attributable to person effects between the original and the scrambled sample. In 9,064 of 10,000 cases, person effects explain a larger share of the variation in the dependent variable in the original sample as compared to the scrambled sample. In more than half of the cases (5,865 of 9,064 iterations), the discrepancy in this measure between the original and the scrambled sample is larger than 0.05. Overall, this measure represents an informative indicator whether the model assumption of additive person effects is supported by the data. It might be used in addition to the comparison between p-values for joint significance of person effects.

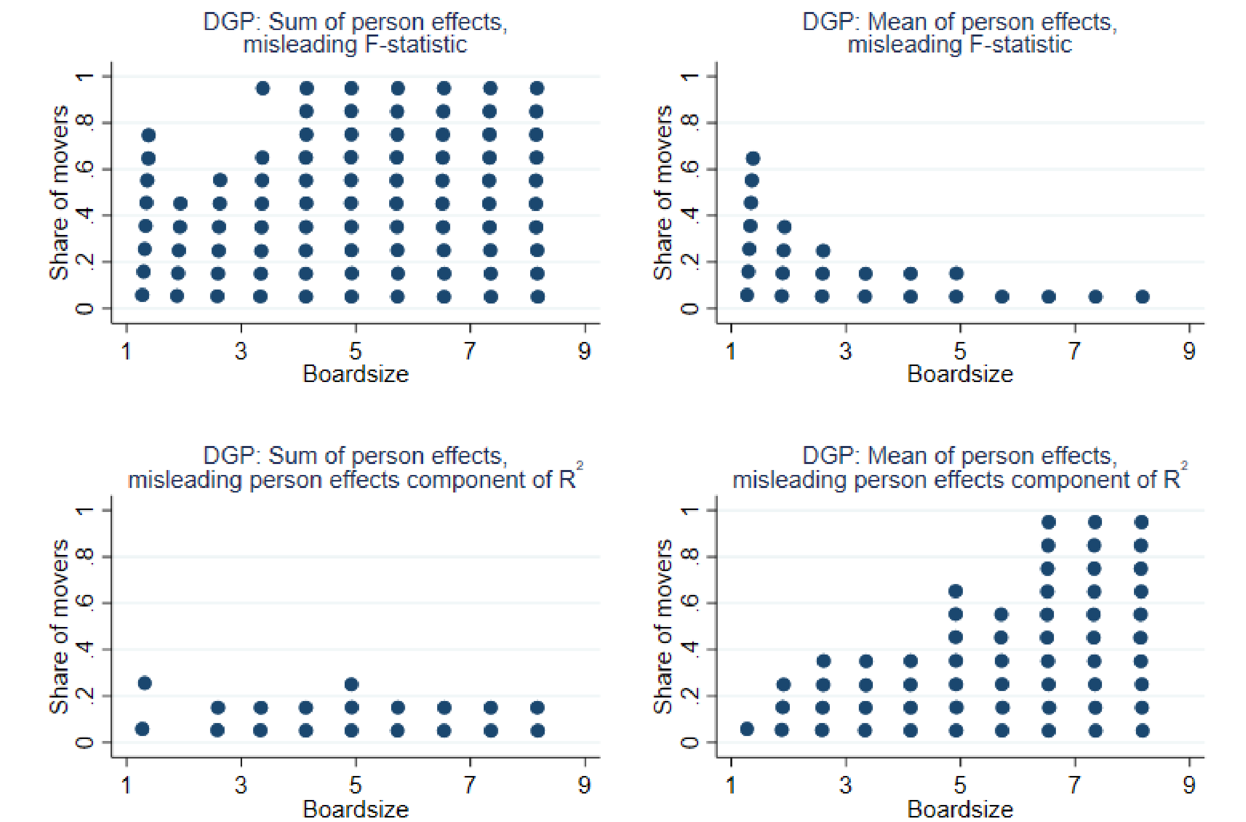
The second column of Table 4 assumes that the average of person effects θ_i enters the DGP for the dependent variable. As they enter into DGP from equation (8) with the factor $\frac{1}{n_j t}$, they lead to a more compressed distribution of dependent variables and, hence, to a more compressed distribution of person effects. This contributes to considerably less iterations with person effects jointly different from zero: Person effects are significant on the 5% level in only 4,174 of 10,000 iterations. Hence, the F-test only rarely recognises that person effects are part of the underlying DGP. Regarding the scrambling procedure, F-tests indicate significance in 2,655 iterations. Consequently, the number of iterations with p-values rounded to zero in the original and the scrambled sample (2,252) is lower than in the first column. Conditioning on p-values in the original sample below 0.05, a comparison between p-values from the original and the scrambled sample indicates the presence of person effects in 1,887 iterations, i.e. in less then half of the 4,174 cases. This shows that also under the DGP specified in equation (8), the scrambling procedure from Jarosiewicz & Ross (2023) can produce misleading conclusions regarding the existence of person effects.

When excluding person effects from the DGP for the dependent variable in the third column, joint significance of person effects is indicated in 2,489 of 10,000 iterations. Scrambling hardly changes the magnitude of this figure (2,305) as person effects are not part of the underlying DGP so that their random re-allocation to other firms does not systematically change the explanatory power of estimated person effects. However, significance of estimated person effects in more than 20 percent of iterations in spite of a DGP without person effects requires an explanation of the mechanism. First, further disentangling the 2,489 iterations by the share of movers shows (untabulated) that all 1,000 iterations with a mover share of five percent, 582 iterations with a mover share of 15 percent and 327 iterations with a mover share of 25 percent make up for the vast majority of these iterations. Hence, primarily low-mobility iterations generate the spurious result that person effects are jointly significant on the 5 % - level although they are not part of the underlying DGP. Second, it is helpful to compare the number of 2,489 iterations which spuriously imply significance of person effects to the last column where the dependent variable is based on a DGP that only features random shocks. In that case, person effects are significant on the 5 % - level in less than one percent of cases (24 of 10,000 iterations). Given that the only difference between columns (3) and (4) consists in the firm effects included in the DGP for column (3), those firm effects must be responsible for the 2,489 iterations with significant person effects. Combining the observations that (i) low-mobility iterations and (ii) the presence of firm effects lead to 2,489 of 10,000 iterations with significant person effects suggests the following: If mobility is low and firm effects are part of the underlying DGP, estimated person effects might absorb variation that stems from the firm effects in the DGP. All non-movers are de-facto assigned person effects based on the firm effects. If mobility is low, the share of non-movers who are de-facto assigned person effects is high. Thus, the estimation procedure is unable to clearly differentiate between spurious person effects (which are artefacts from firm effects) and genuine person effects, leading to the high number of 2,489 iterations with significant person effects. Further, this artefact of spuriously significant person effects creates a challenge for applied researchers: If the researcher finds p-values close to zero in both the original and the scrambled sample (third row of Table 4), different scenarios are possible. Either genuine person effects might be present (first column with 6,837 iterations or second column with 2,252 iterations) or the significant person effects might result from a low-mobility sample with a dependent variable with firm-level heterogeneity (1,988 iterations from third column) which the estimation procedure spuriously assigns to person effects.

Overall, Table 4 shows that the scrambling of person-firm spells suggested by Jarosiewicz & Ross (2023) often reduces the explanatory power of person effects when person effects are part of the DGP. However, there are different scenarios where the scrambling test might mislead the researcher regarding the existence of time-invariant person effects.

To provide further insights under which scenarios the scrambling procedure might result in misleading conclusions, consider Figure 1. It represents subsamples of all boardsize-mobility combinations from Figure A2 and for both DGPs from equations (7) and (8). For each such boardsize-mobility combination, Figure 1 depicts whether there are more than 20 percent of iterations where the scrambling procedure does not reduce the explanatory power of person effects and, thus, suggests that person effects reflect random noise. In the upper two graphs of Figure 1, the data points indicate that scrambling the person-firm-spells still yields p-values for the significance of person effects of zero (rounded to four digits). Consistent with Table 4, this phenomenon is more pronounced for the DGP featuring the sum of person effects. Still, Figure 1 illustrates that the two DGPs have in common that person effects from both the original and the scrambled sample are often significant for low-mobility iterations. Further, the lower two graphs represent iterations where the component of R^2 attributable to person effects increases when scrambling. Again, the diagrams show that this result occurs particularly for low-mobility iterations. The intuition of this result resembles the explanations of column (3) from Table 4: If mobility is low, there are many non-movers in the sample. Scrambling the non-movers between firms leads only to changes in the explanatory power of the model if the years of observation between the persons are different. When scrambling non-movers who are all observed in the years 2005 − 2009, for instance, the scrambling procedure is meaningless (abstracting from observable characteristics). Scrambling movers has more influence on the structure of the dataset than scrambling non-movers: If assigning movers to new firms, two effects are at play: First, the number of persons observed in a given firm-year changes (as is the case for non-movers). Second, the structure of connections to other firms changes (which is not the case when scrambling only non-movers). Therefore, scrambling a dataset with a low mobility share has less impact as the structure of connections between firms changes only for the low share of movers.

Figure 1: Examples for misleading scrambling iterations



Note: The diagrams present selected boardsize-mobility combinations from the simulations shown in Table 4 where scrambling the person-firm-spells might lead to misleading results regarding the presence of person effects. *Misleading* means that person effects are part of the underlying DGP and that the scrambling procedure does not reduce the explanatory power of person effects in more than 20 percent of iterations within each boardsize-mobility combination. The upper two graphs represent iterations where the p-value associated with the F-statistic for joint significance of person effects is zero (rounded to four digits) both in the original sample and in the scrambled sample (p-value (orig.) $= 0$, p-value (scram.) $= 0$ in Table 4). The lower two graphs refer to iterations where the component of R^2 attributable to person effects increases when scrambling ($\frac{Cov(y_{j(i)t}, \hat{\theta}_i)}{Var(y_{j(i)t})}$ (orig.) - $\frac{Cov(y_{j(i)t}, \hat{\theta}_i)}{Var(y_{j(i)t})}$ (scram.) < 0 in Table 4). The DGPs are given by
(1) Sum of person effects: $y_{j(i)t} = (\sum_{i=1}^{n_{jt}} \theta_{i \in (jt)}) + \psi_j + u_{jt}$, equation (7)
(2) Mean of person effects: $y_{j(i)t} = (\frac{1}{n_{jt}} * \sum_{i=1}^{n_{jt}} \theta_{i \in (jt)}) + \psi_j + u_{jt}$, equation (8)

In sum, structural parameters of the dataset like boardsize and the share of movers determine whether the scrambling procedure from Jarosiewicz & Ross (2023) is able to detect or reject the presence of genuine person effects. Further, the common practice in the literature of comparing the number of significant results to the benchmark implied by the significance level (as in Figure 1 from Schoar et al. (2023)) can be misleading as this type of analysis ignores that structural parameters of the data determine how often significant person effects can be expected.[22]

[22] A frequently mentioned advantage of the AKM approach from Abowd et al. (1999) compared to the MDV approach from Bertrand & Schoar (2003) consists in the ability of the AKM approach to estimate person effects also for non-movers as long as they are part of a set connected by movers. However, this can lead the researcher to analyse low-mobility samples. For instance, mobility both in the real-world sample of publicly listed German firms from the present study (mover share of 11 percent) as well as the sample from Francis et al. (2020) (mover share of 15 percent) are characterised by low mobility.

When carrying out the scrambling test from Jarosiewicz & Ross (2023), it is advisable to run multiple iterations for the dataset under consideration. Although parameters of the dataset like boardsize and the mobility share are important determinants, there might be also variation within each boardsize-mobility combination regarding the power of the scrambling test. Figure A3 provides examples for boardsize-mobility combinations with a large range of p-values of the F-test for joint significance of person effects in the scrambled samples. If scrambling the dataset only once, one might find significant person effects due to chance although person effects would be insignificant in the majority of further iterations.

5.3 Testing strategy for the presence of person effects in a given dataset

In this section, I recommend a strategy to test for the presence of idiosyncratic person effects. It combines the scrambling procedure recommended by Jarosiewicz & Ross (2023) with the findings from Section 5.2. The suggested strategy therefore takes into account that the sensitivity of the scrambling test depends on the structure of the dataset, indicating that the scrambling test can perform poorly for low mobility-datasets (as shown by Figure 1). Therefore, I suggest the following strategy which is illustrated in the remainder of this section: First, scramble the person-firm-spells to other firms and compute the change in explanatory power of person effects (either based on the joint significance for person effects or on the component of R^2 attributable to person effects) in response to scrambling. Second, compare this change to the change in explanatory power when the dependent variable is replaced by simulated data with a DGP that includes idiosyncratic person effects.

5.3.1 Step I: Change in explanatory power when scrambling the actual dataset

In the first step, I apply the scrambling procedure recommended by Jarosiewicz & Ross (2023) to the dataset of 156 publicly listed German firms as described in Section 3. I consider three commonly used measures for firm performance: ROA, ROE and stock return. Similar to Section 5.2, I consider F-tests on the joint significance of person effects (Table 5) and the component of R^2 attributable to person effects (Figure 2) as measures for the explanatory power of person effects. Table 5 considers the joint significance of person effects in the original and in the scrambled sample. Hence, it compares p-values of the F-tests on the null hypothesis that person effects have jointly no influence on the respective firm performance measure. For both ROA and ROE, person effects are significant both in the original and in the scrambled sample, indicated by p-values of zero (rounded to four digits). This result is in line with previous methodological criticism: Similar to the analysis in Fee et al. (2013) and Jarosiewicz & Ross (2023), scrambling in the present dataset for ROA and ROE does not lead to a decline in the explanatory power of person effects. This supports their interpretation that person effects only reflect spurious variation as they maintain their statistical significance when person-firm spells are randomly matched to other firms. However, the simulations in Table 4 indicate multiple possible explanations for this finding: In the presence of genuine person effects, the estimated person effects can maintain their significance after scrambling (columns (1) and (2) of Table 4). Alternatively, person effects might reflect spurious variation and still, they are found to be significant both in the original sample and in the scrambled sample (column (3) of Table 4). Hence, based on Table 5, no definite conclusion regarding the existence of idiosyncratic person effects for ROA and ROE can be drawn. In contrast to ROA and ROE, the person effects in the model for stock return are insignificant in both the original and the scrambled sample. Given the concern of Jarosiewicz & Ross (2023) that many fixed-effects often lead to an overparametrisation and to significance of spurious predictors, the insignificant person effects in the case of stock return are a strong indicator that person effects should not be included in the respective regression model.

Table 5: Scrambling for different dependent variables.

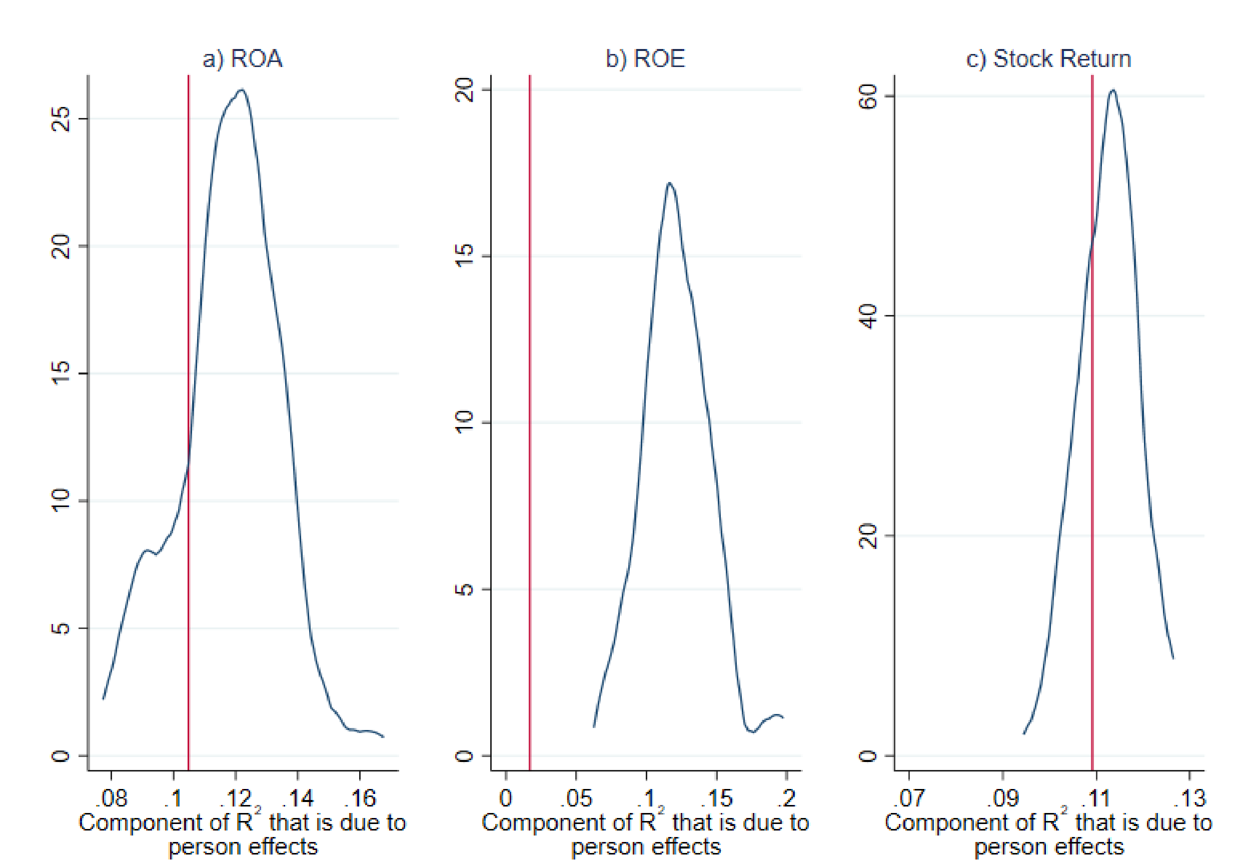
	(1) ROA	(2) ROE	(3) Stock Return
p-value for person effects (original)	0	0	1
p-value for person effects (scrambled)	0	0	1

Note: The p-values (rounded to four digits) are based on F-tests, as computed by equation (11), on the joint significance of person effects estimated via the AKM approach. The table is based on the sample of 4,076 supervisory directors from 156 publicly listed German firms described in Section 3, denoted the original sample, and 100 iterations of the scrambling procedure.

