Empirical Essays on Household Finance and Incentives

Martin G. Becker

Doctoral Thesis

submitted to

Justus-Liebig-University Gießen,

Department of Business Administration and Economics

Supervisors:

Prof. Dr. Andreas Walter

Prof. Dr. Oscar Stolper

Date of submission:

13th of February 2026

Justus-Liebig-University Gießen | Department of Financial Services

Licher Str. 74 | 35394 Gießen

E-mail: Martin.Becker@wirtschaft.uni-giessen.de

Contents

List of Tables, Figures and Appendices **iv**

I. Investing by Example: Leveraging Peer Information in Digital Banking **I-1**

1. Introduction I-3
2. Experimental Design and Data I-8
3. Results I-12
4. Discussion and Conclusion I-21
5. References I-25
6. Appendix I-28

II. From Branches to Browsers: A Comparative Analysis of Direct and Traditional Bank Clients in Germany **II-35**

1. Introduction II-37
2. Retail Banking Market in Germany II-40
3. Data II-41
4. Empirical Results II-47
5. Robustness Checks II-58
6. Discussion and Implications II-63
7. References II-68
8. Appendix II-71

III. Anlageverhalten und Kundenprofile im Vergleich: Unterschiede zwischen Sparkassen, Genossenschaftsbanken und Großbanken **III-78**

1. Einführung III-80
2. Institutionelle Rahmenbedingungen und Forschungsfragen III-81
3. Daten und deskriptive Statistik III-84
4. Empirische Ergebnisse III-87
5. Fazit III-100
6. Literaturverzeichnis III-101
7. Appendix III-103

IV. Financial Support Among Siblings: The Relevance of Personal and Family Characteristics **IV-104**

1. Introduction IV-106
2. Theoretical and Empirical Background IV-108
3. Methods IV-119
4. Results IV-126
5. Discussion IV-134

6. References IV-139
7. Appendix IV-145
V. Perception vs. Reality: Wealth Disparities Between Urban and Rural Households in Germany V-151
1. Introduction V-153
2. Data and Methodology V-157
3. Results V-164
4. Discussion and Conclusion V-176
5. References V-179
6. Appendix V-186
VI. The Power of ESG Transparency: The Effect of the New SFDR Sustainability Labels on Mutual Funds and Individual Investors VI-197
1. Introduction VI-199
2. Data VI-202
3. The Influence of the EU Directive on Sustainability Scores VI-204
4. Impact on Investors VI-206
5. Conclusion VI-207
6. References VI-209
VII. When a B Becomes an A: Causal Evidence on the Effects of a Journal Ranking Update on Academics' Publication Behavior VII-210
1. Introduction VII-212
2. Setting and Identification VII-214
3. The Effects of a Journal Upgrade VII-216
4. Data and Methodology VII-217
5. Results VII-223
6. Robustness Tests and Additional Analyses VII-227
7. Discussion and Conclusion VII-235
8. References VII-240
9. Appendix VII-243
Affidavit ix

List of Tables, Figures and Appendices

Tables

Table I-1: Descriptive Statistics I-12

Table I-2: Summary Stats by Treatment Group I-12

Table I-3: Outcome by Treatment Group I-13

Table I-4: Regression on uCTR I-15

Table I-5: Heterogeneous Treatment Effects on uCTR I-17

Table I-6: Effect of Channel on eCTR I-20

Table II-1: Descriptive Statistics II-45

Table II-2: Correlations II-46

Table II-3: Differences in Client Characteristics II-47

Table II-4: Differences between Generations II-48

Table II-5: Logistic Regression on Direct Bank II-50

Table II-6: Differences in Investment Behavior II-51

Table II-7: Regression on Investment Variables II-53

Table II-8: Regression on Investment Variables II-56

Table II-9: Regression on Investment Variables with Matched Sample II-60

Tabelle III-1: Deskriptive Statistiken III-87

Tabelle III-2: Charakteristika nach Bankengruppe III-90

Tabelle III-3: Anlage nach Bankengruppe III-92

Tabelle III-4: Aktienmarktpartizipation III-95

Tabelle III-5: Einlagenquote III-97

Tabelle III-6: Aktienquote III-99

Table IV-1: Robustness of Variable Selection for Predictors of Received Sibling Support IV-131

Table V-1: Descriptive Statistics V-160

Table V-2: Correlation Matrix V-163

Table V-3: Linear Regression on Net Wealth (k) V-165

Table V-4: Linear Regression on Net Wealth Decile Misperception V-169

Table VI-1: Summary Statistics VI-203

Table VI-2: The Influence of the EU Directive on Sustainability Scores VI-205

Table VI-3: The Influence of the EU Directive on Fund Flows VI-207

Table VII-1: Descriptive Statistics VII-220

Table VII-2: Panel Regressions – Main Effect VII-224

Table VII-3: Panel Regressions – Ranking Heterogeneity VII-226

Table VII-4: Panel Regressions – Main Effect – Robustness Check (U.S. Based Authors) VII-228

Table VII-5: Panel Regressions – Ranking Heterogeneity – Robustness Checks (U.S. Based Authors) VII-229

Table VII-6: Panel Regressions – Main Effect – Robustness Check (Journals Upgraded in the Academic Journal Guide) VII-230

Table VII-7: OLS Regressions – Researcher Heterogeneity VII-234

Figures

Figure I-2: Peer Information Nudge in Online Banking Channel for Women Aged 26-30 I-28

Figure I-3: Neutral Prompt in Online Banking Channel for Women Aged 26-30 I-29

Figure I-4: Peer Information Nudge in the Mobile App Channel for Women Aged 26-30 I-30

Figure IV-1: Determinants of Sibling Financial Support (LPM Estimates) IV-127

Figure IV-2: Determinants of Sibling Financial Support by Gender (LPM Estimates) IV-132

Figure VII-1: Estimated Effect of the Update of the VHB-Journal-Ranking VII-222

Figure VII-2: SJR Score of all upgrades VII-232

Appendices

Appendix I-A: Examples of the Interventions I-28

Appendix I-B: Reference Values per Peer Group I-31

Appendix I-C: Outcome by Treatment Group for the Full Sample I-32

Appendix I-D: Investment Plan Sign-up Rate by Treatment Group I-33

Appendix I-E: Causal Forest Calibration Test I-34

Appendix II-A: Variable Dimensions and Descriptions II-71

Appendix II-B: Post Matching Diagnostic Test II-72

Appendix II-C: Logistic Regression on Direct Bank for Single-Person Households II-73

Appendix II-D: Regression on Investment Variables for Single-Person Households II-74

Appendix II-E: Logistic Regression on Stock Market Participation II-75

Appendix II-F: Logistic Regression on Direct Bank for the Extended Sample II-76

Appendix II-G: Regression on Investment Variables for the Extended Sample II-77

Appendix III-A: Variablen Definition und Ausprägung III-103

Appendix IV-A: Sample Selection Process IV-145

Appendix IV-B: Summary Statistics IV-146

Appendix IV-C: LPM Regressions Predicting the Receipt of Financial Support from a Sibling IV-147

Appendix IV-D: LPM Regressions Predicting the Receipt of Financial Support from a Sibling by Gender IV-148

Appendix IV-E: Logistic Regressions Predicting the Receipt of Financial Support from a Sibling IV-149

Appendix V-A: Variable Dimensions and Descriptions V-186

Appendix V-B: Mean Differences Between Urban and Rural Households V-187

Appendix V-C: Linear Regression on Net Wealth with Matched Sample V-188

Appendix V-D: Linear Regression on Net Wealth Decile Misperception with Matched Sample V-189

Appendix V-E: Linear Regression on Net Wealth – Extended Sample Incl. 20–100k Municipalities V-190

Appendix V-F: Linear Regression on Net Wealth Decile Misperception – Extended Sample Incl. 20–100k Municipalities V-191

Appendix V-G: Linear Regression on Net Wealth – Spatial Extremes Sample (<5k vs. >500k Municipalities) V-192

Appendix V-H: Linear Regression on Net Wealth Decile Misperception – Spatial Extremes Sample (<5k vs. >500k Municipalities) V-193

Appendix V-I: Linear Regression on Net Wealth – Single Household Sample V-194

Appendix V-J: Linear Regression on Net Wealth Decile Misperception – Single Household Sample V-195

Appendix V-K: Linear Regression on Net Wealth Decile Misperception – Interquartile Sample V-196

Appendix VII-A: Event Study – Test for Parallel Trends VII-243

Appendix VII-B: Panel Regressions – Main Effect – Robustness Check (France Based Authors) VII-244

Appendix VII-C: Panel Regressions – Ranking Heterogeneity – Robustness Checks (France Based Authors) VII-245

Appendix VII-D: Panel Regressions – Ranking Heterogeneity – Upgrades to A+ VII-246

Appendix VII-E: Overlap of Ranking Changes between the VHB-journal-ranking and the AJG VII-247

I. Investing by Example: Leveraging Peer Information in Digital Banking

Co-authors:

Oscar Stolper, Andreas Walter

Own share:

60%

This article has received a Revise and Resubmit from:

Journal of Behavioral and Experimental Economics

Previous versions of this paper have been presented at the following conferences and workshops:

- 19th International Behavioural Finance Conference, 2025, Chicago, United States of America

Investing by Example: Leveraging Peer Information in Digital Banking

Martin G. Becker[a] Oscar Stolper[b] Andreas Walter[c]

Abstract: We show that a subgroup peer information nudge can significantly increase the initial engagement with investing among online and mobile bank clients. In a randomized controlled field experiment involving over 50,000 bank clients, we provided participants with descriptive peer information, namely the monthly investment amount that three-quarters of investing individuals in their age- and gender-matched peer group exceed. We find that this intervention increases the user-level click-through rate for additional information by 23%, although there is no corresponding increase in the sign-up rate for investment plans. Our findings reveal a dichotomy: while peer information successfully increases initial engagement, it is insufficient to overcome the frictions required to change actual investment behavior.