Figure 2 extends the analysis on whether director effects should be estimated for ROA, ROE and stock return in the present dataset. As shown by Table 4, not only the F-test (as in Table 5), but also the component of explained variation attributable to person effects indicates whether person effects might be part of the underlying DGP and, hence, whether the estimation of a two-way fixed-effects model is meaningful.

For each of the three dependent variables, Figure 2 provides the actual explanatory power of person effects (red line) as measured by the component of R^2 that is attributable to person effects, as computed by equation (5). Further, the diagrams contain the densities of the explanatory power of person effects over 100 iterations of the scrambling procedure (blue line). Regarding ROA, the scrambling procedure leads to an increase in the explanatory power of person effects in the majority of cases. Hence, the variable ROA does not yield clear indication that its underlying DGP is consistent with a two-way fixed-effects model. Put differently, ROA does not pass the scrambling test by Jarosiewicz & Ross (2023). Consequently, the ROA-based person effects might reflect random noise rather than actual influence of the directors. This pattern is even more pronounced for ROE. Person effects account for 2 percent of the variation in the dependent variable. However, when scrambling the person-firm spells, this share increases for all 100 iterations. This should be interpreted as strong evidence that person effects are not part of the underlying DGP for ROE and, thus, should not be considered as proxies for director style regarding ROE. Panel C describes the results for stock return. In the original sample, the component of R^2 attributable to person amounts to 11 percent. This share increases in more than half of the iterations when scrambling the person-firm spells. In combination with the finding from Table 5 that person effects are jointly insignificant, stock return appears not to be consistent with a two-way fixed-effects DGP. As an interim conclusion, neither of the variables ROA, ROE and stock return seems to be consistent with an underlying two-way fixed-effects DGP with non-zero person effects.

Figure 2: Density plots: component of R^2 that is due to person effects



Note: The three diagrams plot the distribution of the component of R^2 that is attributable to person effects, as computed by equation (5). The graphs are based on the sample of 4,076 supervisory directors from 156 publicly listed German firms described in Section 3. The red line plots the explanatory power of person effects as measured in the original sample for the respective dependent variable. The blue line plots the densities of the explanatory power of person effects, each computed over 100 iterations of the scrambling procedure.

5.3.2 Step II: Change in explanatory power when scrambling the dataset with a simulated dependent variable

In the second step, the results from the first step as described in Section 5.3.1 are compared against a benchmark to assess which sensitivity of the scrambling test can be expected for the given structure of the data. Therefore, I follow studies like Fitza (2014), Quigley & Graffin (2017) and Fitza (2017) in replacing the dependent variable of interest with a DGP that is compatible with idiosyncratic person effects. I analyse how the component of R^2 attributable to person effects reacts to scrambling for the given structure of data. This improves upon the analysis in Table 4 which features different simulations on the sensitivity of the scrambling test but does not take the structure of the present dataset into account. As Table 5 yields no clear pattern regarding the change in p-values when scrambling, I focus the subsequent analysis on explanatory power of person effects as measured by the component of R^2 attributable to person effects. While Figure 2 suggests that scrambling does not reduce the explanatory power of person effects, I analyse how the explanatory power of person effects changes in response to scrambling if person effects are part of the underlying DGP.

The comparison of explanatory power of person effects between the original data and scrambled data is presented in Table 6. Panels A, B and C each represent 100 iterations of simulations based on the DGPs specified in equations (7), (8) and (9). While Panels A and B feature the inclusion of person effects in the DGP, Panel C uses a DGP without idiosyncratic person effects. This approach of substituting an existing variable with simulated data with no explanatory power of interest, like explanatory power of person effects, is common in the literature on person effects in models of firm-level dependent variables (e.g. in variance decomposition analyses by Fitza (2014), Quigley & Graffin (2017), Fitza (2017) or Krause et al. (2019)). In Table 6, the components of R^2 attributable to person effects are presented according to the descriptive statistics mean, standard deviation, minimum and maximum. While the explanatory power of person effects $\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ is of primary interest, I also report the component of R^2 attributable to firm effects $\frac{Cov(y_{j(i)t},\hat{\psi}_j)}{Var(y_{j(i)t})}$ to demonstrate that, on average, firm effects do not absorb the explanatory power from person effects.[23] First, it is instructive to compare the means of $\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ from the original data across Panels A, B and C of Table 6. It can be clearly seen that this component of R^2 attributable to person effects is always the highest in Panel A. While its mean amounts to 0.19 for the DGP with a two-way fixed-effects model with the sum of person effects in Panel A, the magnitudes from Panel B and Panel C amount to 0.06 and 0.05. Consistent with Table 4, the explanatory power of person effects is more pronounced for a DGP with the sum of person effects than for a DGP with the mean of person effects.

Additionally, I investigate how variably the explanatory power of person effects reacts to randomly scrambling person-firm-spells to other firms. In Panel A, the scrambling procedure reduces the share explained by person-level heterogeneity from 0.19 to 0.15. The distribution of the latter ranges from 0.10 to 0.21 as indicated by the minimum and maximum. Those figures are derived for the structure of the present dataset and a DGP that assumes the existence of two-way fixed-effects for a firm-level dependent variable. Consequently, they should be used as comparison benchmarks for the real-world variables ROA, ROE and stock return shown in Figure 2. Untabulated analyses show a decline in $\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ in 90 of 100 cases when scrambling.[24] Such a pattern cannot be found in Figure 2. Hence, Table 6 provides further evidence that neither ROA, ROE nor stock return in the real-world data of the present study are consistent with idiosyncratic styles of supervisory board directors in the underlying DGP. The results from Panel A in Table 6 further indicate that each dataset with the dependent variable of interest has to be tested individually whether it is consistent with an underlying two-way fixed-effects - DGP. Therefore, it is not possible to generalise the finding of spurious person effects from Jarosiewicz & Ross (2023) to all settings with person effects for a firm-level dependent variable. The findings from Panel A highlight that the dependent variable can indeed follow a DGP with genuine rather than spurious person effects.

Panels B and C of Table 6 are insightful regarding how the DGP for the simulated dependent variable should be specified. The DGPs with the mean of person effects (Panel B) or without

[23] Section A.3 discusses the scrambling procedure for a person-level dependent variable. In this case, the explanatory power of firm effects entirely disappears and is absorbed by the person effects when scrambling.
[24] This decline exceeds the value 0.02 in 68 iterations while it is larger than 0.05 in 31 iterations.

Table 6: Component of R^2 attributable to person effects and firm effects for firm-level dependent variables, based on equation (5). Comparison between original data and randomly scrambled data.

Panel A: The firm-level dependent variable follows a two-way fixed-effects model including the sum of person effects:
$y_{j(i)t} = (\sum_{i=1}^{n_{jt}} \theta_{i \in (jt)}) + \psi_j + u_{jt}$

	Mean	SD	Min	Max
$\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (original)	0.19	0.03	0.12	0.26
$\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (scrambled)	0.15	0.02	0.10	0.21
$\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (orig.) - $\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (scram.)	0.03	0.03	-0.04	0.10
$\frac{Cov(y_{j(i)t},\hat{\psi}_j)}{Var(y_{j(i)t})}$ (original)	0.53	0.05	0.40	0.63
$\frac{Cov(y_{j(i)t},\hat{\psi}_j)}{Var(y_{j(i)t})}$ (scrambled)	0.53	0.05	0.42	0.65
$\frac{Cov(y_{j(i)t},\hat{\psi}_j)}{Var(y_{j(i)t})}$ (orig.) - $\frac{Cov(y_{j(i)t},\hat{\psi}_j)}{Var(y_{j(i)t})}$ (scram.)	-0.00	0.03	-0.06	0.08

Panel B: The firm-level dependent variable follows a two-way fixed-effects model including the mean of person effects:
$y_{j(i)t} = (\frac{1}{n_{jt}} \sum_{i=1}^{n_{jt}} \theta_{i \in (jt)}) + \psi_j + u_{jt}$

	Mean	SD	Min	Max
$\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (original)	0.06	0.02	0.02	0.09
$\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (scrambled)	0.06	0.01	0.04	0.09
$\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (orig.) - $\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (scram.)	-0.00	0.02	-0.05	0.05
$\frac{Cov(y_{j(i)t},\hat{\psi}_j)}{Var(y_{j(i)t})}$ (original)	0.53	0.03	0.44	0.61
$\frac{Cov(y_{j(i)t},\hat{\psi}_j)}{Var(y_{j(i)t})}$ (scrambled)	0.53	0.03	0.44	0.61
$\frac{Cov(y_{j(i)t},\hat{\psi}_j)}{Var(y_{j(i)t})}$ (orig.) - $\frac{Cov(y_{j(i)t},\hat{\psi}_j)}{Var(y_{j(i)t})}$ (scram.)	0.00	0.02	-0.04	0.06

Panel C: The firm-level dependent variable follows a one-way fixed -effects model:
$y_{j(i)t} = \psi_j + u_{jt}$

	Mean	SD	Min	Max
$\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (original)	0.05	0.01	0.02	0.09
$\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (scrambled)	0.06	0.01	0.03	0.09
$\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (orig.) - $\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (scram.)	-0.01	0.02	-0.05	0.04
$\frac{Cov(y_{j(i)t},\hat{\psi}_j)}{Var(y_{j(i)t})}$ (original)	0.54	0.03	0.47	0.63
$\frac{Cov(y_{j(i)t},\hat{\psi}_j)}{Var(y_{j(i)t})}$ (scrambled)	0.53	0.04	0.46	0.63
$\frac{Cov(y_{j(i)t},\hat{\psi}_j)}{Var(y_{j(i)t})}$ (orig.) - $\frac{Cov(y_{j(i)t},\hat{\psi}_j)}{Var(y_{j(i)t})}$ (scram.)	0.01	0.02	-0.04	0.06

Note: The sample presents descriptive statistics for the share in R^2 attributable to person effects and firm effects, as defined by equation (5). The data structure is based on the sample of 4,076 supervisory directors from 156 publicly listed German firms described in Section 3 to maintain properties of a real-world sample. The scrambling procedure randomly assigns person spells to other firms. The real-world dependent variable is replaced by simulated data according to three different DGPs as defined in the panel headers for panels A, B and C. Each panel refers to 100 iterations with different simulated dependent variables with one scrambling iteration for each of these 100 datasets. All components from the DGP, i.e. θ_i, ψ_j and u_{jt}, are drawn from independent standard normal distributions.

person effects (Panel C) have different implications for the scrambling test than Panel A. Under both DGPs, the average components of R^2 attributable to person effects are unchanged (Panel B) or even increase slightly (Panel C).[25] The finding from Panel B is consistent with Figure 1: When assuming that the mean of person effects is part of the underlying DGP, there are many boardsize-mobility-combinations where scrambling does not lead to a decline in the component of R^2 attributable to person effects. Thus, in the structure of the present dataset, a DGP with the mean of person effects (instead of the sum of person effects) does not induce the scrambling test to detect idiosyncratic person effects. Hence, for the structure of the present dataset, specifying a DGP with the sum of person effects should be preferred to a DGP with the mean of person effects. Further, the absence of person-level heterogeneity in Panel C of Table 6 induces the scrambling not to reduce the explanatory power of person effects. This is the scenario sketched in Jarosiewicz & Ross (2023): When the underlying DGP for the dependent variable does not account for person effects, the random allocation of firm-person spells to other firms does not reduce the explanatory power of person effects. Thus, in case of Panel C, it is appropriate to interpret person effects as random noise.

In all three Panels A, B and C, the firm effects make up for more than half of the explained variation as the component of R^2 attributable to firm effects exceeds 0.50. This share appears to be invariant to the underlying DGP and to the application of the scrambling procedure. This is due to the draws of firm effects and person effects from independent distributions and due to the unchanged patterns of firm-level outcomes in each firm as the scrambling procedure changes only the person-composition at each firm. Put differently, the firm component ψ_j specified in the DGP has similar explanatory power in both the original and the scrambled sample.

Overall, the scrambling check suggested by Jarosiewicz & Ross (2023) (and similarly used by Bertrand & Schoar (2003), Dyreng et al. (2010), Fee et al. (2013) and Herpfer (2021)) has pros and cons: On the one hand, it provides an important validation tool when deciding on whether a two-way fixed-effects model is appropriate for analysing a firm-level dependent variable in a given dataset. Hence, researchers do not have to discard the use of two-way fixed-effects models for firm-level variables *per se*. Instead, they should address the criticism that person effects might reflect spurious estimates by showing that the scrambling procedure removes their explanatory power. On the other hand, researchers should also take into account that the effectiveness of scrambling depends on the structure of the underlying dataset. More specifically, they have to be aware of the shortcomings of the scrambling test as shown in Table 4 and in Figure 1. Especially datasets with low mobility can mislead the researcher when testing for the presence of person effects. If scrambling the spells in a dataset with low mobility, person effects can maintain their explanatory power even if the dependent variable follows a DGP with person effects.

The suggested strategy to test for genuine person effects attempts to overcome the limitations of the scrambling check from Jarosiewicz & Ross (2023). Once this strategy has been carried out successfully, researchers can argue that the underlying DGP satisfies a two-way fixed-effects

[25]The component of R^2 attributable to person effects reduces in response to scrambling in 47 of 100 iterations of Panel B and in 32 of 100 iterations of Panel C.

model and, thus, includes person effects that are different from zero. In this case, it is legitimate to estimate a two-way fixed-effects model and interpret the estimated person effects as idiosyncratic styles. Despite being a recent contribution, Jarosiewicz & Ross (2023) are already cited by other studies which refer especially to the recommendation of scrambling the person-firm spells to check for spurious results (Coles & Li, 2020; Dang et al., 2021). It should be considered good practice to follow this procedure and apply the scrambling as well. However, comparing the resulting change in explanatory power to a benchmark from a simulated dependent variable is essential. Otherwise, the structure of the dataset can drive the result of the scrambling test, making it impossible for the applied researcher to draw conclusions on the existence of idiosyncratic person effects.

5.4 Properties of estimated person effects

Applying the two-way fixed-effects model to a firm-level dependent variable has implications for the estimated person effects. To begin with, one should note that a firm-level dependent variable $y_{j(i)t}$ is a special case of a person-level dependent variable y_{it} such as wages. The firm-level dependent variable imposes the restriction that the dependent variable is equal between all persons observed in one firm-year. One implication from this restriction is that persons obtain the same person effects if they have the same spell pattern, abstracting from individual-level covariates.