JEL Codes: C93, D14, G11, G41, G51

*Keywords***:** digital nudging, field experiment, peer information, investment behavior, retail investing

[a] martin.becker@wirtschaft.uni-giessen.de, Justus Liebig University Gießen, Department of Financial Services, Licher Straße 74, 35394 Gießen, Germany

[b] oscar.stolper@wiwi.uni-marburg.de, Philipps University Marburg, Behavioral Finance Research Group, Am Plan 1, 35032 Marburg, Germany

[c] andreas.walter@wirtschaft.uni-giessen.de, Justus Liebig University Gießen, Department of Financial Services, Licher Straße 74, 35394 Gießen, Germany

1. Introduction

Persistent underinvestment in equities poses a continuing challenge for household wealth accumulation and retirement preparedness. Policymakers and financial institutions have explored various interventions – including default enrollment (e.g., Choi et al., 2002; Madrian & Shea, 2001), tax incentives (e.g., Chetty et al., 2014; Duflo et al., 2006), and targeted messaging (e.g., Duflo & Saez, 2003; Goda et al., 2014) – to boost individuals' engagement with the stock market. At the same time, peer information strategies have been shown to alter behaviors in domains such as energy conservation (e.g., Allcott, 2011), tax compliance (e.g., Hallsworth et al., 2017), and charitable giving (e.g., Shang & Croson, 2009). This suggests social norms could be a potent lever for encouraging increased investment. Yet empirical studies applying these insights to savings and investment decisions report mixed results (e.g., Bauer et al., 2022; Beshears et al., 2015; Døskeland et al., 2025; Dur et al., 2021; Kast et al., 2018), raising the question of whether and under what conditions peer information nudges might effectively motivate individuals to invest.

In this paper, we address this gap by conducting a large-scale randomized controlled field experiment with over 50,000 clients at a German regional bank. Specifically, we test whether providing descriptive peer information – highlighting the monthly contribution levels of existing investors – can enhance both the initial engagement with investing (click-through rate) and actual investment behavior (investment take-up rate). By examining these factors separately, we aim to pinpoint the conditions under which peer information nudges effectively translate into real changes in financial decision-making.

We conducted a randomized controlled field experiment in collaboration with a regional German bank to examine whether providing peer information can influence clients' behavior. Specifically, our sample consists of over 50,000 clients who did not previously have an investment plan at the bank. These clients were randomly assigned to one of two groups, ensuring that any observed differences in outcomes could be attributed to our intervention rather than to pre-existing differences between individuals. We then examined the effect of the message “Did you know that 3 out of 4 of our female/male investment plan clients in your age group invest [Placeholder]€ or more each month?”[1] on the Treatment group. The Control group received a neutral prompt, which informed them of the possibilities of investing at the bank. Both the peer information nudge and the neutral prompt were implemented in the digital channels of the bank.[2]

Our main findings show that clients who received the peer information nudge are 23% more likely to click through for additional details, demonstrating that descriptive peer information can enhance initial engagement. However, we detect no accompanying increase in actual investment plan sign-ups, indicating that a higher click-through rate does not necessarily translate into greater uptake of investment products. Thus, while peer information nudges can generate immediate interest, complementary measures may be required to convert that interest into actual investment behavior. Further analysis suggests that older clients and users of the mobile app exhibit higher overall click-through tendencies. However, these results do not differ between the Treatment and the Control group.

These mixed results resonate with the existing literature, which paints a complex picture of peer effects in financial decision-making: while social interactions undeniably shape financial behavior, targeted peer-based interventions yield highly variable outcomes. On one hand, observational studies confirm that social proximity drives investment participation. Hong et al. (2004) demonstrate that households interacting more frequently with neighbors are significantly more likely to invest in the stock market, an effect amplified in states with higher overall participation. Similarly, Ouimet and Tate (2020) find that coworkers' behavior substantially influences individual participation in employee stock purchase programs. Beyond

[1] The exact values for each peer group can be found in Appendix I-B.
[2] Illustrative examples of the provided nudges can be found in Appendix I-A.

investing, peer effects have shown potency in broader contexts, such as insurance choices (Hu, 2022; Sorensen, 2006) and retirement savings participation (Duflo & Saez, 2002).

However, the literature also clarifies that this influence is neither universal nor easily harnessed. Lieber and Skimmyhorn (2018), for instance, find peer influence to be a strong driver of charitable giving among U.S. Army soldiers, yet notably absent in their retirement savings or life insurance decisions. This suggests that peer effects may not uniformly translate across different types of financial choices.

Furthermore, when researchers attempt to apply these insights through targeted information interventions, the results often fail to reproduce the positive findings of observational studies. Beshears et al. (2015) find that providing information on coworkers' participation in 401(k) plans can, counterintuitively, reduce savings for nonparticipants, highlighting the potential for negative effects if descriptive norms seem unachievable. Bauer et al. (2022) find no significant effect of peer information on retirement savings, while Dur et al. (2021) observe that although such nudges boost short-term interest – evidenced by a higher click-through rate and more website visits – they do not translate into increases in actual savings. Meanwhile, Kast et al. (2018) provide evidence of a successful peer information intervention delivered via text messages that significantly increased savings deposits among microcredit clients in Chile. Similarly, Døskeland et al. (2025) find that descriptive norm nudges delivered through an online banking interface successfully increased actual retirement savings, albeit with small effect sizes. This inconclusive pattern underscores the need to clarify the conditions under which peer information nudges succeed or fail in prompting meaningful financial actions.

To that end, we draw on two complementary psychological mechanisms that explain how peer information can influence behavior. First, it may act as a form of social learning (Bandura, 1977; Bursztyn et al., 2014; Ellison & Fudenberg, 1993), providing individuals with information they previously lacked. For example, a person who had not considered investing at their bank may revise their beliefs about the viability or popularity of that option upon learning that others are doing so. In this way, descriptive peer information can function as a heuristic signal of product relevance or trustworthiness.

Second, peer information may trigger a desire to conform to perceived social norms (Bicchieri & Dimant, 2022; Cialdini et al., 1990; Schultz et al., 2007). Learning that others – especially demographically similar peers – are behaving in a certain way can increase psychological pressure to align one's behavior with that norm, even absent new factual information. The mere awareness that a behavioral pattern exists within one's peer group can therefore elicit normative compliance, particularly when the action is framed as common, appropriate, or expected.

Both mechanisms suggest that peer information should at least increase initial engagement (e.g., information-seeking), with downstream behavior depending on how actionable and attainable the referenced norm appears. We empirically test this distinction to determine if social signals are strong enough to not just spark initial interest but to overcome the frictions that prevent financial commitment.

Our study contributes to the literature in multiple ways. First, we extend the discussion on the effectiveness of peer information nudges in saving and investing, particularly in contexts where existing evidence has yielded conflicting results (e.g., Bauer et al., 2022; Kast et al., 2018). By separately examining initial engagement and final take-up in a randomized controlled field experiment, we help clarify these mixed findings. We provide evidence of a significant disconnect between engagement and action, showing that while peer information is effective at generating initial engagement (increasing click-through rates by 23%), it possesses little marginal power to overcome the frictions required to finalize an investment plan in our setting. In doing so, we provide empirical support for recent arguments that nudges often succeed at changing proximate behaviors but might be ineffective at affecting downstream economic results (e.g., Beshears & Kosowsky, 2020; Guttman-Kenney et al., 2025).

Second, we add to the growing body of work on digital nudging (e.g., Mirsch et al., 2017; Weinmann et al., 2016) – where strategic prompts or design choices are integrated into online or mobile interfaces (Meske & Potthoff, 2017; Valta & Maier, 2025) – by investigating how nudges perform across different devices and channels. While prior work suggests mobile users may be more impulsive (e.g., Mograbi, 2022), our encounter-level analysis reveals that although the mobile app drives significantly higher baseline engagement, the incremental impact of the peer information nudge remains stable across both mobile and online banking channels.

Third, we shed light on the implications of displaying a relatively modest descriptive norm (e.g., Carattini & Blasch, 2024). Rather than emphasizing how much the "average" client invests, our intervention highlighted the reachable 25th-percentile contribution of existing investors in the peer group. This design choice was explicitly intended to make the norm appear attainable without justifying inertia (Bicchieri & Dimant, 2022). The finding that even this attainable norm failed to increase investment uptake highlights the limits of purely informational nudges. It suggests that optimizing the content of the social signal may not be enough to change financial behavior on its own, implying that future interventions might benefit from simultaneously reducing transactional frictions.

The remainder of this paper is structured as follows. Section 2 discusses the experimental design and data. Section 3 presents our main results, explores heterogeneous treatment effects, and examines whether the channel of implementation affects the nudge's effectiveness. Section 4 concludes with a discussion on the broader implications of our findings and suggests directions for future research.

2. Experimental Design and Data

Our experiment was conducted in collaboration with a regional German bank as part of a digital campaign delivered via online and mobile banking channels. The campaign ran for a total of three months starting in May 2023. The primary objective of this campaign was to encourage clients without an investment plan to enroll in monthly investment plans.[3] In agreement with the bank, we decided to implement a peer information nudge.