To demonstrate this, consider the explanations of how the two-way fixed-effects model is estimated in Section 4.1. Separate identification of the person effects θ_i and firm effects ψ_j requires a demeaning procedure first. As datasets typically include considerably more persons than firms, it is computationally more efficient to eliminate the person-identifiers first to obtain firm effect estimates and only then to recover person effects. However, for illustration, consider the case where the firm identifiers are eliminated first. In this case, equation (1) does not lead to equation (2) but to the following expression, abstracting from covariates and time effects:

$$y_{j(i)t} - \overline{y}_j = (\theta_i - \overline{\theta}_{ij}) + (u_{ijt} - \overline{u}_j) \tag{12}$$

For two persons that are observed in the same firm(s) during the same year(s), this equation always produces the same estimate $\hat{\theta}_i$. This is due to the identical outcome variable $y_{j(i)t}$ and, according to affiliations with the same firms, due to identical firm-level averages of the dependent variable $\overline{y}_j$. Further, both the idiosyncratic error term u_{ijt} and the firm-average error term $\overline{u}_j$ would be invariant between both persons.

To analyse the consequences of a firm-year-level dependent variable with the aforementioned implications, I carry out simulations with 10,000 iterations, stratified by 10 boardsizes (ranging from 1.3 to about 8), 10 mobility shares (0.05, 0.15, ..., 0.25) per boardsize and 100 iterations per boardsize-mobility combinations. Each iteration features a dataset of 100 firms observed over 15 years. The DGPs are given by equations (7) and (8) where person effects θ_i, firm effects ψ_j and error terms u_{jt} are drawn from independent standard normal distributions. The result of these 10,000 iterations is shown in Figure 3. The left graph is based on the DGP where the sum of person effects enters the model whereas the right graph features a DGP with the average of person effects. Both panels of this figure aggregate the simulation results via binned scatterplots of the DGP-based true person effect θ_i against the difference between true person effects and estimated person effects.

The binned scatterplots from both DGPs in Figure 3 indicate a positive relation between true person effects and the difference plotted on the vertical axis. High values of θ_i coincide, on average, with underestimated person effects while low person effects are overestimated. This points to attenuation bias in the estimation of person-level heterogeneity for a firm-level dependent variable, indicating that estimated person effects are biased towards zero. The finding of attenuation bias holds although the underlying DGP assumes the two-way fixed-effects model to be true. However, this pattern is not observed when we have a dependent variable at the person-level: Figure A5 considers person effects derived from a person-level dependent variable. It does not feature systematic deviations between true and estimated person effects, resulting in a flat line approximating the relation between person effects and their estimation error. Hence, the finding of attenuation bias does not apply to the general class of two-way fixed-effects models but to those applications with a firm-year-level dependent variable.

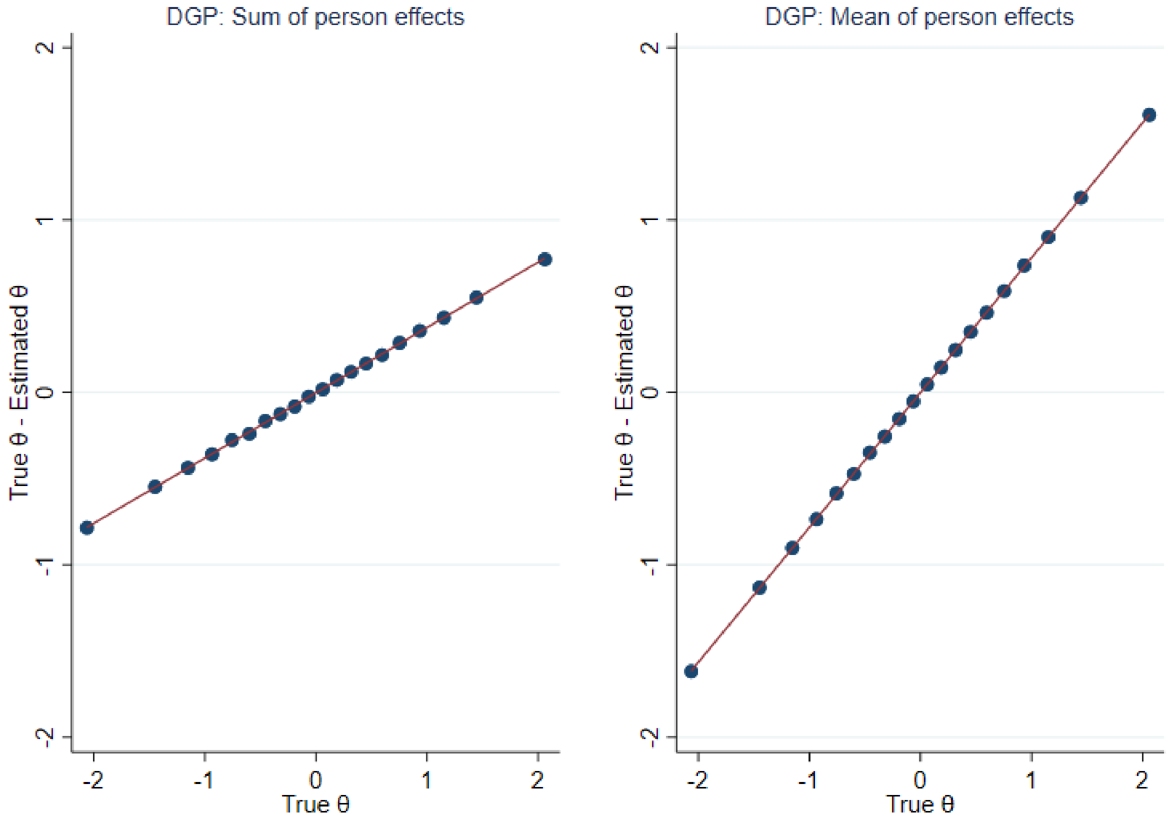
To provide more insights on determinants for the attenuation bias, Figure 4 plots its average magnitude for different boardsizes. The two graphs on the left refer to a DGP with the sum of person effects while the two graphs on the right are based on a DGP that features the average person effects per firm-year. Further, I stratify each of the 10,000 iterations per DGP by negative and positive true person effects as otherwise, positive and negative attenuation biases would neutralise each other. For all four graphs within Figure 4, the attenuation bias as indicated by the vertical axis is smallest in absolute terms for small boardsizes.[26] For instance, the upper right graph of Figure 4 shows an average attenuation bias of approximately 0.32 for an average boardsize of 1.3 while the bias amounts to 0.74 for a boardsize of 8, on average. For both DGPs, the magnitude of attenuation bias appears to converge towards an asymptote. Overall, the dependence of attenuation bias on the boardsize exists in all four graphs while it is particularly pronounced for DGPs based on the mean of person effects (two graphs on the right). Additional analyses in Figure A6 show that attenuation bias exists for all shares of movers.[27]

The finding of attenuation bias in the estimation of person effects can be explained by the structure of the estimation model: As outlined in the context of equation (12), two persons obtain the same estimated person effect if they have the same spell structure, i.e. if they are observed in the same firms during the same years. As the dependent variable $y_{j(i)t}$ is invariant

[26] This finding is supported by Figure A7 which plots the correlation of true person effects with the difference between true and estimated person effects for different boardsizes. Consistent with Figure 3, this correlation is positive. Additionally, this correlation increases in boardsize until it reaches an asymptote (for the DGP with the mean of person effects) or slightly declines (for the DGP with the sum of person effects).

[27] For the DGP including the sum of person effects in Figure A6, attenuation bias first increases in absolute terms in the share of movers. However, this relation reverses at a mover share of approximately 0.35 when increasing mover shares coincide with a decline in attenuation bias in absolute terms. For the DGP featuring the mean of person effects in Figure A6, the magnitude of attenuation bias is almost constant across the different mover shares.

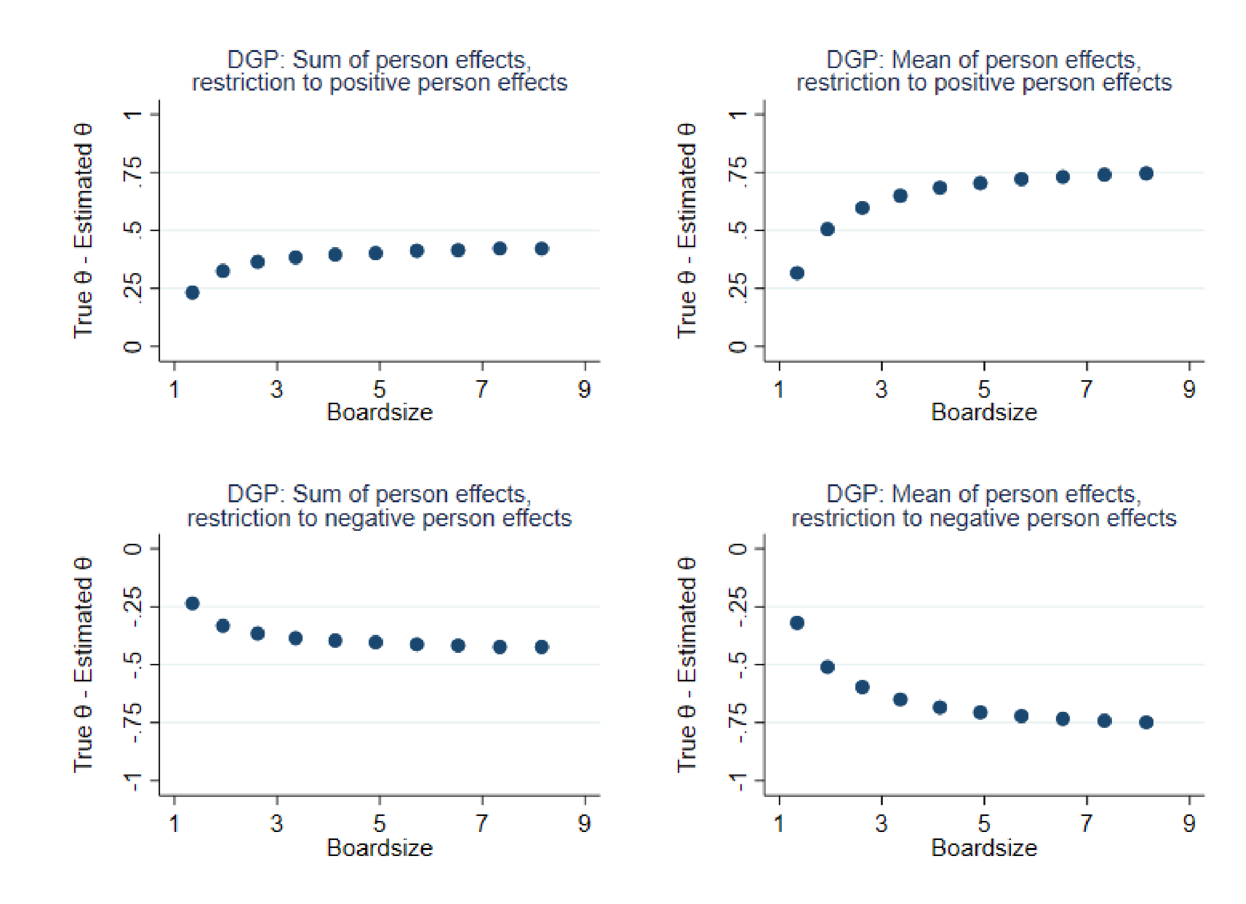
Figure 3: Estimation error in person effects by true person effects



Note: The graphs show binned scatterplots of true person effects (x-axis) and the difference between true person effects and estimated person effects (y-axis), generated by simulations. They are based on 10,000 iterations, stratified by 10 boardsizes and 10 mobility shares and 100 iterations per boardsize-mobility combination. The underlying datasets generated for each iteration consist of 100 firms, observed over 15 years. The DGPs are given by equations (7) (left graph) and (8) (right graph) where person effects θ_i, firm effects ψ_j and error terms u_{jt} are drawn from independent standard normal distributions.

per firm-year, the two-way fixed-effects model is unable to differentiate high-ability and low-ability persons (measured by person effects θ_i) as long as both persons have the same spell-structure and, hence, the same identifying variation. This provides an intuitive interpretation of attenuation bias: Assuming the draws from the ability distribution of persons affiliated with one firm are independent and identically distributed, each person faces the same expected ability distribution among coworkers, regardless of the own ability. Hence, a high value of θ_i is expected to coincide with, on average, lower person effects of coworkers. Yet, the ability of all persons observed in one firm-year contributes to the generation of the dependent variable as specified by the DGP. Thus, θ_i of high-ability persons is identified by firm-years where coworkers with lower person effects contribute to the dependent variable. This generates attenuation bias as the two-way fixed-effects model is not able to fully recover the high person effects specified in the DGP. Note that this reasoning for the existence of attenuation bias in case of firm-level dependent variables applies to both DGPs as specified in equations (7) and (8). The reason that attenuation bias is stronger for the DGP from equation (8) which contains the average of person effects (right panels in both Figures 3 and 4) is the multiplication with factor $\frac{1}{n_{jt}}$ in equation (8). This compresses the distribution of the dependent variables and, hence, of the estimated

Figure 4: Estimation error in person effects by boardsize



Note: The graphs plot the difference between true person effects and estimated person effects (y-axis) for different boardsizes (x-axis). The data is based on simulations with 10,000 iterations for each of the two DGPs, stratified by 10 boardsizes, 10 mover shares and 100 iterations per boardsize-mobility combination. The underlying datasets generated for each iteration consist of 100 firms, observed over 15 years. The DGPs are given by equations (7) (two graphs on the left) and (8) (two graphs on the right) where person effects θ_i, firm effects ψ_j and error terms u_{jt} are drawn from independent standard normal distributions.

person effects. Consequently, large person effects are unlikely to be recovered when assuming the mean of person effects to be part of the DGP and estimating a two-way fixed-effects model for the respective firm-year-level dependent variable. Relating this reasoning to the number of coworkers, it is possible to rationalise Figure 4. It shows that attenuation bias is larger in absolute terms for larger boardsizes and, equivalently, for a larger number of coworkers. Hence, the influence of the high-ability person on the generation of the dependent variable diminishes once the boardsize increases. In this case, the estimated person effect of the high-ability person is biased towards the mean of coworkers' abilities, leading to an underestimation of the person effect.

As a consequence, second stage analyses (i.e. relating θ_i to other covariates) suffer from attenuation bias. Previous literature provides several examples for this type of analysis: Table 6 from Gul et al. (2013) relates demographic and career characteristics to auditor fixed-effects, suggesting inter alia that Big N experience coincides with less auditor aggressiveness. Additionally, Table 8 from Hagendorff et al. (2021) explores whether observable manager characteristics are correlated with manager effects from bank risk measures, indicating that work experience during a banking crisis coincides with lower risk-taking. Similarly, Schoar et al. (2023) show in Table 9 that early career experiences, proxied by a dummy whether the manager entered the labour market during a recession, are negatively correlated with manager effects on risk. As illustrated by the simulations throughout Section 5.4, the magnitude of the coefficients reported in these studies might be understated in absolute value.

Table 6 from Cavaco et al. (2017) represents a similar setting which might be subject to attenuation bias. They compare affiliated and independent directors in listed French companies. By relating person effects derived from firm performance measures to directors' characteristics, Cavaco et al. (2017) attempt to draw conclusions on the selection process of board members. Their study uses a sample with large corporate boards, making the phenomenon of attenuation bias a serious concern as attenuation bias increases in boardsize until it might converge towards an asymptote (as shown in Figure 4).

However, there is an additional concern in quantile regressions of person effects on observable characteristics (like Table 6 from Cavaco et al. (2017)). In the remaining part of Section 5.4, I show that these results are likely subject to a statistical artefact at the upper and lower tail of the distribution. Cavaco et al. (2017) focus on the entire distribution of person effects from independent and affiliated directors. While they do not specify which kind of quantile regressions they estimate, Figure 1 from Cavaco et al. (2017) suggests that they posit different unconditional distributions from which individual director ability is drawn. Therefore, I base the subsequent analysis on the estimation technique of Unconditional Quantile Regressions as introduced by Firpo et al. (2009). This seminal methodological contribution shows that the partial effect of the explanatory variable on a distributional statistic of the dependent variable can be derived by modifying the dependent variable. More specifically, one has to compute the recentered influence function (RIF) in order to use it as dependent variable in the regression. For the quantile τ, RIF is defined as follows:

$$RIF(y, q_\tau) = q_\tau + \frac{(\tau - \mathbf{1}\{y \leq q_\tau\})}{f_y(q_\tau)} \tag{13}$$

In the present case, I apply Unconditional Quantile Regressions to relate person effects from a two-way fixed-effects model to observable characteristics of board members. I focus on characteristics that refer to the spell structure of observed directors, namely the appointment to multiple boards either simultaneously or in sequence.

Table 7 shows t-statistics for the coefficients from quantile regressions of the RIF of quantiles $\tau = 0.10, ..., 0.90$ of person effects. Panel A considers regressions on a dummy indicating that a director is appointed to at least two boards at the same time. For the case of corporate boards, the relevance of this analysis is mirrored by a strand of literature on *director busyness*, focusing on the implications of directors that serve on multiple boards for firm performance.[28] In contrast,

[28] The effect of busy directors on firm performance has been studied in a number of contexts with still inconclusive

Panel B focuses on a dummy indicating that a director is observed in two different boards during the estimation window which is equivalently used in Table 6 from Cavaco et al. (2017). Thus, the explanatory variable from Panel A varies within individuals while the explanatory variable from Panel B is time-invariant. The table is based on 100 scrambling iterations from the original sample presented in Section 3 where the dependent variable upon which the person effects are based was generated by the DGP from equation (7). Note that the DGP does not implement any systematic pattern along the unconditional distribution of person effects. Hence, one might expect insignificant coefficients of both *Busy* and *At least two firms*, on average.

Across both panels A and B of Table 7, it is evident that the t-statistics of the coefficients pertaining to lower quantiles exceed, on average, the critical value of 1.96, indicating significance on the 5% - level. In contrast, the t-statistics for the highest quantiles are all below -1.96, implying coefficients that are significant on the 5% - level as well. The stability of these findings is indicated by the minima and maxima of the t-statistics at the upper and lower tail of the distributions, indicating that one almost always obtains significant coefficients for the lowest and highest quantiles. If an estimation method produces certain results regardless of the underlying dataset (as indicated by the 100 random scramblings), its results should be considered spurious, resulting from a statistical artefact. Table A3 reproduces this finding for ROA (rather than a simulated variable) as dependent variable. In untabulated analysis, I further show that this finding holds for the other dependent variables ROE and stock return as well.

The emergence of this artefact is due to the spell-structure of persons that are observed in multiple firms. The quantile regression compares those persons with non-movers at a given quantile of the unconditional distribution of person effects. When placed at the upper tail of the ability distribution, multi-firm persons obtain lower person effects than single-firm persons. If observed at multiple firms, persons face, on average, a larger set of coworkers. Assuming that the abilities of all those persons are part of the DGP for the dependent variable, the persons from the upper tail of the distribution have, on average, higher ability than their coworkers. As shown in Figure 3, there is attenuation bias in the estimation of person effects, resulting in a conversion to the mean. If exposed to more firms and, thus, to more coworkers, the conversion to the mean is even more pronounced, spuriously leading to positive coefficients for low quantiles and negative coefficients for high quantiles. This is illustrated by Table 6, Panels A and B, from Cavaco et al. (2017) which includes a dummy for multi-directorships which follows a similar definition as the explanatory variable in Panel B of Table 7. It is no coincidence that for the five (Panel A) or three (Panel B) lowest deciles, the coefficients are positive and significant while the two highest deciles feature negative coefficients. Instead, those coefficients are subject to the statistical artefact described above.

results. The *reputation hypothesis* posits a positive link between busyness and firm performance, resulting from additional experience at other boards, increased networks and improved information flows (Pfeffer & Salancik, 1978; Ferris et al., 2003; Kiel & Nicholson, 2006; Masulis & Mobbs, 2011). In contrast, the *busyness hypothesis* suggests that excessive board responsibilities lead to a decline of monitoring effort and are, thus, detrimental for firm performance (Lipton & Lorsch, 1992; Fich & Shivdasani, 2006; Falato et al., 2014). There is a broad range of strategies to identify the effect of busyness, for instance via the usage of GMM approaches (Latif et al., 2020), via instrumental variables like the number of firms with a headquarter in the same country (Trinh et al., 2020) or by exploiting exogenous variation in busyness associated with mergers (Hauser, 2018).

Table 7: Spurious findings in Unconditional Quantile Regressions

Panel A: Explanatory variable: Busy				
	Mean	SD	Min	Max
t-statistic at quantile 10	5.18	1.90	1.22	10.36
t-statistic at quantile 20	4.12	2.15	-0.14	9.82
t-statistic at quantile 30	2.61	1.77	-1.25	6.90
t-statistic at quantile 40	1.25	1.62	-2.13	5.36
t-statistic at quantile 50	-0.12	1.51	-4.17	3.98
t-statistic at quantile 60	-1.42	1.48	-5.02	1.79
t-statistic at quantile 70	-2.75	1.80	-7.25	0.87
t-statistic at quantile 80	-4.11	1.95	-9.38	0.50
t-statistic at quantile 90	-5.23	1.97	-9.94	-1.10
Panel B: Explanatory variable: At least two firms				
	Mean	SD	Min	Max
t-statistic at quantile 10	5.64	1.92	0.36	10.13
t-statistic at quantile 20	4.51	1.80	0.45	10.33
t-statistic at quantile 30	3.01	1.67	-0.55	7.75
t-statistic at quantile 40	1.39	1.47	-1.55	4.75
t-statistic at quantile 50	-0.09	1.49	-3.17	3.39
t-statistic at quantile 60	-1.49	1.51	-4.56	1.83
t-statistic at quantile 70	-3.07	1.66	-7.85	0.40
t-statistic at quantile 80	-4.24	1.63	-10.07	-0.79
t-statistic at quantile 90	-5.53	1.89	-10.79	0.33

Note: The table presents t-statistics for the coefficients from Unconditional Quantile Regressions (UQR) as introduced by Firpo et al. (2009). The t-statistics refer to coefficients from regressions where the dependent variable is the recentered influence function (RIF) as defined in (13) of person effects that refer to a dependent variable generated by the DGP from equation (7). The structure of the underlying dataset is a scrambled version of the sample of 4,076 supervisory directors from 156 publicly listed German firms described in Section 3. Both panels are based on 100 such scrambling iterations with a simulated dependent variable. In Panel A, the dependent variable is regressed on *Busy*, a dummy indicating that a director serves on two supervisory boards at the same time. In Panel B, the explanatory variable is a dummy indicating that a director is observed in at least two firms in the sample, either simultaneously or in sequence.

Overall, quantile regressions using person effects derived with the two-way fixed-effects model should be treated with caution. If still relying on quantile regressions of person effects, practitioners should run simulations to show that their results do not stem from a statistical artefact. Even if a researcher does not analyse the heterogeneity with respect to variables such as director busyness (which describes the spell structure), the subgroups of interest might differ regarding their spell structure, leading again to a statistical artefact.[29]

[29] For instance, if a researcher intends to analyse the implications of independent directors by applying Unconditional Quantile Regressions with the RIF of estimated person effects as dependent variable and an independence

5.5 Correlation of person effects of different dependent variables

Multiple studies classify directors according to their styles. These studies consider datasets where different firm-level outcome variables are available. In the first step, for each of those outcome variables, person effects are computed. In the second step, the correlation between the person effects of one outcome variable and the person effects of the other outcome variable is examined. Examples include Table VII from Bertrand & Schoar (2003), Table 2 from Francis et al. (2020) and Table 5 from Schoar et al. (2023). This section reiterates this sort of analysis.[30]

To inspect properties of the correlation of estimated person effects, I use simulation analysis. For two different DGPs, I run 20,000 iterations each. Those are stratified by 20 different levels for the correlation such that $Corr(\theta_{i1}, \theta_{i2})$ = -0.95, -0.85, ..., 0.75, 0.85, 0.95, 10 different levels of boardsize illustrated in Figure A2, 10 different mobility shares per boardsize (0.05, 0.15, ..., 0.95) and 10 iterations per configuration of the above factors, yielding a total of $20*10*10*10 = 20,000$ iterations. Further, I specify a total number of 100 firms per dataset observed over 15 years. The DGPs are given by equations (7) and (8) where person effects θ_i, firm effects ψ_j and error terms u_{jt} are drawn from independent standard normal distributions.

To unveil the relation between the correlation of true person effects and estimation bias in the correlation, consider Figure 5. It shows the above specified 20 correlation levels between person effects of different dependent variables as compiled by the DGP. Each data point reflects 1,000 iterations (combinations of 10 different shares of moving directors over 10 boardsizes and 10 iterations). The graph provides clear evidence that high correlation coefficients between person effects are understated when considering the correlation between estimated person effects. On the contrary, the estimates of negative correlations are, on average, too small in absolute values. Hence, the simulation results are consistent with attenuation bias in the correlation of person effects, implying that the correlation coefficients are biased towards zero when computed over the estimated person effects. Figures A8 and A9 analyse the sensitivity of this attenuation bias to different boardsizes and mover shares. Based on the first graph in Figure A8, the attenuation bias is particularly strong for small boardsizes in the case of a DGP which features the sum of person effects. However, the second graph of Figure A8 indicates that attenuation bias is particularly pronounced for large boardsizes. Similarly, Figure A9 suggests that attenuation bias in the correlation of estimated person effects is existent regardless of mobility in the sample.

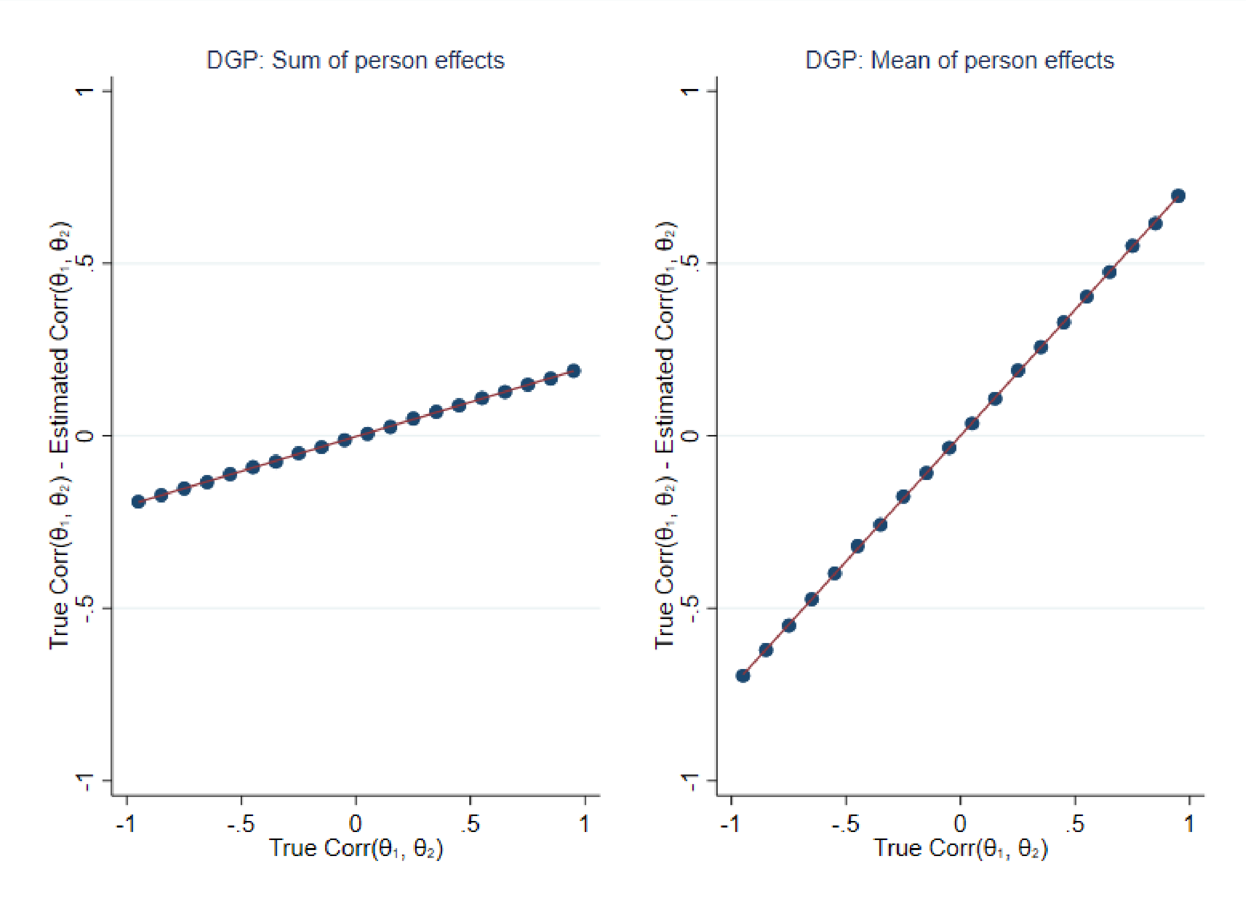
The underestimation of correlations in absolute terms when switching from correlations of true person effects to correlations of estimated person effects has implications for empirical studies: Table 5 from Schoar et al. (2023) reports a regression coefficient of 0.138 of the manager effect for systematic risk in a model for the manager effect of stock return where fixed-effects are normalised to a mean of zero and a variance of one. Given the attenuation bias demonstrated in

dummy as explanatory variable of interest, the possibility of an artefact should be taken into account. For instance, it would occur if independent directors were movers more frequently than affiliated directors. In this case, the researcher should at least show that both groups of directors have similar mover shares.

[30] Section A.5 combines analytical and simulation analysis to investigate the correlation of estimated person effects of different dependent variables.

this section, this correlation might still be understated. Further, Table 2, Panel B of Francis et al. (2020) considers the correlation of person effect of different loan terms. Although they mostly report correlations that are significantly different from zero, those might well be understated in absolute terms as demonstrated by the present simulation results.

Figure 5: Estimation error in the correlation of person effects of different dependent variables



Note: The graphs show binned scatterplots of true person effects (x-axis) and the difference between correlations of true person effects and the correlations of estimated person effects (y-axis), generated by simulations. The data considers 20 different levels of correlations $Corr(\theta_{i1}, \theta_{i2})$ = -0.95, -0.85, ..., 0.75, 0.85, 0.95 between person effects θ_i estimated from two-way fixed-effects models of two different dependent variables y_1 and y_2. Each data point reflects 1,000 iterations (combinations of 10 different shares of moving directors over 10 boardsizes and 10 iterations per boardsize-mobility combination), yielding a total of 20,000 iterations per graph. The underlying datasets generated for each iteration consist of 100 firms, observed over 15 years.
The DGPs are specified as follows
(1) Sum of person effects: $y_{j(i)t} = (\sum_{i=1}^{n_{jt}} \theta_{i \in (jt)}) + \psi_j + u_{jt}$
(2) Mean of person effects: $y_{j(i)t} = (\frac{1}{n_{jt}} * \sum_{i=1}^{n_{jt}} \theta_{i \in (jt)}) + \psi_j + u_{jt}$
Person effects θ_i, firm effects ψ_j and error terms u_{jt} as specified in the DGP are drawn from independent standard normal distributions.

6 Conclusion

The seminal study by Bertrand & Schoar (2003) has spawned a new and still growing strand of research that provides insights on the effect of individual managers or directors on firm-level variables. More recently, this type of analysis has raised serious methodological concerns, arguing that person effects from the two-way fixed-effects model with a firm-level dependent variable only reflect spurious variation. The present study sheds light on the controversial nature of this class of models.

My results confirm that the existing criticism represents a considerable limitation to the two-way fixed-effects model for firm-level outcomes. However, I do not discard its application *per se*. Instead, I show that the presence of idiosyncratic person effects has to be tested for every dataset and every dependent variable individually. I recommend a two-step strategy to verify the existence of person effects. It starts with a scrambling procedure to break the person-firm-connections in the dataset. Next, it compares the associated change in explanatory power of person effects to a benchmark where a simulated dependent variable is generated via a DGP that contains person-level heterogeneity. This takes into account that the sensitivity of the scrambling test depends on the structure of the underlying dataset. Ignoring this finding might produce misleading results regarding the presence of genuine person effects, especially when the share of movers in the sample is low. Every study that estimates a two-way fixed-effects model for a firm-level dependent variable should include this type of analysis to rule out that person effects only reflect random noise.

Further, I consider the case where person effects are part of the underlying DGP and, thus, do not capture random noise. In this case, I show that the estimation of person effects is subject to attenuation bias which increases in the number of persons observed per firm-year. Hence, when relating person effects to other observable characteristics, the correlations are understated in absolute value. Additionally, I demonstrate that the application of Unconditional Quantile Regressions to estimated person effects can generate statistical artefacts at the upper and lower tails of the distribution. Another contribution concerns analyses of the correlation of person effects of different outcome variables. I show that attenuation bias also impairs this type of investigation. Thus, actual trade-offs and complementarities between distinct firm-level variables might be understated by this analysis.