A significant challenge when applying peer effects theory to financial nudges is the anonymity and perceived distance of referenced peer groups. Whereas neighbors (e.g., Hong et al., 2004) and close coworkers (e.g., Ouimet & Tate, 2020) are identifiable individuals whose behaviors are directly observable, bank clients (e.g., Bauer et al., 2022; Dur et al., 2021) referenced in digital nudging typically appear as anonymous aggregates. According to social-comparison theory (Festinger, 1954; Raineau et al., 2025), the influence of social norms diminishes as social distance grows and strengthens as perceived similarity and psychological proximity increase. This difference in perceived relatedness likely contributes to the differences in effectiveness observed in peer information interventions across different contexts (e.g., Bauer et al., 2022; Kast et al., 2018).

Although we cannot remove anonymity among bank clients, we attempted to enhance perceived similarity and connection through deliberate intervention design choices. Specifically, we segmented peer groups according to age and gender[4], thereby increasing the demographic similarity between message recipients and the referenced peers (Kang & Chung, 2017; McPherson et al., 2001). Furthermore, to reinforce psychological proximity and identification, each prompt was accompanied by a representative photo depicting a typical individual from the respective peer group. These design choices aimed to mitigate the inherent

[3] The prompts were also shown to clients with existing investment-plan contributions below €50 per month. Contributions at this level are considered too small to constitute active participation in the bank's investment offerings, so these clients were treated as functionally similar to non-investors for the purposes of the campaign.
[4] For the purpose of this study, we rely on the binary classification (male/female) recorded in the bank's administrative database. We use this legal sex marker as a proxy for gender throughout the analysis.

anonymity of digital banking contexts and to strengthen the influence of peer information through increased relatability (Cyr et al., 2009; Hassanein & Head, 2007).

Given that only a small proportion of the bank's clientele already had an investment plan, directly referencing low participation rates might have undermined the effectiveness of a descriptive norm nudge (Carattini & Blasch, 2024). To address this concern, instead of omitting any mention of baseline rates altogether, we focused on a narrower reference group – clients who already maintained an investment plan at the bank. By restricting the reference group to existing investors, we avoided signaling the low overall participation rate that can legitimize inertia. While acknowledging that this subgroup does not precisely represent the broader client base, we believe the approach remains valid and transparent, clearly informing clients that the statistics provided reflect investors specifically.

Furthermore, to make sure that our nudge did not end up discouraging clients (e.g., Beshears et al., 2015), we decided not to use the median investment value. Instead, for each subgroup, we calculated the 25th percentile of monthly investment amounts, representing the figure that three-quarters of investors in that subgroup exceeded. These values formed the basis of our treatment message.

Clients in the Treatment group saw the message: "Did you know that 3 out of 4 of our female/male investment plan clients in your age group invest [Placeholder]€ or more each month?"[5] In contrast, the Control group received a more neutral prompt, which informed them about the opportunities to invest at the bank. Framing the statistic as "3 out of 4" and tying it to the 25th-percentile contribution set a threshold most recipients would view as reachable, mitigating the backfire risk observed when norms feel unattainable.

Except for the peer information text, all elements were held constant between the Treatment and Control groups within each age–gender subgroup. For example, a given subgroup saw the

[5] Since the clients were German, they received the message: "Wussten Sie, dass 3 von 4 unser Sparplankunden(-innen) in Ihrer Altersgruppe monatlich [Placeholder]€ oder mehr investieren?" which means the same in German.

same demographically tailored image in both treatment arms, though images differed across subgroups. Across channels, the content was identical; only the standard platform layout varied slightly between online banking and mobile banking.[6]

The prompts were presented each time a client logged in, via the online or mobile channel, until the earliest of the following stopping criteria was met: (i) the client clicked the link; (ii) the client had accumulated 12 prompt exposures in that channel; or (iii) the three-month intervention period elapsed. Exposures were tracked separately for the online and mobile channels.

Initially, 112,456 bank clients who did not have an investment plan were randomly assigned with equal probability (simple randomization) to either the Treatment or Control group. Subsequently, the bank removed clients without online or mobile banking access and those with negative account balances, applying identical exclusion criteria to both groups, leaving 78,413 eligible clients who were scheduled to receive the prompts.

In our analysis, we further excluded clients who did not access their online or mobile banking accounts at all during the treatment period. Additionally, we excluded clients without reported income or an account balance of zero prior to the treatment period, as these clients are likely to use a different bank for daily banking activities. This results in a final sample of 52,917 clients.

To capture how clients responded to the intervention, we focus on two key outcome measures: the user-level click-through rate ($uCTR$)[7] and the investment take-up rate (ITR). $uCTR$ represents the proportion of clients who engaged with the campaign by clicking the provided link, directing them to a webpage detailing the process of initiating an investment plan. Subsequently, clients could contact the bank to establish an investment account or

[6] Appendix I-A contains examples of the nudges presented to the clients.

[7] Ordinarily click-through rates are defined as the number of clicks divided by the number of impressions. However, for our setting we are more interested in share of clients having clicked on the link. We therefore define the measure user-level click-through rate which is defined as the number of clients who have clicked on the link divided by the total number of clients in the sample.

alternatively set up an investment account online. Only after opening the investment account could they set up an investment plan.[8]

ITR encompasses any client who signed up for an investment plan during the treatment period, as well as any client who invested directly in assets also available via an investment plan. This broader definition accounts for the possibility that the intervention could prompt general interest in investing, leading some clients to invest directly rather than setting up a recurring plan. Because ITR records every investment executed during the observation window – regardless of what prompted it – it also inevitably captures actions unrelated to the intervention. However, this noise affects both the Treatment and Control groups in the same way and thus should not lead to systematic distortions. Both $uCTR$ and ITR are treated as binary indicators at the client-level – once a client takes a particular action (e.g., clicks on the link, signs up for any plan), they are counted as 1 for that outcome, regardless of the subsequent behavior. This approach to defining outcomes allows us to distinguish between two critical stages of client engagement: the initial decision to learn more and the subsequent decision to make an actual investment.

Table I-1 presents the descriptive statistics for the final sample. Overall, men and women are nearly equally represented. The average client is 43.8 years old (median = 42), with a mean monthly net income of €2,670 (median = €2,270) and an average account balance of €27,150 (median = €6,540). Around 17.1% of clients in our sample already hold investment funds. Across the entire sample, approximately 2.0% of clients clicked on the link for more information, and 0.5% either signed up for an investment plan or made a direct investment.

[8] This step is of course not necessary for clients who already have an investment account at the bank but currently do not have any active investment plan.

Table I-1: Descriptive Statistics

Variable	Mean	25%-Quantile	Median	75%-Quantile	N
Age	43.80	30.00	42.00	57.00	52,917
Male	49.85%	0.00	0.00	1.00	52,917
Monthly Income (€k)	2.67	1.43	2.27	3.21	52,917
Account Balance (€k)	27.15	2.02	6.54	23.27	52,917
Fund Owner	17.06%	0.00	0.00	0.00	52,917
uCTR	2.02%	0.00	0.00	0.00	52,917
ITR	0.50%	0.00	0.00	0.00	52,917
Treatment	50.36%	0.00	1.00	1.00	52,917

Notes: This table presents the summary statistics of our dataset. Age indicates the age of the client measured in years at the start of the treatment period. $Male$ is a dummy variable equal to 1 if the client is male and 0 otherwise. $Monthly\ Income$ (in thousands of euros) is the average monthly income of the client into the bank accounts at the bank over the last three months before the start of the treatment period. $Account\ Balance$ (in thousands of euros) is the client's account balance in all his accounts with the bank at the start of the treatment period. $Fund\ Owner$ is a dummy variable indicating whether the client already owns any kind of investment funds at the bank. $uCTR$ is a dummy variable indicating whether the client clicked on the link to obtain additional information, and ITR is a dummy variable indicating whether the client created an investment plan or otherwise purchased funds during the treatment period. $Treatment$ is a dummy variable equal to one for all clients who received the peer information nudge.

Table I-2 presents the descriptive statistics for both groups, showing minor imbalances in monthly income (+€60) and fund ownership (+0.8%) for the Treatment group. Therefore, we include these variables as controls in all regressions. Other covariates do not differ significantly.

Table I-2: Summary Stats by Treatment Group

	Treatment N=26,649	Control N=26,268	Difference
Age	43.8321	43.7630	0.0691
Male	49.93%	49.77%	0.16%
Monthly Income	2.7040	2.6438	0.0601**
Account Balance	27.4674	26.8357	0.6316
Fund Owner	17.43%	16.67%	0.76%**

Notes: This table reports the mean of each of the control variables for the Treatment and Control groups. The difference column reports the mean difference across groups, with significance assessed via a chi-square test of independence for binary variables and via a two-sided Welch two-sample t-test for continuous variables. Significance levels for p-values < 0.10, 0.05, and 0.01 are denoted by *, **, and ***, respectively.

3. Results

3.1 Baseline Results

To gauge the effectiveness of our peer information nudge, we first examine $uCTR$ and ITR between the Treatment and Control groups. Table I-3 provides the means and standard deviations of these outcomes for both groups. We formally test for differences in proportions using chi-square tests for binary variables.

Table I-3: Outcome by Treatment Group

	Treatment N=26,649	Control N=26,268	Difference
uCTR	2.23%	1.82%	0.41%***
	(14.75%)	(13.37%)	
ITR	0.49%	0.52%	-0.03%
	(6.97%)	(7.18%)	

Notes: This table reports the mean of each outcome for the Treatment and Control groups, with standard deviations in parentheses. $uCTR$ is a dummy variable indicating whether the client clicked the link for more information. ITR is a dummy variable for whether the client signed up for an investment plan or directly invested in assets also available through the plan during the treatment period. The difference column reports the mean difference in percentage points across groups, with significance assessed via a chi-square test of independence. Significance levels for p-values < 0.10, 0.05, and 0.01 are denoted by *, **, and ***, respectively.

We find that $uCTR$ is notably higher in the Treatment group (2.23%) than in the Control group (1.82%), reflecting a statistically significant and economically meaningful relative increase of approximately 23%. In contrast, ITR shows no significant difference between groups, with 0.49% in the Treatment group versus 0.52% in the Control group.[9] These results suggest that although the peer information nudge effectively increases initial engagement with investing, it does not translate into meaningful changes in actual investment behavior.[10]

Thus, we provide mixed evidence supporting peer information's potential to drive initial engagement but simultaneously highlight its limitations in influencing concrete financial decisions. These findings align closely with those of Dur et al. (2021), who similarly report that peer information nudges can increase short-term engagement with saving (evidenced by a higher click-through rate) without necessarily affecting actual behavior. Conversely, our findings diverge from studies reporting tangible improvements in savings and investment behaviors following peer nudges (e.g., Døskeland et al., 2025; Kast et al., 2018). A plausible explanation for the null result is that peer information by itself is not strong enough to overcome

[9] As a robustness check we repeated the analysis including all 78,413 bank clients who were scheduled by the bank to receive the prompt. This ensures the effect is not driven by us excluding clients who did not access their banking during the treatment period or were exclude from the analysis due to a lack of income and or a reported account balance of 0. While the overall magnitude of $uCTR$ and ITR is naturally smaller, the difference in $uCTR$ remains highly significant with a 23% higher $uCTR$ for the Treatment Group. The difference in ITR remains insignificant. The results are displayed in Appendix I-C.

[10] As an additional robustness check, we further perform the same test with the Investment plan sign-up rate which only includes clients which set up an investment plan in the treatment period. The result remain virtually unchanged and are tabulated in Appendix I-D.

the cognitive and transactional frictions that keep clients from starting an investment plan (see also Bauer et al., 2022; Beshears et al., 2015; Dur et al., 2021).[11]

In the following sections, we examine whether these engagement gains persist when controlling for demographic characteristics (Section 3.2), if the peer information affects different demographic subgroups differently (Section 3.3), and explore whether the prompts operate differently in online vs. mobile channels (Section 3.4).

3.2 The Impact of Demographic Factors on the Click-Through Rate

To assess whether the observed treatment effect holds after accounting for demographic and financial characteristics, we employ a standard logistic regression in which the dependent variable is $uCTR$.

[11] Alternatively, the results could also be explained by a lack of statistical power. With a baseline ITR of only 0.50%, even a moderate true effect would require a considerably larger sample to be detected with conventional power. With a baseline ITR of 0.50% and roughly 26,400 observations per arm, our 80 %-powered minimum detectable effect at α=0.05 is approximately 0.19 percentage points (from 0.50% → 0.69 %). Detecting smaller effects, e.g., a 0.10 percentage point increase would require nearly 86,000 clients per group. These figures illustrate that, given the low baseline, our sample is underpowered to reject small but potentially meaningful treatment effects.

Table I-4: Regression on uCTR

	(1) Simplified Model	(2) Full Model
Treatment	1.2279***	1.2323***
	(0.0762)	(0.0766)
Age		1.0042**
		(0.0020)
Male		1.3239***
		(0.0852)
Monthly Income		0.9087***
		(0.0190)
Account Balance		1.0001
		(0.0005)
Fund Owner		0.8571*
		(0.0751)
Constant	0.0185***	0.0172***
	(0.0009)	(0.0018)
Observations	52,917	52,917
Pseudo R^2	0.0010	0.0055

Notes: This table reports odds ratios from two logistic regressions with $uCTR$ as the dependent variable (1 = client clicked the link). Specification (1) includes only the $Treatment$ dummy (1 = client received the peer nudge) and a constant, while specification (2) adds age, $gender$, $monthly\ income$, $account\ balance$ and $fund\ ownership$ as controls. Heteroskedasticity-robust standard errors are shown in parentheses. Significance levels for p-values < 0.10, 0.05, 0.01 are denoted by *, **, ***, respectively. The Pseudo R^2 reported is the McFadden R^2.

In the first specification, we include only the $Treatment$ indicator and a constant. We find a significant positive effect: the odds ratio indicates that treated clients have 22.79% higher odds of clicking the link compared to those in the Control group.

In the second specification, we add controls for age, gender, monthly income, account balance at the beginning of the treatment period, and ownership of investment funds at the bank. The treatment effect remains positive, statistically significant, and virtually unchanged in magnitude, suggesting the results are robust to these additional factors.

With respect to demographic factors, older and male clients show a higher propensity to click on the link than younger and female clients. By contrast, higher monthly income and pre-existing fund ownership both reduce the likelihood of clicking. One plausible explanation is that higher-income clients may have already made a more deliberate decision regarding investing – possibly at another institution – making them less responsive to the prompts.

Similarly, clients who own funds at the bank are presumably already aware of investment options, reducing their need to click for further information.

3.3 Heterogeneous Treatment Effects

Beyond our primary analysis, we also investigate whether client characteristics influence the impact of the peer information nudge. Previous studies suggest that peer information interventions may be especially effective among younger, less financially experienced individuals, who often rely on social cues to make investment decisions (see, for example, Bursztyn et al., 2014; Ouimet & Tate, 2020).

To explore potential heterogeneity, we again employ a logistic regression with $uCTR$ as the dependent variable, this time including interaction terms between $Treatment$ and Age, as well as between $Treatment$ and $Gender$. In the baseline specifications, we include only the intercept, the $Treatment$ dummy, the demographic variable (Age or $Gender$), as well as the relevant interaction term. Next, we add the control variables from Table I-4 for a more comprehensive specification.

Table I-5: Heterogeneous Treatment Effects on uCTR

	(1) Simplified Model	(2) Full Model	(3) Simplified Model	(4) Full Model
Treatment	1.1402	1.1394	1.3616***	1.3652***
	(0.2115)	(0.2027)	(0.1268)	(0.1272)
Age	1.0012	1.0032		1.0042**
	(0.0029)	(0.0028)		(0.0020)
Treatment*Age	1.0017	1.0018		
	(0.0039)	(0.0038)		
Male		1.3242***	1.3750***	1.4676***
		(0.0852)	(0.1284)	(0.1385)
Treatment*Male			0.8289	0.8306
			(0.1037)	(0.1039)
Monthly Income		0.9087***		0.9087***
		(0.0190)		(0.0190)
Account Balance		1.0001		1.0001
		(0.0005)		(0.0005)
Fund Owner		0.8572*		0.8572*
		(0.0751)		(0.0751)
Constant	0.0176***	0.0179***	0.0156***	0.0162***
	(0.0024)	(0.0024)	(0.0011)	(0.0018)
Observations	52,917	52,917	52,917	52,917
Pseudo R^2	0.0012	0.0056	0.0024	0.0058

Notes: This table reports odds ratios from four logistic regressions with $uCTR$ as the dependent variable (1 = client clicked the link). Columns (1) and (2) examine heterogeneous treatment effects by age, while columns (3) and (4) examine heterogeneous treatment effects by gender. In each pair, the specification with the simplified model includes only the $Treatment$ dummy, the heterogeneity variable, and their interaction as well as a constant; The full model adds the other heterogeneity variable ($male$ and age respectively), as well as $monthly\ income, account\ balance$, and $fund\ ownership$ as controls. Heteroskedasticity-robust standard errors are shown in parentheses. Significance levels for p-values < 0.10, 0.05, 0.01 are denoted by *, **, ***, respectively. The Pseudo R^2 reported is the McFadden R^2.

Our results indicate no statistically significant interaction between the $Treatment$ and either Age or $Gender$. In other words, while overall click-through rates vary according to demographic factors, the incremental effect of the peer information nudge itself appears to be independent of these attributes. This suggests that although client characteristics shape general responsiveness to a digital campaign, the specific impact of peer information does not differ substantially by age or gender.

To ensure we do not overlook heterogeneity driven by complex, nonlinear combinations of characteristics, we additionally employ causal forests following the methodology of Athey and Wager (2019). This nonparametric approach allows us to observe heterogeneity regarding potential subgroups without restricting the analysis to pre-specified interaction terms. As demonstrated in recent household finance literature (e.g., Burke et al., 2023; Medina & Pagel, 2025), causal forests are particularly effective at identifying individuals with the largest response to a treatment while mitigating the risk of invalid inference due to overfitting.

We train a generalized random forest on the full set of pretreatment covariates. To assess the robustness of the heterogeneity, we conducted an omnibus test for the presence of heterogeneity based on the best linear fit of the forest's predictions.

The results of this calibration test, reported in Appendix I-E, strongly fail to reject the null hypothesis of homogeneous treatment effects. While the "mean forest prediction" remains significant – indicating the model correctly captures the average treatment effect – the "differential forest prediction" is not statistically significant, implying that the treatment effect applies uniformly across the population.

Consequently, we conclude that the peer information nudge exerts a consistent positive influence on engagement across the diverse client base, rather than being driven by specific subgroups such as the "younger, less experienced" investors initially hypothesized.

3.4 Channel Effects

Thus far, our analysis has examined the intervention at the client level, where the dependent variable $uCTR$ equals 1 if a client ever clicked the link during the three-month campaign. This allowed us to estimate the effectiveness of the peer information nudge and to investigate which demographic factors might influence the response.