Applied future research on manager effects should take these limitations into account. Methodological demands are high as the appropriate use of a two-way fixed-effects model requires intensive scrambling and simulation checks before applying it to a new dataset. Due to the documented and previously under-researched patterns of attenuation biases, however, a researcher might be rewarded by unveiling new patterns regarding person effects.

References

Abernethy, M. A., & Wallis, M. S. (2019). Critique on the "manager effects" research and implications for management accounting research. *Journal of Management Accounting Research*, *31*(1), 3–40.

Abowd, J. M., Creecy, R. H., & Kramarz, F. (2002). Computing person and firm effects using linked longitudinal employer-employee data. *Longitudinal Employer - Household Dynamics Technical paper No. TP-2002-06 Center for Economic Studies, US Census Bureau.*

Abowd, J. M., Kramarz, F., & Margolis, D. N. (1999). High wage workers and high wage firms. *Econometrica*, *67*(2), 251–333.

Andrews, M., Gill, L., Schank, T., & Upward, R. (2008). High wage workers and low wage firms: negative assortative matching or limited mobility bias? *Journal of the Royal Statistical Society Series A: Statistics in Society*, *171*(3), 673–697.

Andrews, M., Schank, T., & Upward, R. (2006). Practical fixed-effects estimation methods for the three-way error-components model. *The Stata Journal*, *6*(4), 461–481.

Baltrunaite, A., Bovini, G., & Mocetti, S. (2023). Managerial talent and managerial practices: Are they complements? *Journal of Corporate Finance*, *79*, 102348.

Bamber, L. S., Jiang, J., & Wang, I. Y. (2010). What's my style? The influence of top managers on voluntary corporate financial disclosure. *The Accounting Review*, *85*(4), 1131–1162.

Bertrand, M., & Schoar, A. (2003). Managing with style: The effect of managers on firm policies. *The Quarterly Journal of Economics*, *118*(4), 1169–1208.

Blettner, D. P., Chaddad, F. R., & Bettis, R. A. (2012). The CEO performance effect: Statistical issues and a complex fit perspective. *Strategic Management Journal*, *33*(8), 986–999.

Bozhinov, V. (2019). Women on boards: An empirical analysis of role and effects in Germany. *Doctoral Thesis at Johannes Gutenberg-University Mainz.*

Card, D., Heining, J., & Kline, P. (2013). Workplace heterogeneity and the rise of West German wage inequality. *The Quarterly Journal of Economics*, *128*(3), 967–1015.

Cavaco, S., Crifo, P., Rebérioux, A., & Roudaut, G. (2017). Independent directors: Less informed but better selected than affiliated board members? *Journal of Corporate Finance*, *43*, 106–121.

Cho, C., Halford, J. T., Hsu, S., & Ng, L. (2016). Do managers matter for corporate innovation? *Journal of Corporate Finance*, *36*, 206–229.

Coles, J. L., & Li, Z. (2020). Managerial attributes, incentives, and performance. *The Review of Corporate Finance Studies*, *9*(2), 256–301.

Dang, C., Foerster, S., Li, Z. F., & Tang, Z. (2021). Analyst talent, information, and insider trading. *Journal of Corporate Finance*, *67*, 101803.

Davidson, R. H., Dey, A., & Smith, A. J. (2019). CEO materialism and corporate social responsibility. *The Accounting Review*, *94*(1), 101–126.

Dejong, D., & Ling, Z. (2013). Managers: Their effects on accruals and firm policies. *Journal of Business Finance & Accounting*, *40*(1-2), 82–114.

Devereux, K. (2018). Identifying the value of teamwork: Application to professional tennis. *Working Paper Series, No. 14, University of Waterloo, Canadian Labour Economics Forum (CLEF), Waterloo.*

Dyreng, S. D., Hanlon, M., & Maydew, E. L. (2010). The effects of executives on corporate tax avoidance. *The Accounting Review*, *85*(4), 1163–1189.

Ewens, M., & Rhodes-Kropf, M. (2015). Is a VC partnership greater than the sum of its partners? *The Journal of Finance*, *70*(3), 1081–1113.

Falato, A., Kadyrzhanova, D., & Lel, U. (2014). Distracted directors: Does board busyness hurt shareholder value? *Journal of Financial Economics*, *113*(3), 404–426.

Fee, C. E., Hadlock, C. J., & Pierce, J. R. (2013). Managers with and without style: Evidence using exogenous variation. *The Review of Financial Studies*, *26*(3), 567–601.

Fenizia, A. (2022). Managers and productivity in the public sector. *Econometrica*, *90*(3), 1063–1084.

Ferris, S. P., Jagannathan, M., & Pritchard, A. C. (2003). Too busy to mind the business? Monitoring by directors with multiple board appointments. *The Journal of Finance*, *58*(3), 1087–1111.

Fich, E. M., & Shivdasani, A. (2006). Are busy boards effective monitors? *The Journal of Finance*, *61*(2).

Firpo, S., Fortin, N. M., & Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, *77*(3), 953–973.

Fitza, M. A. (2014). The use of variance decomposition in the investigation of CEO effects: How large must the CEO effect be to rule out chance? *Strategic Management Journal*, *35*(12), 1839–1852.

Fitza, M. A. (2017). How much do CEOs really matter? Reaffirming that the CEO effect is mostly due to chance. *Strategic Management Journal*, *38*(3), 802–811.

Foerster, S., Linnainmaa, J. T., Melzer, B. T., & Previtero, A. (2017). Retail financial advice: Does one size fit all? *The Journal of Finance*, *72*(4), 1441–1482.

Francis, B. B., Hasan, I., & Zhu, Y. (2020). Managerial effect or firm effect: Evidence from the private debt market. *Financial Review*, *55*(1), 25–59.

Graham, J. R., Li, S., & Qiu, J. (2012). Managerial attributes and executive compensation. *The Review of Financial Studies*, *25*(1), 144–186.

Gul, F. A., Wu, D., & Yang, Z. (2013). Do individual auditors affect audit quality? Evidence from archival data. *The Accounting Review*, *88*(6), 1993–2023.

Hagendorff, J., Saunders, A., Steffen, S., & Vallascas, F. (2021). The wolves of Wall Street? Managerial attributes and bank risk. *Journal of Financial Intermediation*, *47*, 100921.

Hauser, R. (2018). Busy directors and firm performance: Evidence from mergers. *Journal of Financial Economics*, *128*(1), 16–37.

Herpfer, C. (2021). The role of bankers in the US syndicated loan market. *Journal of Accounting and Economics*, *71*(2-3), 101383.

Huang, J., & Wang, A. Y. (2015). The predictability of managerial heterogeneities in mutual funds. *Financial Management*, *44*(4), 947–979.

Jarosiewicz, V. E., & Ross, D. G. (2023). Revisiting managerial "style": The replicability and falsifiability of manager fixed effects for firm policies. *Strategic Management Journal*, *44*(3), 858–886.

Kiel, G. C., & Nicholson, G. J. (2006). Multiple directorships and corporate performance in Australian listed companies. *Corporate Governance: An International Review*, *14*(6), 530–546.

Krause, R., Li, W., Ma, X., & Bruton, G. D. (2019). The board chair effect across countries: An institutional view. *Strategic Management Journal*, *40*(10), 1570–1592.

Latif, B., Voordeckers, W., Lambrechts, F., & Hendriks, W. (2020). Multiple directorships in emerging countries: Fiduciary duties at stake? *Business Ethics: A European Review*, *29*(3), 629–645.

Lipton, M., & Lorsch, J. W. (1992). A modest proposal for improved corporate governance. *The Business Lawyer*, 59–77.

Ma, L. (2018). Essays on mutual funds and fund managers. *Doctoral Thesis at Humboldt-Universität zu Berlin*.

Masulis, R. W., & Mobbs, S. (2011). Are all inside directors the same? Evidence from the external directorship market. *The Journal of Finance*, *66*(3), 823–872.

Minaker, B. (2021). How effective are charity managers?: Evidence from a panel of charities. *Journal of Human Resources*, *56*(2), 632–654.

Pfeffer, J., & Salancik, G. R. (1978). *The external control of organizations: A resource dependence perspective*. New York: Harper Row.

Quigley, T. J., & Graffin, S. D. (2017). Reaffirming the CEO effect is significant and much larger than chance: A comment on Fitza (2014). *Strategic Management Journal*, *38*(3), 793–801.

Schoar, A., Yeung, K., & Zuo, L. (2023). The effect of managers on systematic risk. *Management Science*.

Trinh, V. Q., Aljughaiman, A. A., & Cao, N. D. (2020). Fetching better deals from creditors: Board busyness, agency relationships and the bank cost of debt. *International Review of Financial Analysis*, *69*, 101472.

Wells, K. (2020). Who manages the firm matters: The incremental effect of individual managers on accounting quality. *The Accounting Review*, *95*(2), 365–384.

A Appendix

A.1 Exogenous mobility

Adding to the explanations in Section 4.1, the present section discusses the assumption of exogenous mobility for the dataset of publicly listed German firms during the years 2005 – 2019 as described in Section 3. This assumption posits that the error term from the two-way fixed-effects model as presented in (1) is uncorrelated to mobility patterns (Card et al., 2013). If including firm effects and person effects separately in a model, time-invariant characteristics of a spell, i.e. a person-firm-match are not controlled for. For this reason, previous studies estimating two-way fixed-effects models examine whether exogenous mobility is satisfied in their datasets. A widely used test consists in comparing R^2 between the two-way fixed-effects model and a model with spell fixed-effects. Examples include Section 6.1 from Cavaco et al. (2017) and Table IV from Fenizia (2022). In Table A1, I run this test for the different dependent variables (ROA, ROE, Stock Return) from the present dataset. For all three dependent variables, the change in R^2 amounts to less than 0.01 when switching from the two-way fixed-effects specification to the spell fixed-effects specification. This suggests that the idiosyncratic match component does not play a major role, mitigating the concerns of endogenous mobility.

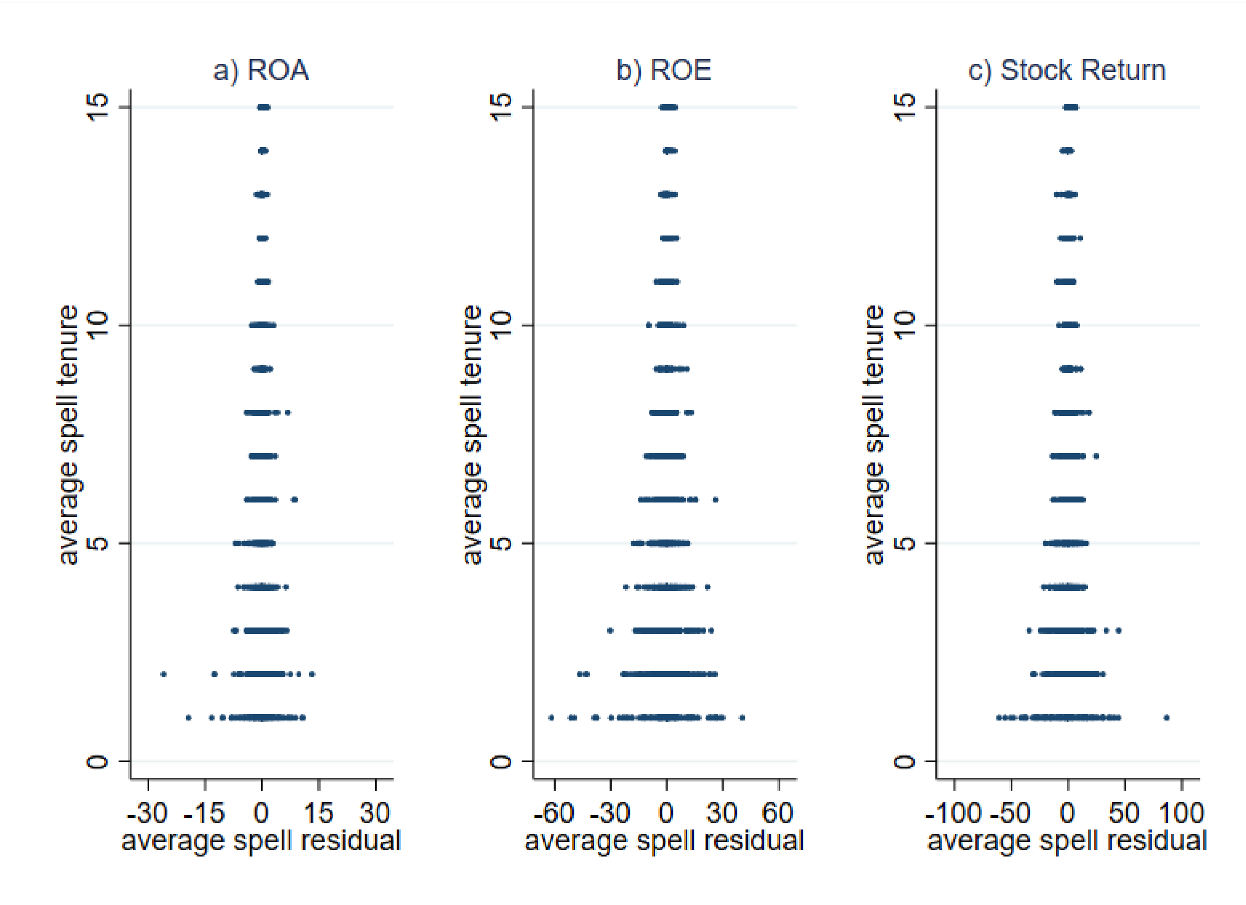
Table A1: Tests for exogenous mobility.

	(1) ROA	(2) ROA	(3) ROE	(4) ROE	(5) Stock Return	(6) Stock Return
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Person FE	Yes	No	Yes	No	Yes	No
Firm FE	Yes	No	Yes	No	Yes	No
Interacted Person-Firm FE	No	Yes	No	Yes	No	Yes
Observations	24036	24036	24036	24036	24036	24036
R^2	0.669	0.678	0.659	0.667	0.448	0.455

Note: The table compares R^2 between specifications with additively separable person fixed-effects and firm fixed-effects versus interacted person-firm fixed-effects, i.e. spell fixed-effects. The dependent variable of each specification is indicated by the column header (ROA, ROE, Stock Return). The table is based on the sample of 4,076 supervisory directors from 156 publicly listed German firms described in Section 3.

Further, in a setting of sports economics, Devereux (2018) analyses whether the match component is correlated with the duration of the existence of this match. The weaker these are correlated, the lower the relevance of endogenous mobility patterns is. He interprets the match component as the average residual per i-j-match. In Figure A1, I plot the relation between the match-level average residual and the average tenure of each match for the three dependent variables. The distribution of the average residual becomes more compressed once match tenure increases. This is due to a decreasing number of observations for higher levels of match tenure. Apart from that, no relation between both variables is visible. This is consistent with very low correlation coefficients of 0.0185 for ROE, 0.0137 for ROA and 0.0150 for Stock Return, fostering the assumption of exogenous mobility.