However, we now want to investigate whether the channel (online banking or mobile app) through which a client encounters the prompt influences responsiveness. Since the same client could access their banking through both the mobile app and traditional online banking, we cannot study this effect using client-level data. Therefore, we now turn to the full dataset of all encounters. This encounter-level dataset contains more than 670,000 individual login-level observations. For every encounter with the prompt, we define the encounter-level click-through rate ($eCTR$) as 1 when the click occurred during this specific encounter and 0 otherwise. This approach enables us to estimate the model for each individual encounter while simultaneously allowing us to distinguish between encounters across different channels. Notably, the content of the peer information message remained identical for each individual across all encounters and channels. Concretely, clients in the Treatment group always saw the same statement highlighting that "3 out of 4 of our female/male investment plan clients in your age group invest at least [Placeholder]€ per month", regardless of whether they accessed their banking through the app or online banking. Meanwhile, clients in the Control group always received the same neutral prompt.

Table I-6 reports the results of three different specifications of logistic regressions. In the first specification – our benchmark specification – we include only the $Treatment$ dummy and a constant. We again find the treatment effect to be highly significant. To explore whether the effectiveness of the prompts varies by channel, we add a dummy variable, $In\ App$, which equals 1 if the specific encounter occurred in the mobile app and 0 if it occurred in online banking, along with the same control variables used in earlier specifications.

Table I-6: Effect of Channel on eCTR

	(1) Simplified Model	(2) Full Model	(3) Full Model
Treatment	1.2138***	1.2189***	1.3269**
	(0.0748)	(0.0751)	(0.1465)
In App		2.1633***	2.3169***
		(0.1577)	(0.2386)
Treatment*In App			0.8841
			(0.1169)
Age		1.0139***	1.0139***
		(0.0022)	(0.0022)
Male		1.2695***	1.2694***
		(0.0817)	(0.0817)
Monthly Income		0.8897***	0.8897***
		(0.0190)	(0.0190)
Account Balance		1.0003	1.0003
		(0.0003)	(0.0003)
Fund Owner		0.8825	0.8823
		(0.0770)	(0.0770)
Constant	0.0014***	0.0006***	0.0006***
	(0.0001)	(0.0001)	(0.0001)
Observations	676,107	676,107	676,107
Pseudo R^2	0.0006	0.0123	0.0124

Notes: This table reports odds ratios from three logistic regressions with $eCTR$ as the dependent variable (1 = client clicked the link in this encounter with the prompt). Unlike previous regressions based on individual clients, we now examine each encounter between clients and the prompt. Specification (1) includes only the $Treatment$ dummy (1 = client received the peer nudge). Specification (2) adds $age, gender, monthly\ income, account\ balance, fund\ ownership$, and $In\ App$ (1 = encounter in the bank's app, 0 = online banking). Specification (3) introduces an interaction term between $Treatment$ and $In\ App$. Standard errors are clustered at the client-level and are shown in parentheses. Significance levels for p-values < 0.10, 0.05, 0.01 are denoted by *, **, ***, respectively. The reported Pseudo R^2 is the McFadden R^2.

Our results show that clients are significantly more likely to click on the prompt when using the mobile app than when using online banking, aligning with broader findings that mobile users often display higher engagement levels (see also Jiang et al., 2018). One potential explanation for the higher click-through rate in the mobile app is that users checking their phone may be more open to spontaneous actions, whereas online banking usage on a computer may be more goal-directed or scheduled (see also Mograbi, 2022; Wang et al., 2023).

In the next step, we examine the effectiveness of the peer information nudge in different settings. Specifically, we include an interaction term, $Treatment * In\ App$, which equals 1 for the encounters that include the peer information and occur in the app and 0 otherwise. The main effect of $In\ App$ remains significant, indicating that clients are more likely to engage with prompts presented inside the mobile app. By contrast, the interaction term is not significant, indicating that while the mobile app drives higher baseline engagement, the peer information nudge is equally effective across both channels.

4. Discussion and Conclusion

This paper examines whether a digitally delivered and demographically tailored peer information nudge can move retail bank clients along the investment decision funnel. In our experiment, the peer information nudge successfully increased initial engagement with investing. User-level click-through ($uCTR$) increased from 1.82% in the Control group to 2.23% in the Treatment group. This is both statistically significant and economically meaningful, as it constitutes an increase in engagement of around 23%. By contrast, we observe no detectable effect on investment take-up (ITR: 0.52% Control vs. 0.49% Treatment). This suggests that descriptive norms alone are insufficient to overcome the cognitive and transactional frictions inherent in setting up an investment plan. Dur et al. (2021) similarly find that descriptive peer information often can garner attention without ultimately leading to effective changes in behavior.

Why might peer information increase initial engagement but stop short of leading to actual behavior change? Our design is motivated by two mechanisms through which peer information might influence investment behavior. First, a social learning mechanism: observing peer behavior may update investors' beliefs about product relevance or expected participation (Bursztyn et al., 2014). Second, a normative mechanism: individuals often align behavior with perceived descriptive norms, particularly when the reference group feels proximate or when deviation from the behavior feels salient (Festinger, 1954; Bicchieri & Dimant, 2022). Our implementation attempted to strengthen both mechanisms by (i) segmenting messages by age and gender to reduce social distance, and (ii) presenting a reachable threshold drawn from the lower quartile of existing investors, thereby avoiding the discouragement that can arise when peers appear far ahead (Beshears et al., 2015). The observed increase in $uCTR$ suggests these choices succeeded in increasing low-friction information-seeking. Yet the absence of downstream behavior change implies that peer information nudges, even when demographically matched and aspirationally modest, may not by themselves overcome the behavioral frictions, cognitive load, risk assessment, paperwork, or liquidity concerns that are involved in starting an investment plan – barriers also highlighted in field evidence on savings and retirement decisions (see also Bauer et al., 2022; Dur et al., 2021).

Our heterogeneity analyses provide further insights into the effectiveness of peer information nudges and prompt engagement in general. Older and male clients are significantly more likely to engage with the prompts. However, the treatment effects of the peer information nudge do not significantly differ by age or gender. This pattern suggests that banks might not need to microtarget peer information content itself to specific demographics; instead, targeting should focus on who is most likely to engage with any campaign message, while other interventions should focus on harder-to-influence groups. At the encounter level, prompts delivered through the mobile app produce substantially higher $eCTR$ than those in online banking, consistent with broader evidence of greater engagement on mobile channels (Jiang et al., 2018; Mograbi, 2022; Wang et al., 2023). However, we do not find that the marginal treatment effect of adding peer information differs by channel. Practically speaking, this means that if banks wish to maximize their campaign impact, ensuring high mobile penetration may yield larger absolute numbers of engaged clients, while the relative advantage of including peer information is similar across platforms.

Our study is, of course, not without its limitations. First, our ITR metric necessarily aggregates any investment plan set up and investment during the treatment period, regardless of whether it is causally linked to the intervention or not. It also does not capture investments made after the treatment period that were nonetheless triggered by the treatment. Second, our engagement measures lack granularity. We observe only binary click and take-up indicators, not intermediate steps or click quality. Accidental or low-attention clicks may inflate $uCTR$ and we cannot pinpoint intermediate drop-off points in the decision funnel. Third, external validity is bounded by geography and institutional context: results from one German regional bank may not extend to institutions with different client mixes, product menus, or baseline investing norms. Finally, despite demographic tailoring, clients still receive anonymized aggregate peer statistics. Prior work shows that the effectiveness of peer effects increases with richer, more relational context (Hong et al., 2004; Lieber & Skimmyhorn, 2018), raising the possibility that stronger identity cues, such as focusing on people living in the same zip code or in a similar phase of life, could increase the treatment effect.

Despite these limitations, our findings suggest that peer information can raise initial engagement with investing and lead to information-seeking among underinvested clients but should be accompanied by other measures, which help reduce behavioral friction. Such measures could, for example, be a streamlined digital process to open an investment account, proactive outreach from financial advisors, further information on the benefits of investing, matched contributions, or other financial benefits. Designing multistage interventions that focus on specific behavioral barriers might be necessary to translate initial engagement into actual investment behavior change.

Future research can build on our findings in several directions. First, experiments that include both peer information and friction-reducing or incentive treatments would test for complementary effects and could identify minimal effective bundles. Second, systematically varying the level (mean vs. quartile) and scope (investors only vs. full client base) of the referenced norm can illuminate backfire thresholds (e.g., Beshears et al., 2015) and the conditions under which modest, attainable norms (e.g., Carattini & Blasch, 2024) outperform more ambitious benchmarks. Third, longitudinal follow-ups would detect whether early engagement seeds later investing, even if short-run effects are muted. Finally, integrating detailed clickstream and process data would allow true funnel analytics, enabling researchers to diagnose precisely where clients drop out and which intervention elements remove which frictions.

Taken together, our results contribute to the emerging view that peer information nudges – and nudges more broadly – are particularly effective at changing proximate behavior such as initial engagement but often struggle to deliver fundamental changes in downstream economic results. To bridge the gap between initial engagement and action, these interventions must be complemented by mechanisms to lower the psychological, informational, and transactional costs of actual investing. Understanding how to design such complementary measures remains an important question for both researchers and practitioners.

5. References

Allcott, H. (2011). Social norms and energy conservation. Journal of Public Economics, 95(9-10), 1082–1095. https://doi.org/10.1016/j.jpubeco.2011.03.003

Athey, S., & Wager, S. (2019, February 20). Estimating treatment effects with causal forests: An application. https://arxiv.org/pdf/1902.07409v1

Bandura, A. (1977). Social learning theory. Prentice-Hall.

Bauer, R., Eberhardt, I., & Smeets, P. (2022). A fistful of dollars: Financial incentives, peer information, and retirement savings. The Review of Financial Studies, 35(6), 2981–3020. https://doi.org/10.1093/rfs/hhab088

Beshears, J., Choi, J. J., Laibson, D., Madrian, B. C., & Milkman, K. L. (2015). The effect of providing peer information on retirement savings decisions. The Journal of Finance, 70(3), 1161-1201. https://doi.org/10.1111/jofi.12258

Beshears, J., & Kosowsky, H. (2020). Nudging: Progress to date and future directions. Organizational Behavior and Human Decision Processes, 161, 3–19. https://doi.org/10.1016/j.obhdp.2020.09.001

Bicchieri, C., & Dimant, E. (2022). Nudging with care: The risks and benefits of social information. Public Choice, 191(3-4), 443–464. https://doi.org/10.1007/s11127-019-00684-6

Burke, J., Jamison, J., Karlan, D., Mihaly, K., & Zinman, J. (2023). Credit building or credit crumbling? A credit builder loan’s effects on consumer behavior and market efficiency in the United States. The Review of Financial Studies, 36(4), 1585–1620. https://doi.org/10.1093/rfs/hhac060

Bursztyn, L., Ederer, F., Ferman, B., & Yuchtman, N. (2014). Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions. Econometrica, 82(4), 1273–1301. https://doi.org/10.3982/ECTA11991

Carattini, S., & Blasch, J. (2024). Nudging when the descriptive norm is low: Evidence from a carbon offsetting field experiment. Journal of Behavioral and Experimental Economics, 110, 102194. https://doi.org/10.1016/j.socec.2024.102194

Chetty, R., Friedman, J. N., Leth-Petersen, S., Nielsen, T. H., & Olsen, T. (2014). Active vs. passive decisions and crowd-out in retirement savings accounts: Evidence from Denmark. The Quarterly Journal of Economics, 129(3), 1141–1219. https://doi.org/10.1093/qje/qju013

Choi, J., Laibson, D., Madrian, B., & Metrick, A. (2002). Defined contribution pensions: Plan rules, participant choices, and the path of least resistance. In J. M. Poterba (Ed.), Tax Policy and the Economy, Volume 16 (pp. 67–114). National Bureau of Economic Research, Inc. https://doi.org/10.1086/654750

Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. Journal of Personality and Social Psychology, 58(6), 1015–1026. https://doi.org/10.1037/0022-3514.58.6.1015

Cyr, D., Head, M., Larios, H., & Pan, B. (2009). Exploring human images in website design: A multi-method approach. MIS Quarterly, 33(3), 539–566. https://doi.org/10.2307/20650308

Døskeland, T., Martuza, J. B., Pedersen, L. J. T., Santos, F., Sjåstad, H., & Thorbjørnsen, H. (2025). The role of social norms in retirement saving: Evidence from two natural field experiments. Journal of Business Research, 190, 115201. https://doi.org/10.1016/j.jbusres.2025.115201

Duflo, E., Gale, W., Liebman, J., Orszag, P., & Saez, E. (2006). Saving incentives for low- and middle-income families: Evidence from a field experiment with H&R Block. The Quarterly Journal of Economics, 121(4), 1311–1346. https://doi.org/10.1093/qje/121.4.1311

Duflo, E., & Saez, E. (2002). Participation and investment decisions in a retirement plan: The influence of colleagues' choices. Journal of Public Economics, 85(1), 121–148. https://doi.org/10.1016/S0047-2727(01)00098-6

Duflo, E., & Saez, E. (2003). The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. The Quarterly Journal of Economics, 118(3), 815–842. https://doi.org/10.1162/00335530360698432

Dur, R., Fleming, D., van Garderen, M., & van Lent, M. (2021). A social norm nudge to save more: A field experiment at a retail bank. Journal of Public Economics, 200, 104443. https://doi.org/10.1016/j.jpubeco.2021.104443

Ellison, G., & Fudenberg, D. (1993). Rules of thumb for social learning. Journal of Political Economy, 101(4), 612–643. https://doi.org/10.1086/261890

Festinger, L. (1954). A theory of social comparison processes. Human Relations, 7(2), 117–140. https://doi.org/10.1177/001872675400700202

Goda, G. S., Manchester, C. F., & Sojourner, A. J. (2014). What will my account really be worth? Experimental evidence on how retirement income projections affect saving. Journal of Public Economics, 119, 80–92. https://doi.org/10.1016/j.jpubeco.2014.08.005

Guttman-Kenney, B., Adams, P., Hunt, S., Laibson, D., Stewart, N., & Leary, J. (2025). The semblance of success in nudging consumers to pay down credit card debt. American Economic Journal: Economic Policy, 17(4), 72–105. https://doi.org/10.1257/pol.20230568

Hallsworth, M., List, J. A., Metcalfe, R. D., & Vlaev, I. (2017). The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. Journal of Public Economics, 148, 14–31. https://doi.org/10.1016/j.jpubeco.2017.02.003

Hassanein, K., & Head, M. (2007). Manipulating perceived social presence through the web interface and its impact on attitude towards online shopping. International Journal of Human-Computer Studies, 65(8), 689–708. https://doi.org/10.1016/j.ijhcs.2006.11.018

Hong, H., Kubik, J. D., & Stein, J. C. (2004). Social interaction and stock-market participation. The Journal of Finance, 59(1), 137–163. https://doi.org/10.1111/j.1540-6261.2004.00629.x

Hu, Z. (2022). Social interactions and households' flood insurance decisions. Journal of Financial Economics, 144(2), 414–432. https://doi.org/10.1016/j.jfineco.2022.02.004

Jiang, T., Yang, J., Yu, C., & Sang, Y. (2018). A clickstream data analysis of the differences between visiting behaviors of desktop and mobile users. Data and Information Management, 2(3), 130–140. https://doi.org/10.2478/dim-2018-0012

Kang, J.-H., & Chung, D. Y. (2017). Homophily in an anonymous online community: Sociodemographic versus personality traits. Cyberpsychology, Behavior, and Social Networking, 20(6), 376–381. https://doi.org/10.1089/cyber.2016.0227

Kast, F., Meier, S., & Pomeranz, D. (2018). Saving more in groups: Field experimental evidence from Chile. Journal of Development Economics, 133, 275–294. https://doi.org/10.1016/j.jdeveco.2018.01.006

Lieber, E. M., & Skimmyhorn, W. (2018). Peer effects in financial decision-making. Journal of Public Economics, 163, 37–59. https://doi.org/10.1016/j.jpubeco.2018.05.001

Madrian, B. C., & Shea, D. F. (2001). The power of suggestion: Inertia in 401(k) participation and savings behavior. The Quarterly Journal of Economics, 116(4), 1149–1187. https://doi.org/10.1162/003355301753265543

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. Annual Review of Sociology, 27, 415–444. https://doi.org/10.1146/annurev.soc.27.1.415

Medina, P. C., & Pagel, M. (2025). Does saving cause borrowing? Implications for the coholding puzzle. The Journal of Finance, 80(5), 2689–2738. https://doi.org/10.1111/jofi.13466

Meske, C., & Potthoff, T. (2017). The Dinu-model - a process model for the design of nudges. In Proceedings of the 25th European Conference on Information Systems (ECIS), Guimarães, Portugal, June 5-10, 2017 (pp. 2587–2597). https://api.semanticscholar.org/CorpusID:37020288

Mirsch, T., Lehrer, C., & Jung, R. (2017). Digital nudging: Altering user behavior in digital environments. In Proceedings of the 13th International Conference on Wirtschaftsinformatik (WI) 2017 (pp. 634–648).

Mograbi, E. (2022). Decision-makers are more impulsive on smartphones than on computers. Journal of Behavioral and Experimental Economics, 100, 101916. https://doi.org/10.1016/j.socec.2022.101916

Ouimet, P., & Tate, G. (2020). Learning from coworkers: Peer effects on individual investment decisions. The Journal of Finance, 75(1), 133–172. https://doi.org/10.1111/jofi.12830

Raineau, Y., Giraud-Héraud, É., & Lecocq, S. (2025). Social comparison nudges: What actually happens when we are told what others do? Ecological Economics, 228, 108436. https://doi.org/10.1016/j.ecolecon.2024.108436

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. Psychological Science, 18(5), 429–434. https://doi.org/10.1111/j.1467-9280.2007.01917.x

Shang, J., & Croson, R. (2009). A field experiment in charitable contribution: The impact of social information on the voluntary provision of public goods. The Economic Journal, 119(540), 1422–1439. https://doi.org/10.1111/j.1468-0297.2009.02267.x

Sorensen, A. T. (2006). Social learning and health plan choice. The RAND Journal of Economics, 37(4), 929–945. https://doi.org/10.1111/j.1756-2171.2006.tb00064.x

Valta, M., & Maier, C. (2025). Digital nudging: A systematic literature review, taxonomy, and future research directions. ACM SIGMIS Database: The DATABASE for Advances in Information Systems, 56(1), 101–125. https://doi.org/10.1145/3715966.3715973

Wang, T., Chen, W.-F., Wang, X., & Fan, X. (2023). Smartphone use increases the likelihood of making short-sighted financial decisions. Journal of Pacific Rim Psychology, 17, 18344909221147782. https://doi.org/10.1177/18344909221147782

Weinmann, M., Schneider, C., & vom Brocke, J. (2016). Digital nudging. Business & Information Systems Engineering, 58(6), 433–436. https://doi.org/10.1007/s12599-016-0453-1

6. Appendix

Appendix I-A: Examples of the Interventions

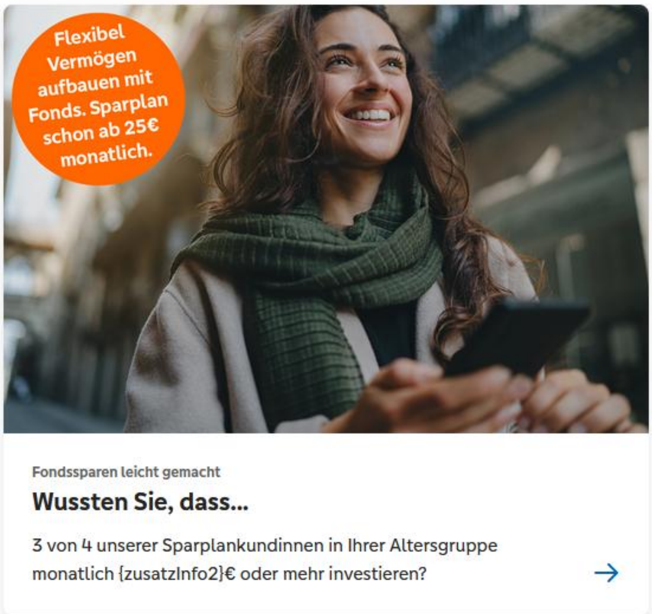

Figure I-1: Peer Information Nudge in Online Banking Channel for Women Aged 26-30

Translation:

Text in orange bubble: *Flexible wealth building with a fund savings plan starting from just €25 per month.*

Text at the bottom: *Fund saving made easy*

Did you know that…

3 out of 4 of our female investment plan clients in your age group invest [Placeholder]€ or more each month?