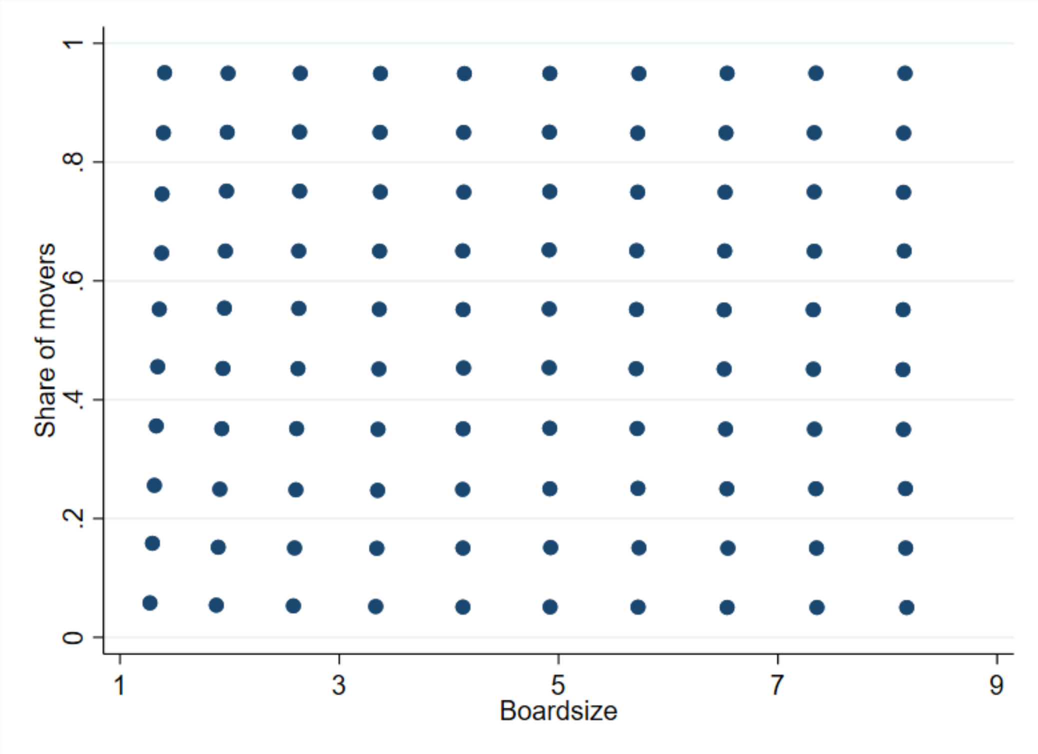
Figure A1: Average spell tenure and average spell residuals



Note: The graph shows a scatterplot of the average tenure of each director-firm-combination (i.e. for each spell) and the average residual from a two-way fixed-effects estimation of the respective director-firm-combination. The dependent variable of each specification is indicated by the column header (ROA, ROE, Stock Return). The table is based on the sample of 4,076 supervisory directors from 156 publicly listed German firms described in Section 3. The correlation between the average residual and the average match tenure amounts to 0.0185 (ROA), 0.0137 (ROE) and 0.0150 (Stock Return).

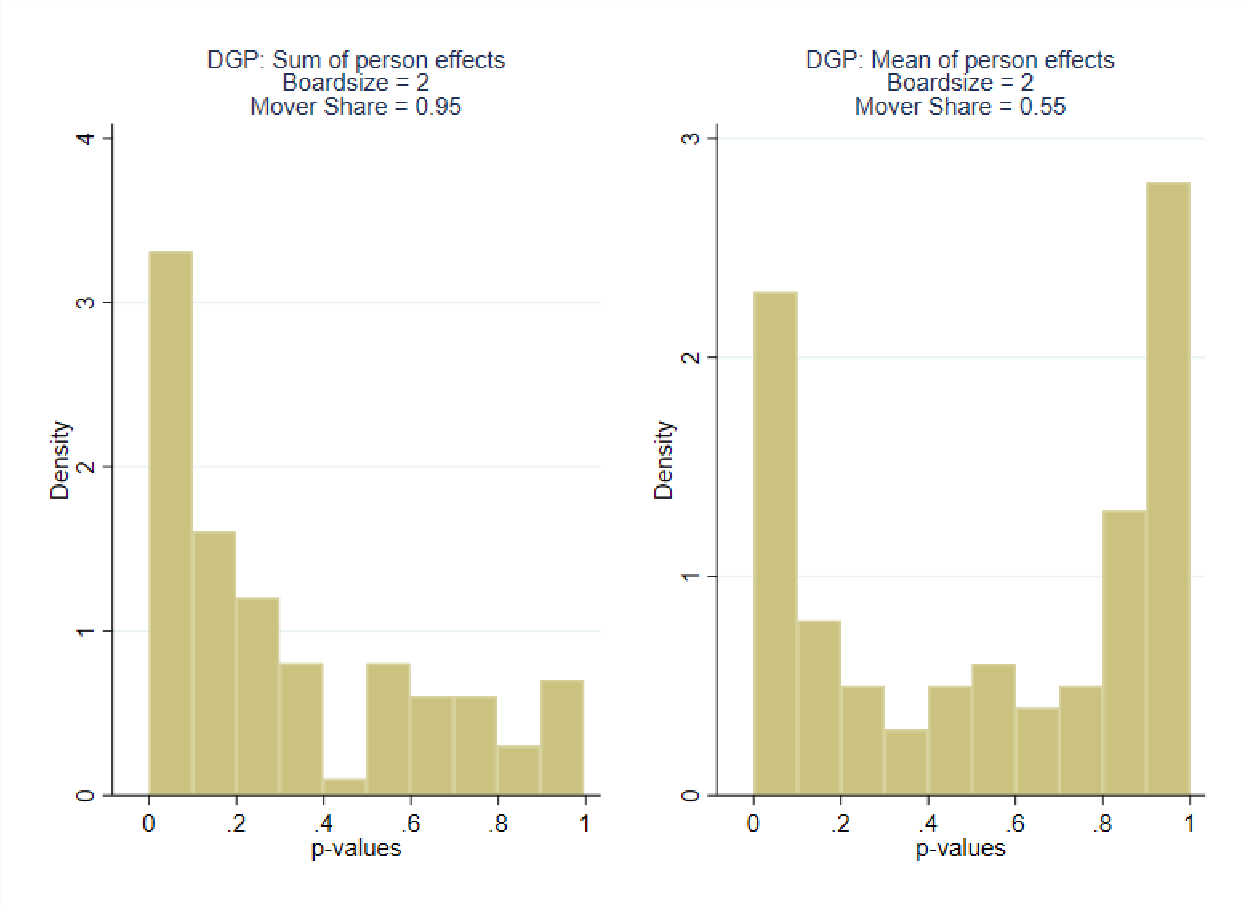
A.2 Testing for the presence of person effects: Insights from simulations (additional analyses)

Figure A2: Boardsize-mobility combinations from simulations



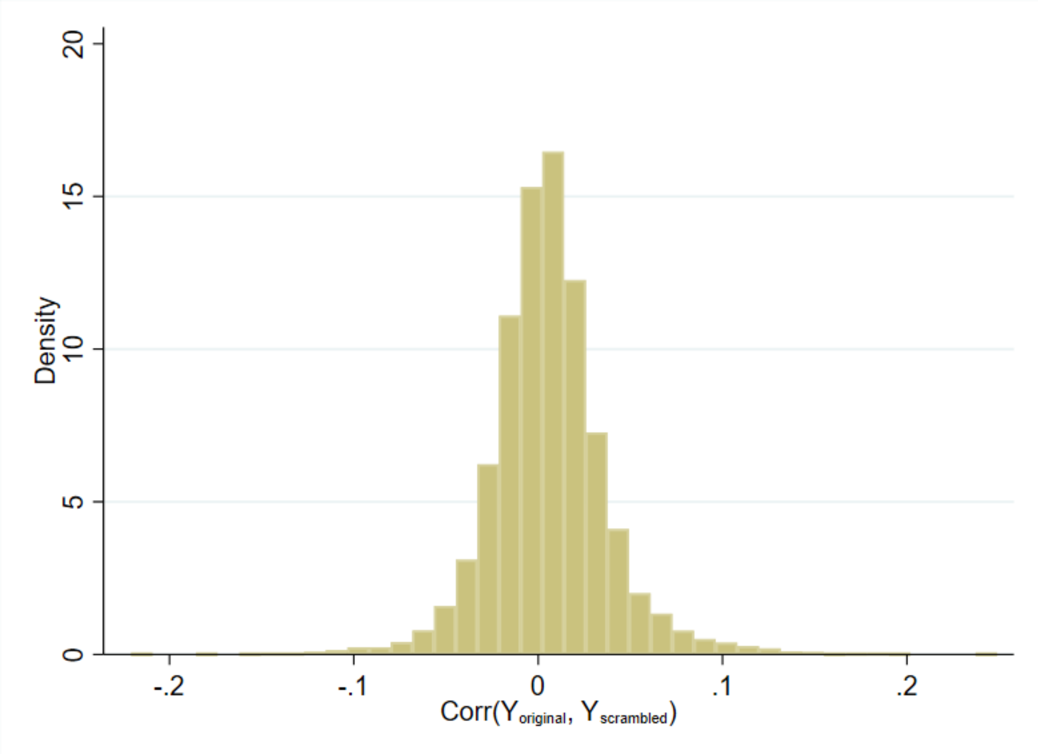
Note: The graph shows a scatterplot of boardsize-mobility combinations used for the simulations. The simulations can be stratified along 10 boardsizes and 10 mover shares, yielding 100 different boardsize-mobility combinations. Each such boardsize-mobility combination is illustrated by a data point which reflects an average over 100 iterations. The underlying dataset of each iteration is based on 100 firms, observed over 15 years.

Figure A3: Examples for variation in p-values within boardsize-mobility combinations



Note: The two histograms represent examples for the distribution of the p-value of the F-test for the joint significance of person effects in scrambled samples. Each histogram is based on 100 iterations of a specific boardsize-mobility combination as indicated by the titles of both graphs. The average p-values in the original simulated samples (without scrambling) amount to 0 (left histogram) and to 0.0026 (right histogram).
The DGPs are given by
(1) Sum of person effects: $y_{j(i)t} = (\sum_{i=1}^{n_{jt}} \theta_{i \in (jt)}) + \psi_j + u_{jt}$, equation (7)
(2) Mean of person effects: $y_{j(i)t} = (\frac{1}{n_{jt}} * \sum_{i=1}^{n_{jt}} \theta_{i \in (jt)}) + \psi_j + u_{jt}$, equation (8)

Figure A4: Correlation between the dependent variable from one simulated sample and one scrambled version of the simulated sample



Note: The graph plots the histogram of the correlation coefficients between the dependent variable from one simulated sample and one scrambled version of the simulated sample. It is based on 10,000 iterations, stratified by 10 boardsizes and 10 mobility shares with 100 iterations per boardsize-mobility combination. The underlying dependent variable is simulated based on the the DGP specified in equation (7) where person effects θ_i, firm effects ψ_j and error terms u_{jt} are drawn from independent standard normal distributions.

A.3 Scrambling with person-level y-variable

To assess the relevance of the scrambling test for firm-level dependent variables suggested by Jarosiewicz & Ross (2023), I compare it to the scrambling in case of person-level dependent variables. In the former case, the scrambling changes only the structure of the dataset. Hence, the composition of persons at each firm changes so that other persons than in the original sample are associated with one firm-year-level outcome. In contrast, when scrambling the data for a person-level dependent variable, two factors are at play: First, persons are randomly assigned to different firms and, hence, different coworkers. Second, as each firm-person-year coincides with another outcome (such as compensation for a board appointment of person i in firm j), each firm receives a new set of outcomes that stem from the new persons assigned to the firm.

For the application of the scrambling with a person-level dependent variable, I consider the following two DGPs where person effects θ_i, firm effects ψ_j and error terms u_{ijt} are drawn from independent standard normal distributions:

1) The person-level dependent variable follows a two-way fixed-effects model:

$$y_{it} = \theta_i + \psi_j + u_{ijt} \tag{14}$$

2) The person-level dependent variable follows a one-way fixed-effects model:

$$y_{it} = \psi_j + u_{ijt} \tag{15}$$

Hence, equations (14) and (15) are equivalent to (7) and (9) with the only difference that (14) and (15) consider person-level outcomes instead of firm-level outcomes and therefore do not aggregate the person effects θ_i per firm-year.

The DGP from equation (14) represents the canonical form of the two-way fixed-effects model where both person effects and firm effects are used to describe a person-level dependent variable such as wages (Abowd et al., 1999). In contrast, model (15) represents a reduced version of a two-way fixed-effects model where person fixed-effects are assumed to equal zero. In the initial AKM setup for wages, this model would assume the existence of high-wage and low-wage firms without implementing person-level heterogeneity. Thus, estimated person effects should reflect random noise. Specifying a DGP for a person-level variable without including person effects might be regarded as a restrictive assumption. Still, it is an important extension to the analysis of the DGPs from equations (7) and (9) as it allows to separate the influence of the level of variation of the dependent variable from the existence of person effects in the model.

The results for the original and scrambled datasets based on 100 iterations of the DGPs from equations (14) and (15) are presented in Table A2. The central difference to Table 6 consists in the role of firm effects. In both Panels A and B of Table A2, the component of R^2 attributable to firm effects reduces markedly when scrambling the data (from 0.33 to 0.02 and from 0.50 to 0.02, on average). This decline is due to the dependent variable varying on the person-level. As outlined above, the scrambling procedure assigns new person-level outcomes to each firm. Since these outcomes originate from many other firms in the dataset, they are not based on the same firm effect. Hence, the scrambling procedure removes any explanatory power of the firm effects when considering a person-level dependent variable.

Table A2: Component of R^2 attributable to person effects and firm effects for person-level dependent variables, based on equation (5). Comparison between original data and randomly scrambled data.

Panel A: The person-level dependent variable follows a two-way fixed-effects model: $y_{it} = \theta_i + \psi_j + u_{ijt}$				
	Mean	SD	Min	Max
$\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (original)	0.39	0.02	0.35	0.42
$\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (scrambled)	0.68	0.01	0.64	0.73
$\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (orig.) - $\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (scram.)	-0.29	0.03	-0.36	-0.23
$\frac{Cov(y_{j(i)t},\hat{\psi}_j)}{Var(y_{j(i)t})}$ (original)	0.33	0.03	0.28	0.39
$\frac{Cov(y_{j(i)t},\hat{\psi}_j)}{Var(y_{j(i)t})}$ (scrambled)	0.02	0.01	0.00	0.03
$\frac{Cov(y_{j(i)t},\hat{\psi}_j)}{Var(y_{j(i)t})}$ (orig.) - $\frac{Cov(y_{j(i)t},\hat{\psi}_j)}{Var(y_{j(i)t})}$ (scram.)	0.32	0.03	0.26	0.39
Panel B: The person-level dependent variable follows a one-way fixed-effects model: $y_{it} = \psi_j + u_{ijt}$				
	Mean	SD	Min	Max
$\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (original)	0.08	0.02	0.04	0.13
$\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (scrambled)	0.51	0.02	0.44	0.57
$\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (orig.) - $\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ (scram.)	-0.43	0.03	-0.51	-0.35
$\frac{Cov(y_{j(i)t},\hat{\psi}_j)}{Var(y_{j(i)t})}$ (original)	0.50	0.04	0.41	0.59
$\frac{Cov(y_{j(i)t},\hat{\psi}_j)}{Var(y_{j(i)t})}$ (scrambled)	0.02	0.01	0.01	0.05
$\frac{Cov(y_{j(i)t},\hat{\psi}_j)}{Var(y_{j(i)t})}$ (orig.) - $\frac{Cov(y_{j(i)t},\hat{\psi}_j)}{Var(y_{j(i)t})}$ (scram.)	0.47	0.04	0.39	0.57

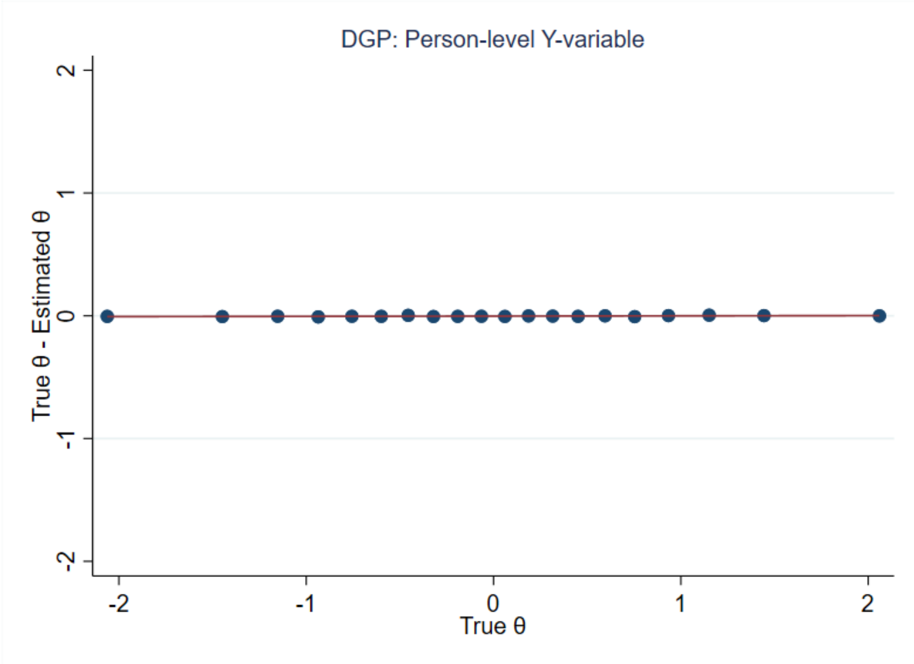
Note: The sample presents descriptive statistics for the share in R^2 attributable to person effects and firm effects, as defined by equation (5). The data structure is based on the sample of 4,076 supervisory directors from 156 publicly listed German firms described in Section 3 to maintain properties of a real-world sample. The scrambling procedure randomly assigns person spells to other firms. The real-world dependent variable is replaced by simulated data according to two different DGPs as defined in the panel headers for panels A and B. Each panel refers to 100 iterations. All components from the DGP, i.e. θ_i, ψ_j and u_{ijt}, are drawn from independent standard normal distributions.