Figure I-2: Neutral Prompt in Online Banking Channel for Women Aged 26-30
Translation:
Text in orange bubble: *Flexible wealth building with funds. Savings plan from just €25 per month.*
Text at the bottom: *Fund saving made easy*
Build your wealth step by step
Saving with funds is possible from just €25 per month – with broad diversification and high flexibility.

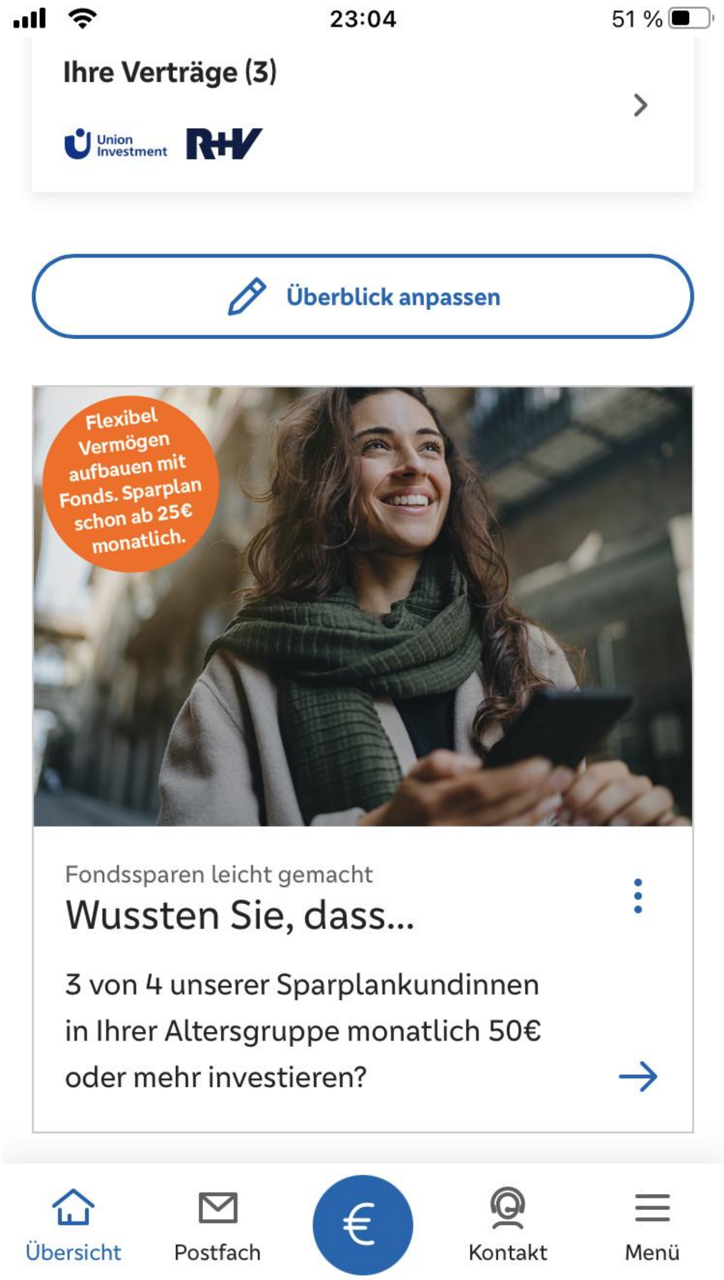

Figure I-3: Peer Information Nudge in the Mobile App Channel for Women Aged 26-30

Translation:
Text in orange bubble: *Flexible wealth building with a fund savings plan starting from just €25 per month.*
Text at the bottom: *Fund saving made easy*
Did you know that…
3 out of 4 of our female investment plan clients in your age group invest €50 or more each month?

Appendix I-B: Reference Values per Peer Group

Age	Male	Female
18-25	50	40
26-30	50	50
31-35	50	50
36-45	50	50
46-55	50	50
56-65	75	50
66+	75	50

Notes: This table reports the 25th-percentile (lower-quartile) of monthly contribution amounts made by existing investment-plan clients within each age–gender peer group. These figures were calculated in the month preceding the field experiment and were quoted verbatim in the treatment prompt, phrased as: “Did you know that 3 out of 4 of our female/male investment-plan clients in your age group invest at least [Placeholder]€ every month?” Hence, [Placeholder]€ equals the amount shown in the table for the participant’s own peer group.

Appendix I-C: Outcome by Treatment Group for the Full Sample

	Treatment N=39,333	Control N=39,080	Difference
uCTR	1.96%	1.60%	0.36%***
	(13.86%)	(12.56%)	
ITR	0.39%	0.44%	-0.06%
	(6.20%)	(6.64%)	

Notes: This table reports the mean of each outcome for the Treatment and Control groups based on all clients who were scheduled to receive the prompt by the bank, with standard deviations in parentheses. $uCTR$ is a dummy variable indicating whether the client clicked the link for more information. ITR is a dummy variable for whether the client signed up for an investment plan or directly invested in assets also available through the plan during the treatment period. The difference column reports the mean difference in percentage points across groups, with significance assessed via a chi-square test of independence. Significance levels for p-values < 0.10, 0.05, and 0.01 are denoted by *, **, and ***, respectively.

Appendix I-D: Investment Plan Sign-up Rate by Treatment Group

	Treatment N=26,649	Control N=26,268	Difference
Investment plan sign-up rate	0.41%	0.46%	-0.06%
	(6.38%)	(6.80%)	

Notes: This table reports the mean of the Investment plan sign-up rate for the Treatment and Control groups, with standard deviations in parentheses. $Investment\ plan\ sign-up\ rate$ is a dummy variable for whether the client signed up for an investment plan during the treatment period. The difference column reports the mean difference in percentage points across groups, with significance assessed via a chi-square test of independence. Significance levels for p-values < 0.10, 0.05, and 0.01 are denoted by *, **, and ***, respectively.

Appendix I-E: Causal Forest Calibration Test

	Estimate	Standard Error	T-Statistic	P-Value
Mean Forest Prediction	1.0292	0.3088	3.3325	0.0004
Differential Forest Prediction	-3.1980	1.8316	-1.7461	0.9596

Notes: This table reports the results of the omnibus test for the presence of heterogeneous treatment effects on $uCTR$. The test is derived from a generalized random forest trained on the full set of pretreatment covariates. The Mean forest prediction assesses whether the estimated average treatment effect is statistically significant and well-calibrated. The Differential forest prediction tests the null hypothesis that the treatment effect is homogeneous across the population; a statistically insignificant coefficient indicates that the model fails to detect significant heterogeneity beyond the average effect.

II. From Branches to Browsers: A Comparative Analysis of Direct and Traditional Bank Clients in Germany

Co-author:

Andreas Walter

Own share:

75%

This article is currently under review at:

Review of Financial Economics (2nd round of review)

Previous versions of this paper have been presented at the following conferences and workshops:

- 18th International Behavioural Finance Conference, 2025, London, United Kingdom
- Research Symposium on Finance and Economics (RSFE), 2025, (virtual)

From Branches to Browsers: A Comparative Analysis of Direct and Traditional Bank Clients in Germany[a]

Martin G. Becker[b]

Andreas Walter[c]

Abstract: This study examines retail banking clients with a house bank relationship either to a traditional bank (with a branch network) or a direct bank (without a branch network). Analyzing a representative sample of over 14,000 household-wave observations from 2011 to 2023 in Germany, we find that direct bank clients tend to be significantly younger, possess higher financial literacy, are better educated, and exhibit higher risk tolerance. Furthermore, direct bank clients demonstrate higher levels of stock market participation and invest a larger proportion of their financial assets in risky financial assets. We conclude that direct bank clients are generally well-equipped to make Do-It-Yourself (DIY) investment decisions.

JEL Codes: D14, G11, G21, G53

Keywords: retail banking, direct bank, DIY investments, financial literacy

[a] The authors gratefully acknowledge funding from the “Stiftung für die Wissenschaft”.