In both Panels A and B, this decrease in the explanatory power of firm effects coincides with a rise in the explanatory power of person effects of almost the same size (0.29 vs. 0.32 in Panel A and 0.43 vs. 0.47 in Panel B). For Panel A, this increase in $\frac{Cov(y_{j(i)t},\hat{\theta}_i)}{Var(y_{j(i)t})}$ is consistent with the DGP containing person effects. Therefore, assigning persons with their person-firm-year-specific outcomes to other firms rules out any explanatory power by the firm effects but does not impair the explanatory power of person effects. However, the findings in Panel B cannot be explained by this reasoning. Similar to Panel A, I document a sharp increase in the component of R^2 attributable to person effects when scrambling the firm-person-spells to other firms. Hence, this result contrasts with Panel C of Table 6 where the explanatory power of person effects remains constant when scrambling the data. As Panel B of Table A2 assumes a DGP without person-level heterogeneity, person effects should not add any explanatory power. However, the DGP contains firm effects. Thus, in the original sample, all non-movers are de facto assigned a person effect equal to the firm effect as the latter is invariant for each non-mover. The movers, in contrast, are de facto assigned weighted averages of all the firms where they were observed before scrambling. For this reason, person effects capture the variation that was explained by firm effects before scrambling the data.

Hence, the idea of scrambling the person-firm-spells from Jarosiewicz & Ross (2023) targets especially settings with firm-level dependent variables where a two-way fixed-effects model shall be estimated. For person-level dependent variables, as in the present case, another scrambling procedure has to be applied. This, in turn, leads to person effects absorbing firm-level heterogeneity. Thus, testing for the existence of idiosyncratic person effects in settings with person-level dependent variables is less meaningful as scrambling these datasets might always result in increased explanatory power of person effects.

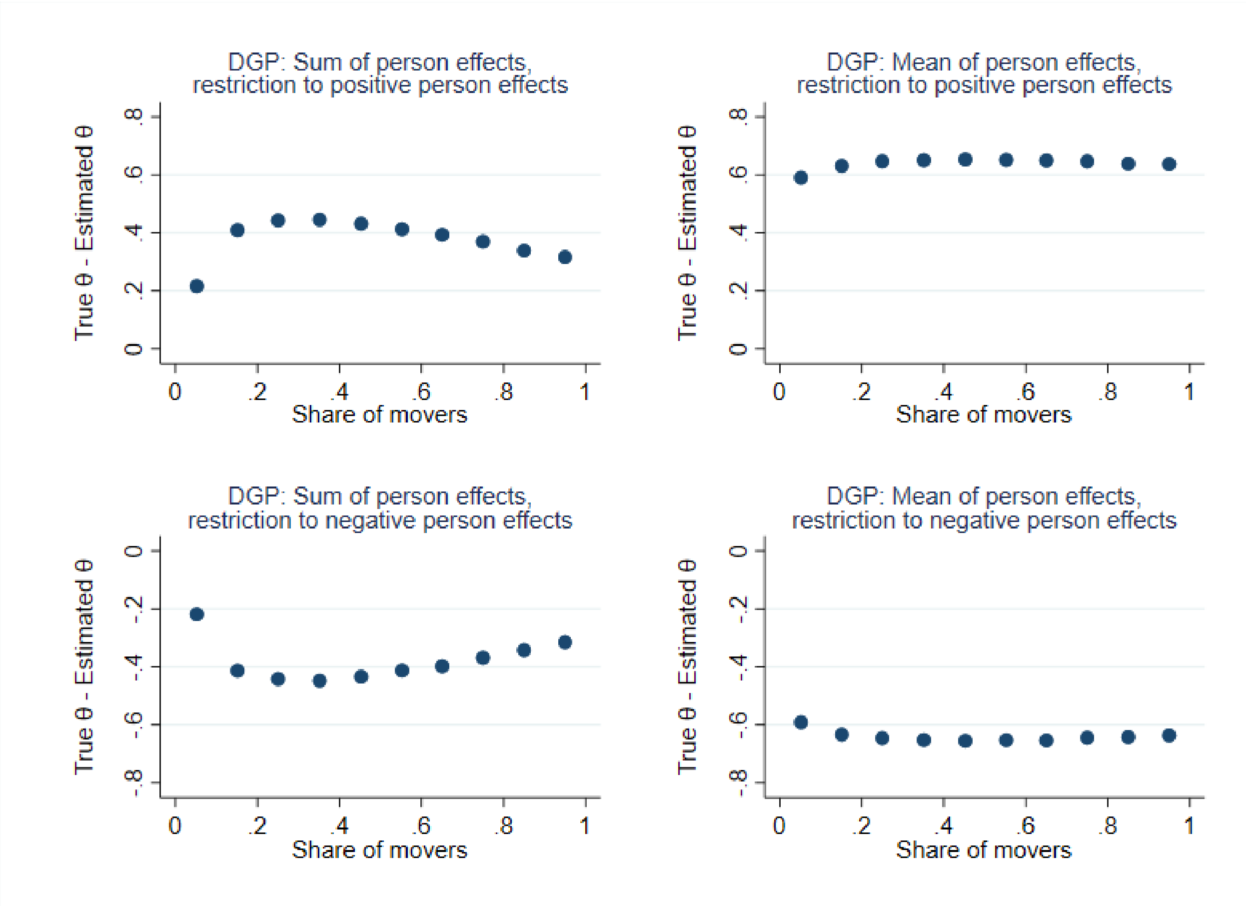
A.4 Properties of estimated person effects (additional analyses)

Figure A5: Estimation error in person effects by true person effects for a person-level Y-variable



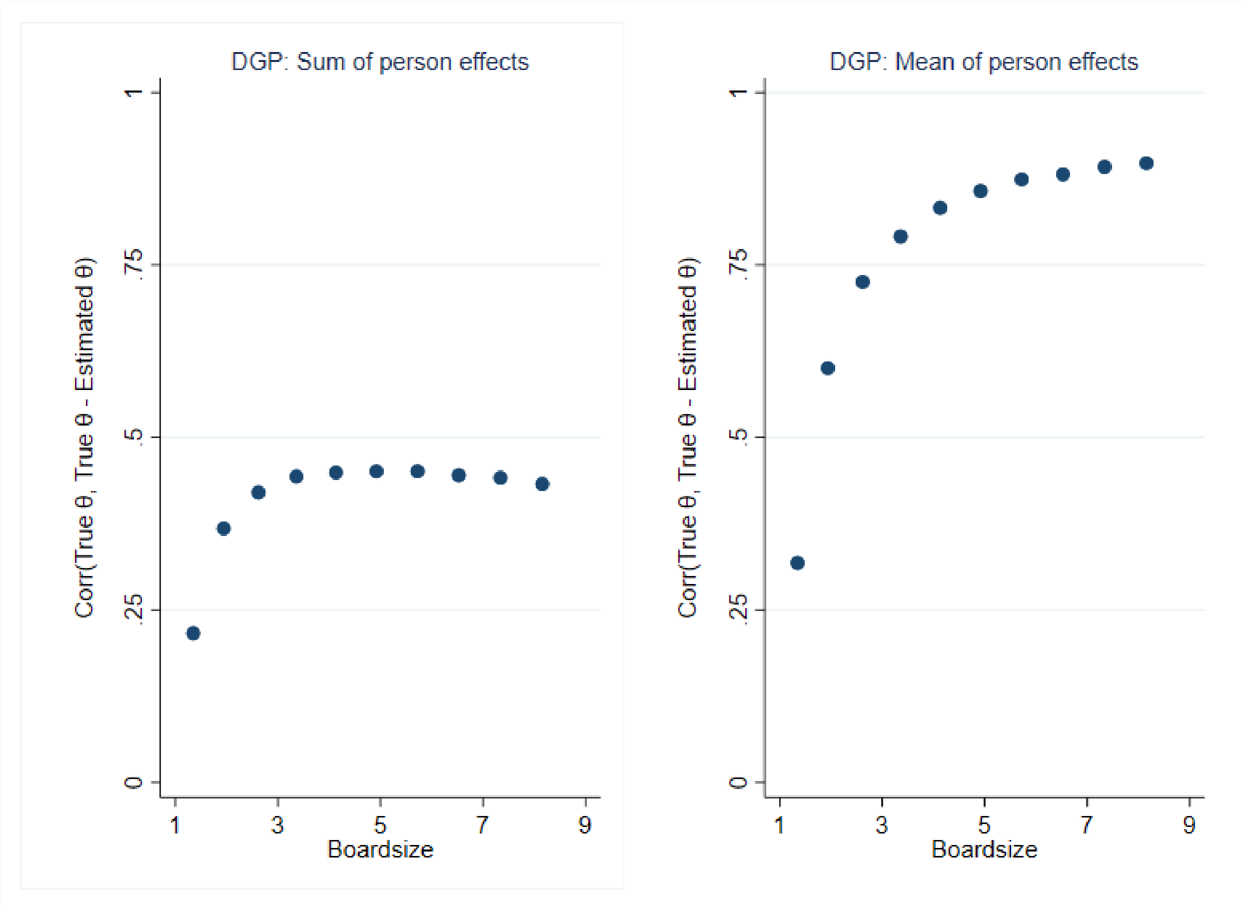
Note: The graph shows a binned scatterplot of true person effects (x-axis) and the difference between true person effects and estimated person effects (y-axis), generated by simulations. It is based on 10,000 iterations, stratified by 10 boardsizes and 10 mover shares and 100 iterations per boardsize-mobility combination. The underlying datasets generated for each iteration consist of 100 firms, observed over 15 years. The DGP for the person-level dependent variable is given by $y_{it} = \theta_i + \psi_j + u_{ijt}$ where person effects θ_i, firm effects ψ_j and error terms u_{ijt} are drawn from independent standard normal distributions.

Figure A6: Estimation error in person effects by mover share



Note: The graphs plot the difference between true person effects and estimated person effects (y-axis) for different mover shares (x-axis). The data is based on simulations with 10,000 iterations for each of the two DGPs, stratified by 10 boardsizes, 10 mover shares and 100 iterations per boardsize-mobility combination. The underlying datasets generated for each iteration consist of 100 firms, observed over 15 years. The DGPs are given by equations (7) (two graphs on the left) and (8) (two graphs on the right) where person effects θ_i, firm effects ψ_j and error terms u_{jt} are drawn from independent standard normal distributions.

Figure A7: Correlation between true person effects and the difference between true person effects and estimated person effects by boardsize



Note: The graphs plot the correlation between true person effects and the difference between true person effects and estimated person effects (y-axis) for different boardsizes (x-axis). The data is based on simulations with 10,000 iterations for each of the two DGPs, stratified by 10 boardsizes, 10 mobility shares per boardsize and 100 iterations per boardsize-mobility combinations. The underlying datasets generated for each iteration consist of 100 firms, observed over 15 years. The DGPs are given by equations (7) (graph on the left) and (8) (graph on the right) where person effects θ_i, firm effects ψ_j and error terms u_{jt} are drawn from independent standard normal distributions.

Table A3: Spurious findings in Unconditional Quantile Regressions, based on ROA-based person effects.

Panel A: Busy				
	Mean	SD	Min	Max
t-statistic at quantile 10	4.45	1.29	1.15	7.47
t-statistic at quantile 20	3.60	1.39	0.27	7.29
t-statistic at quantile 30	2.26	1.06	-0.37	4.79
t-statistic at quantile 40	1.04	1.03	-2.07	3.26
t-statistic at quantile 50	-0.11	0.87	-2.86	1.76
t-statistic at quantile 60	-1.20	0.94	-3.37	0.65
t-statistic at quantile 70	-2.40	1.05	-5.05	-0.19
t-statistic at quantile 80	-3.49	1.14	-7.04	-1.09
t-statistic at quantile 90	-4.54	1.21	-7.87	-1.61
Panel B: At least two firms				
	Mean	SD	Min	Max
t-statistic at quantile 10	4.84	1.28	2.62	8.53
t-statistic at quantile 20	3.71	1.12	1.06	7.09
t-statistic at quantile 30	2.28	1.06	-0.82	4.75
t-statistic at quantile 40	1.07	0.91	-1.44	3.44
t-statistic at quantile 50	0.05	0.99	-2.69	2.17
t-statistic at quantile 60	-1.05	0.94	-3.72	0.64
t-statistic at quantile 70	-2.27	1.01	-4.52	0.17
t-statistic at quantile 80	-3.49	0.94	-5.99	-1.56
t-statistic at quantile 90	-4.78	1.28	-8.11	-2.14

Note: The table presents t-statistics for the coefficients from Unconditional Quantile Regressions (UQR) as introduced by Firpo et al. (2009). The t-statistics refer to coefficients from regressions where the dependent variable is the recentered influence function (RIF) as defined in (13) of ROA-based person effects. The structure of the underlying dataset is a scrambled version of the sample of 4,076 supervisory directors from 156 publicly listed German firms described presented in Section 3. Both panels are based on 100 such scrambling iterations. In Panel A, the dependent variable is regressed on *Busy*, a dummy indicating that a director serves on two supervisory boards at the same time. In Panel B, the explanatory variable is a dummy indicating that a director is observed in at least two firms in the sample, either simultaneously or in sequence.

A.5 Correlation of person effects of different dependent variables (additional analyses)

This section provides a combination of analytical and simulation analysis to investigate the correlation of person effects of different dependent variables. Hence, it provides an extension to the analysis in Section 5.5.

Let y_1 and y_2 denote two different corporate policy variables whose person effects shall be compared. The two corresponding person effects $\hat{\theta}_{i1}$ and $\hat{\theta}_{i2}$ estimated from a two-way fixed-effects model differ from their true values by error terms e_{i1} and e_{i2} such that $\hat{\theta}_{i1} = \theta_{i1} + e_{i1}$ and $\hat{\theta}_{i2} = \theta_{i2} + e_{i2}$. The subsequent analysis shall uncover determinants of the correlation of the estimated person effects $Corr(\hat{\theta}_{i1}, \hat{\theta}_{i2})$ as compared to the correlation of the true person effects $Corr(\theta_{i1}, \theta_{i2})$. Hence, using the definition of a correlation and applying simple rearrangements yields the following formula:

$$
\begin{aligned}
& Corr(\hat{\theta}_{i1}, \hat{\theta}_{i2}) - Corr(\theta_{i1}, \theta_{i2}) \\
=& \frac{Cov(\hat{\theta}_{i1}, \hat{\theta}_{i2})}{\sqrt{Var(\hat{\theta}_{i1}) * Var(\hat{\theta}_{i2})}} - \frac{Cov(\theta_{i1}, \theta_{i2})}{\sqrt{Var(\theta_{i1}) * Var(\theta_{i2})}} \\
=& \frac{Cov(\theta_{i1} + e_{i1}, \theta_{i2} + e_{i2})}{\sqrt{Var(\hat{\theta}_{i1}) * Var(\hat{\theta}_{i2})}} - \frac{Cov(\theta_{i1}, \theta_{i2})}{\sqrt{Var(\theta_{i1}) * Var(\theta_{i2})}} \\
=& \frac{Cov(\theta_{i1}, \theta_{i2}) + Cov(\theta_{i1}, e_{i2}) + Cov(\theta_{i2}, e_{i1}) + Cov(e_{i1}, e_{i2})}{\sqrt{Var(\hat{\theta}_{i1}) * Var(\hat{\theta}_{i2})}} - \frac{Cov(\theta_{i1}, \theta_{i2})}{\sqrt{Var(\theta_{i1}) * Var(\theta_{i2})}} \\
=& \frac{Cov(\theta_{i1}, \theta_{i2})}{\sqrt{Var(\hat{\theta}_{i1}) * Var(\hat{\theta}_{i2})}} - \frac{Cov(\theta_{i1}, \theta_{i2})}{\sqrt{Var(\theta_{i1}) * Var(\theta_{i2})}} + \frac{Cov(\theta_{i1}, e_{i2}) + Cov(\theta_{i2}, e_{i1}) + Cov(e_{i1}, e_{i2})}{\sqrt{Var(\hat{\theta}_{i1}) * Var(\hat{\theta}_{i2})}}
\end{aligned}
\tag{16}
$$

Whether or not the correlation of estimated person effects exceeds the correlation of true person effects is a priori not clear when considering equation (16). Hence, Table A4 analyses the different components from equation (16).

All expressions from the first column in Table A4 are evaluated for (1) all iterations, (2) the iterations with positive correlations according to the DGP and (3) the iterations with negative correlations. The analysis is split into two different DGPs, including the sum of person effects per firm-year (Panel A) and the average of person effects per firm-year (Panel B). Comparing the denominators in the first two fractions in equation (16), I find that $\sqrt{Var(\hat{\theta}_{i1}) * Var(\hat{\theta}_{i2})} > \sqrt{Var(\theta_{i1}) * Var(\theta_{i2})}$ for the majority of iterations in Panel A (13,675 of 20,000) but only for the minority of cases in Panel B (5,145 of 20,000). Hence, the evaluation which of the first two fractions of equation (16) is larger depends on both the assumption on the underlying DGP and on the sign of the numerator.

Table A4: Disentangling the components of the difference in correlations, based on equation (16).

Panel A:
DGP: $y_{j(i)t} = (\sum_{i=1}^{n_{jt}} \theta_{i\,\in\,(jt)}) + \psi_j + u_{jt}$

	(1) All	(2) $Corr(\theta_{i1}, \theta_{i2}) > 0$	(3) $Corr(\theta_{i1}, \theta_{i2}) < 0$
N	20,000	10,000	10,000
$\sqrt{Var(\hat{\theta}_{i1}) * Var(\hat{\theta}_{i2})} > \sqrt{Var(\theta_{i1}) * Var(\theta_{i2})}$	13,675	6,890	6,785
$\frac{Cov(\theta_{i1},e_{i2})+Cov(\theta_{i2},e_{i1})+Cov(e_{i1},e_{i2})}{\sqrt{Var(\hat{\theta}_{i1})*Var(\hat{\theta}_{i2})}}$ and $Cov(e_{i1}, e_{i2})$ have the same sign	10,773	5,378	5,395
$Cov(e_{i1}, e_{i2}) > 0$	10,032	9,388	644
$Corr(\hat{\theta}_{i1}, \hat{\theta}_{i2}) - Corr(\theta_{i1}, \theta_{i2}) < 0$	10,042	7,975	2,067
$Corr(\hat{\theta}_{i1}, \hat{\theta}_{i2}) - Corr(\theta_{i1}, \theta_{i2}) > 0$	9,958	2,025	7,933

Panel B:
DGP: $y_{j(i)t} = (\frac{1}{n_{jt}} * \sum_{i=1}^{n_{jt}} \theta_{i\,\in\,(jt)}) + \psi_j + u_{jt}$

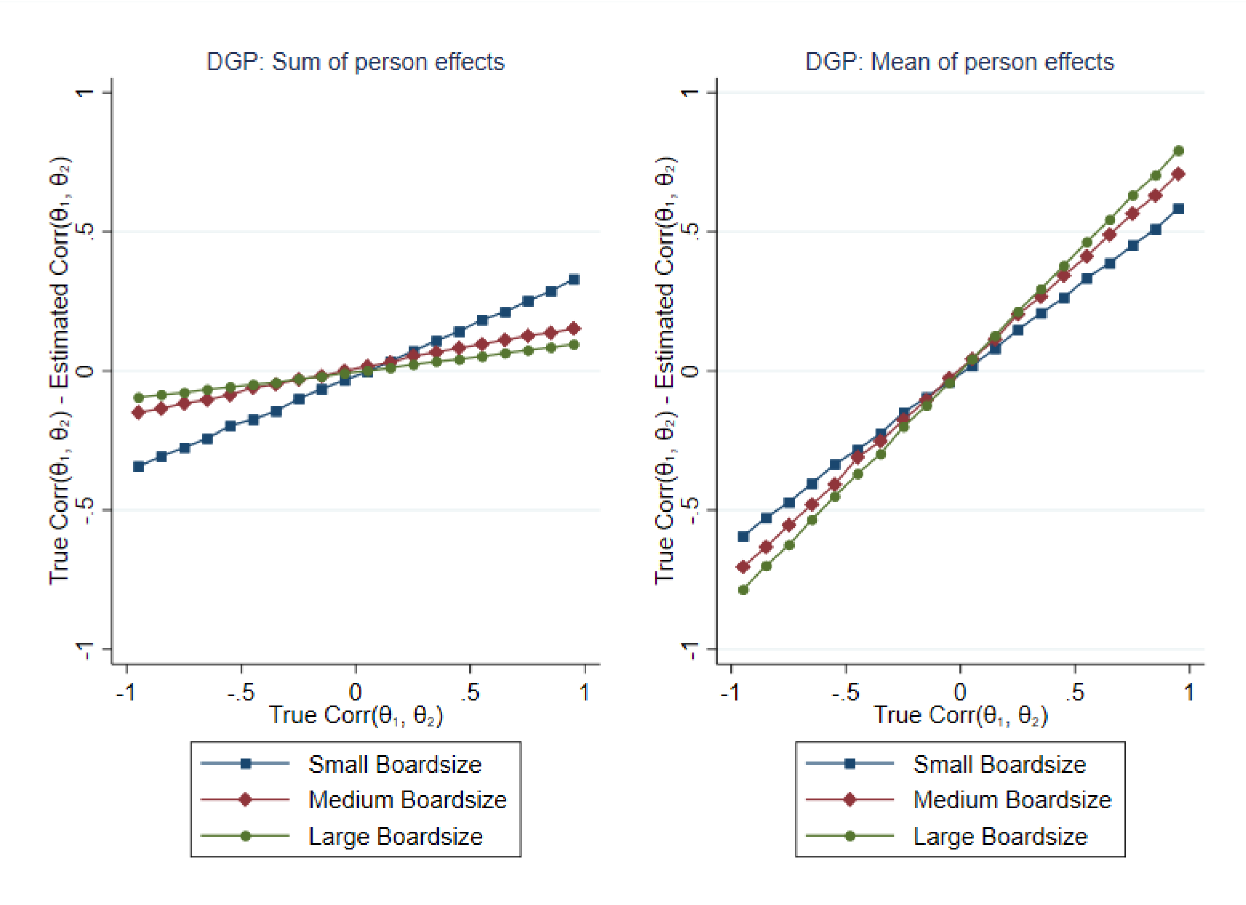
	(1) All	(2) $Corr(\theta_{i1}, \theta_{i2}) > 0$	(3) $Corr(\theta_{i1}, \theta_{i2}) < 0$
N	20,000	10,000	10,000
$\sqrt{Var(\hat{\theta}_{i1}) * Var(\hat{\theta}_{i2})} > \sqrt{Var(\theta_{i1}) * Var(\theta_{i2})}$	5,145	2,572	2,573
$\frac{Cov(\theta_{i1},e_{i2})+Cov(\theta_{i2},e_{i1})+Cov(e_{i1},e_{i2})}{\sqrt{Var(\hat{\theta}_{i1})*Var(\hat{\theta}_{i2})}}$ and $Cov(e_{i1}, e_{i2})$ have the same sign	2,811	1,381	1,430
$Cov(e_{i1}, e_{i2}) > 0$	10,026	9,356	670
$Corr(\hat{\theta}_{i1}, \hat{\theta}_{i2}) - Corr(\theta_{i1}, \theta_{i2}) < 0$	10,060	9,365	695
$Corr(\hat{\theta}_{i1}, \hat{\theta}_{i2}) - Corr(\theta_{i1}, \theta_{i2}) > 0$	9,940	635	9,305

Note: Simulated data based on 10 iterations for 20 different levels of correlations $Corr(\theta_{i1}, \theta_{i2})$ = -0.95, -0.85, ..., 0.75, 0.85, 0.95 between person effects θ_i estimated from two-way fixed-effects models of two dependent variables y_1 and y_2 with 10 different shares of moving directors and 10 different boardsizes, yielding a total of $20 * 10 * 10 * 10 = 20,000$ iterations. The underlying datasets generated for each iteration consist of 100 firms, observed over 15 years. The expressions from the first column refer to different elements of equation (16). The table reports the signs and compares magnitudes of these expressions. Column (1) refers to the pooled dataset of both positive and negative values of $Corr(\theta_{i1}, \theta_{i2})$ while the last two columns consider positive and negative correlations separately. Person effects θ_i, firm effects ψ_j and error terms u_{jt} as specified in the DGP are drawn from independent standard normal distributions.

Further, the last fraction of equation (16) might take on either sign. The denominator is strictly positive as it reflects the product of two standard deviations. As indicated by Table A4, the numerator has the same sign as $Cov(e_{i1}, e_{i2})$ in half of the cases in Panel A (10,773 of 20,000 iterations) and in approximately 14 percent of iterations in Panel B (2,811 of 20,000 iterations). The sign of the covariance of the error terms itself often coincides with the sign of the true correlation $Corr(\theta_{i1}, \theta_{i2})$ (9,388 of 10,000 in the second column of Panel A, 9,356 of 10,000 in the second column of Panel B). In sum, both the sign of the difference between the first and the second fraction as well as the sign of the third fraction are ambivalent, yielding no indication whether the correlation of estimated person effects is overestimated or underestimated.

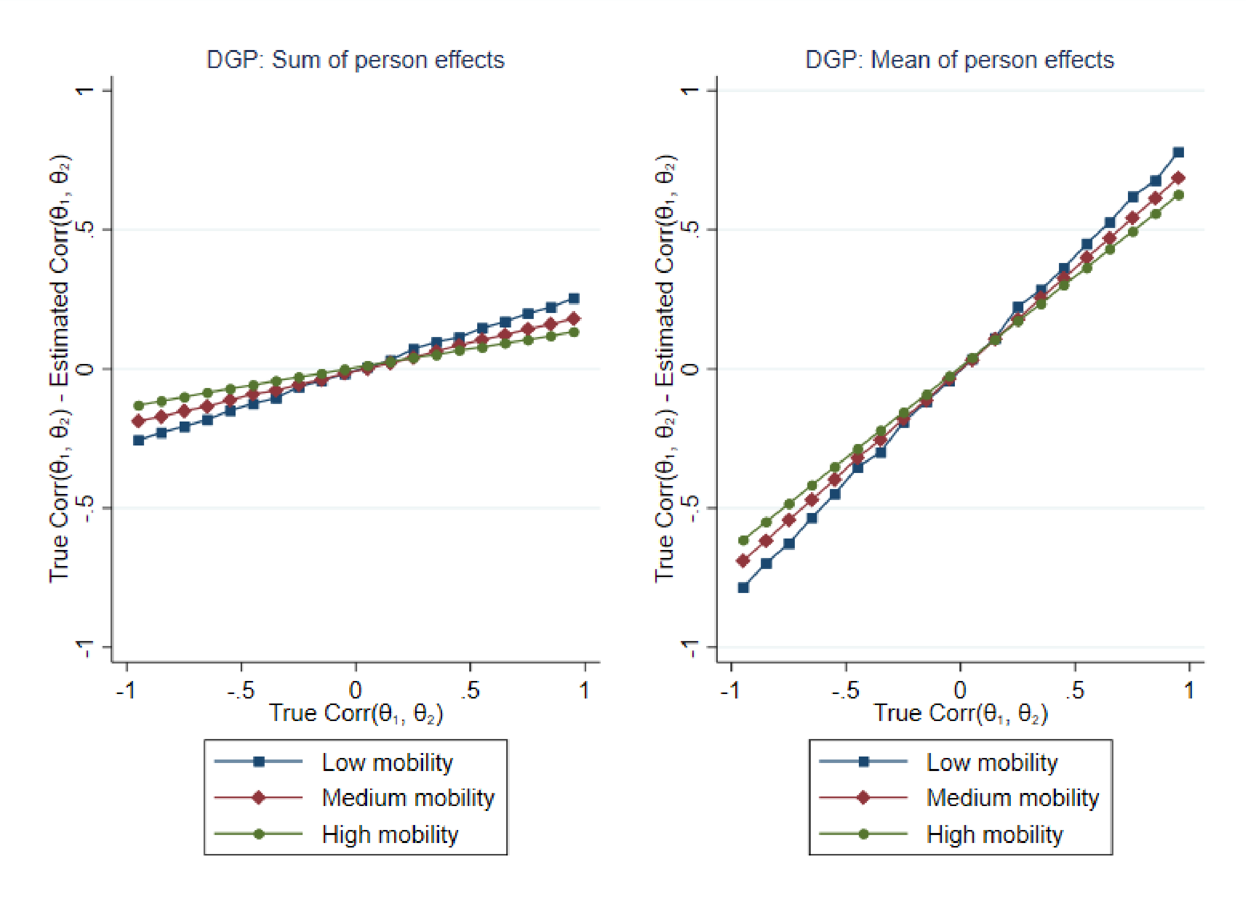
However, the last two lines of both Panel A and Panel B suggest that the correlation of estimated person effects is frequently understated in the positive range (7,975 of 10,000 iterations in Panel A and 9,365 of 10,000 iterations in Panel B) and, equivalently, overstated in the negative range (7,933 of 10,000 iterations in Panel A and 9,305 of 10,000 iterations in Panel B). Hence, this table reiterates the pattern of attenuation bias in the estimated correlation as presented in Figure 5.

Figure A8: Estimation error in the correlation of person effects of different dependent variables by boardsize



Note: The graphs plot the correlation of true person effects (x-axis) and the difference between correlations of true person effects and the correlations of estimated person effects (y-axis), generated by simulations. The data considers 20 different levels of correlations $Corr(\theta_{i1}, \theta_{i2})$ = -0.95, -0.85, ..., 0.75, 0.85, 0.95 between person effects θ_i estimated from two-way fixed-effects models of two different dependent variables y_1 and y_2. Both graphs are based on 20,000 iterations, stratified by 20 levels of correlations, 10 boardsizes, 10 mobility shares and 10 iterations of each boardsize-mobility-correlation-combination. Based on Figure A2, the 10 levels of boardsize are stratified by three levels of small boardsizes, four levels of medium boardsizes and three levels of large boardsizes. The underlying datasets generated for each iteration consist of 100 firms, observed over 15 years.
The DGPs are specified as follows
(1) Sum of person effects: $y_{j(i)t} = (\sum_{i=1}^{n_{jt}} \theta_{i \in (jt)}) + \psi_j + u_{jt}$
(2) Mean of person effects: $y_{j(i)t} = (\frac{1}{n_{jt}} * \sum_{i=1}^{n_{jt}} \theta_{i \in (jt)}) + \psi_j + u_{jt}$
Person effects θ_i, firm effects ψ_j and error terms u_{jt} as specified in the DGP are drawn from independent standard normal distributions.

Figure A9: Estimation error in the correlation of person effects of different dependent variables by mobility



Note: The graphs plot the correlation of true person effects (x-axis) and the difference between correlations of true person effects and the correlations of estimated person effects (y-axis), generated by simulations. The data considers 20 different levels of correlations $Corr(\theta_{i1}, \theta_{i2})$ = -0.95, -0.85, ..., 0.75, 0.85, 0.95 between person effects θ_i estimated from two-way fixed-effects models of two different dependent variables y_1 and y_2. Both graphs are based on 20,000 iterations, stratified by 20 levels of correlations, 10 boardsizes, 10 mobility shares and 10 iterations of each boardsize-mobility-correlation-combination. The 10 levels of mobility are stratified by three levels of low mobility (mover shares from 0.05 to 0.25), four levels of medium mobility (mover shares from 0.35 to 0.65) and three levels of high mobility (mover shares from 0.75 to 0.95). The underlying datasets generated for each iteration consist of 100 firms, observed over 15 years.
The DGPs are specified as follows
(1) Sum of person effects: $y_{j(i)t} = (\sum_{i=1}^{n_{jt}} \theta_{i \in (jt)}) + \psi_j + u_{jt}$
(2) Mean of person effects: $y_{j(i)t} = (\frac{1}{n_{jt}} * \sum_{i=1}^{n_{jt}} \theta_{i \in (jt)}) + \psi_j + u_{jt}$
Person effects θ_i, firm effects ψ_j and error terms u_{jt} as specified in the DGP are drawn from independent standard normal distributions.