[b] martin.becker@wirtschaft.uni-giessen.de, Justus Liebig University Gießen, Department of Financial Services, Licher Straße 74, 35394 Gießen, Germany

[c] andreas.walter@wirtschaft.uni-giessen.de, Justus Liebig University Gießen, Department of Financial Services, Licher Straße 74, 35394 Gießen, Germany

1. Introduction

The market share of direct banks—including neobanks (e.g., Revolut and N26)—has grown significantly in recent years, coinciding with a decline in the branch networks of traditional banks in Germany and globally.[1] These digital-only banks operate without physical branches and typically do not offer personalized financial advisory services. This development raises two key questions: First, how do the clients of direct banks compare with those of traditional banks? Second, given the absence of readily available financial advice, are direct bank clients well-equipped to make sound investment decisions on their own? Despite the increasing importance of digital-only banking, there is little research on the characteristics of these clients (e.g., Filotto et al., 2021) and their financial behavior. It remains unclear whether they exhibit traits that could make them financially vulnerable, such as low financial literacy or limited experience with financial markets. Additionally, little is known about differences in investment behavior between direct and traditional bank clients. Does the lack of personal advice contribute to unhealthy investment behaviors such as underinvestment in risky assets, or are direct bank clients well-positioned to invest successfully on their own? This study aims to fill these gaps by examining whether direct bank clients, as Do-It-Yourself (DIY) investors, face vulnerabilities that may warrant regulatory attention.

To address the research questions raised above, we analyze a representative sample of more than 14,000 household-wave observations in Germany using pooled data from all five waves of the survey-based Panel on Household Finances (PHF), conducted between 2011 and 2023. This dataset includes comprehensive information about investors' characteristics as well as their

[1] See for example:
Statista (2024b) *Global market volume of neobanks*
https://www.statista.com/statistics/1228241/neobanks-global-market-size/
Deutsche Bundesbank (2024) *Payment behaviour in Germany 2023*
https://www.bundesbank.de/en/publications/reports/studies/payment-behaviour-in-germany-in-2023-934896
Deutsche Bundesbank (2023) *Changes in bank office statistics in 2022*
https://www.bundesbank.de/en/press/press-releases/changes-in-bank-office-statistics-in-2022-915502

financial decision-making. With respect to clients' characteristics, we observe significant differences between direct bank clients and those of traditional banks. Specifically, direct bank clients tend to be better educated and more financially literate than traditional bank clients. Additionally, direct bank clients are, on average, 10 years younger, have higher income but lower net wealth, and exhibit higher risk tolerance than clients of traditional banks.

Regarding investment behavior, our dataset includes information on stock market participation and the proportion of total financial assets invested in risky assets such as stocks, bonds, and mutual funds. We find strong evidence that direct bank clients are significantly more likely to invest in the stock market, even after controlling for clients' characteristics. Consequently, they benefit more frequently from higher expected returns associated with investments in the stock market. Exploring differences across subgroups, we observe that direct bank clients consistently exhibit higher stock market participation across various strata of our sample. Thus, this pattern appears to be broadly distributed across our sample of retail clients in Germany.

With regard to our control variables on client characteristics, we find corroborating evidence for factors positively associated with stock market participation. Consistent with the findings of Balloch et al. (2015), Lusardi and Mitchell (2014), and van Rooij et al. (2011), we observe that individuals with higher financial literacy are more likely to invest in the stock market. Moreover, in line with Black et al. (2018) and Cole et al. (2014), our results show that higher levels of formal education are also associated with higher stock market participation. Consistent with the findings of Kaustia et al. (2023), Mankiw and Zeldes (1991), and Shum and Faig (2006), we also observe that income and net wealth are positively associated with holding stocks or mutual funds. Furthermore, we confirm the results of Almenberg and Dreber (2015), Barber and Odean (2001), Ke (2021), and Sundén and Surette (1998), showing that men are more likely to participate in the stock market. Additionally, we provide supporting evidence for

Haliassos and Bertaut (1995) and Shum and Faig (2006) that individuals with higher risk tolerance are more likely to invest in stock markets.

In addition to stock market participation, we examine the share of total financial assets allocated to risky assets. Our results show that direct bank clients, on average, hold a higher proportion of their portfolios in risky assets. This finding remains significant even after controlling for confounding factors such as income, financial literacy, age, and gender. Consequently, based on these two measures of investment behavior, direct bank clients appear to be better positioned compared to traditional bank clients.

Our results contribute to several streams of research. First, we contribute to the literature examining the positive (Chalmers & Reuter, 2020) and negative effects (Mullainathan et al., 2012) of financial advice on retail clients. In particular, we expand this literature as we compare the financial behavior of traditional bank clients who regularly receive financial advice to DIY investors at direct banks, who typically operate without personal advisory services. Our findings indicate that DIY investors at direct banks are more often invested in the stock market and invest a larger proportion of their financial assets in risky assets.

Second, we contribute to the body of research examining the factors associated with stock market participation and investments in risky assets. In line with prior work, we document the importance of factors such as income and wealth (e.g., Kaustia et al., 2023; Mankiw & Zeldes, 1991), formal education (e.g., Black et al., 2018; Cole et al., 2014), risk aversion (e.g., Haliassos & Bertaut, 1995; Shum & Faig, 2006), and financial literacy (e.g., Lusardi & Mitchell, 2014; van Rooij et al., 2011).

Lastly, we contribute by studying the impact of behavioral factors (e.g., Luo & Subrahmanyam, 2019; Patterson & Daigler, 2014), such as life satisfaction or happiness (e.g., Cui & Cho, 2019; Delis & Mylonidis, 2015), trust (e.g., Balloch et al., 2015; Georgarakos & Pasini, 2011; Guiso et al., 2008), as well as impatience (e.g., Benartzi & Thaler, 1995; Frederick et al., 2002). While we find evidence linking higher life satisfaction with stock market participation and a higher share of investment in risky financial assets, we do not find any evidence of the impact of trust and impatience on either investment variable.

The remainder of this paper is structured as follows. Section 2 offers background information on the retail banking market in Germany and differentiates between direct banks and traditional banks. Section 3 introduces the dataset, defines the key variables of interest, and provides summary statistics. In Section 4, we present our empirical results, followed by robustness checks in Section 5. Finally, Section 6 concludes the paper.

2. Retail Banking Market in Germany

Our study focuses on the German retail banking market, the largest in Europe, with a total population of approximately 85 million (Federal Statistical Office, 2024). The German retail banking market is characterized by two key features. First, German clients typically maintain a primary banking relationship with one bank, known as the "Hausbank" (house bank)[2]. Clients refer to their house bank for core financial services such as checking accounts, loans, and investment options. Traditional banks maintain a network of branches where clients can access various banking services, including personalized financial advice. However, due to the costs associated with maintaining branch networks and providing personalized advice, traditional banks tend to charge higher fees and commissions (Sarel & Marmorstein, 2003).

Second, the German retail banking market is characterized by a high level of competition. As of 2023, 1,403 independent banks operate in the market (Deutsche Bundesbank, 2025b), with the majority being credit banks (e.g., Deutsche Bank), savings banks (e.g., Frankfurter Sparkasse), and cooperative banks (e.g., Frankfurter Volksbank). Since the mid-1990s, a new

[2] This strong relationship is also reflected in our data. In the original sample, only 4% of households named more than one bank as their house bank, while only 11% did not name any bank at all. This close relationship is further corroborated by the relatively low willingness to switch primary banks. (e.g., Statista (2024a) *Germany: willingness to change primary bank 2024* https://www.statista.com/statistics/1395372/willingness-to-change-primary-bank-in-germany/)

type of bank has gained prominence in Germany: direct banks (e.g., comdirect, DKB, and ING). These banks do not maintain branch networks but provide financial services through the internet, initially via online banking on desktop PCs and more recently through mobile banking on smartphones. Direct banks typically offer lower fees but do not provide personalized financial advice. Their digital services are generally more convenient than those of traditional banks, particularly for activities such as trading of stocks and ETFs (Citterio et al., 2025). As a result, direct bank clients make their financial decisions without an advisor and can thus be categorized as DIY investors.

In this study, we will explore the differences in client characteristics between direct and traditional banks. Additionally, we will examine whether direct banking clients, as DIY investors, exhibit different investment behaviors compared to traditional banking clients.

3. Data

3.1 Sample Selection

We employ the German Panel on Household Finances (PHF) (Deutsche Bundesbank, 2025a), a longitudinal survey conducted by the Research Centre of the Deutsche Bundesbank. The PHF captures detailed information on household finances, including assets, liabilities, income, financial literacy, and various demographic characteristics. It is designed as a full panel survey, with repeated interviews of the same households over time.[3] It should be noted that the survey is on the household level, but the questions in the survey regarding the whole household are not answered by all household members but by the "financially knowledgeable person", FKP, of that household. The FKP is defined as the household member who is best informed about the household's finances, as determined by the household members. Given that the financial data are only available at the household level, we match the households' finances with the

[3] Although the PHF is designed as a full panel survey, it employs a dynamic panel concept that includes drawing refresher samples (newly selected households) to counteract attrition and to ensure that the sample remains representative of the overall population.

demographic information of the FKP within each household. Consequently, the majority of analyses, such as those examining the correlation between household financial behavior and demographics or financial literacy, are conducted using data for the FKP of the household.

Wealthy households are deliberately oversampled in the PHF survey to ensure an accurate representation of wealth distribution across the population (Schmidt & Eisele, 2013). To obtain results that are representative of the German population, we apply survey weights in all analyses and thereby correct for the deliberate oversampling. We also adjust our standard errors for imputation error and survey design by using the replication weights included in the sample. For further information on the survey and replication weights and the data collection process, see von Kalckreuth et al. (2012).

Our analysis is based on the pooled sample of all five waves, with the first wave conducted in 2011 and the most recent wave conducted in 2023. For the purpose of this study, we limit our sample to households that report maintaining a single house bank relationship and for which the dataset is complete for all variables of interest[4], thus excluding those observations of households that report not maintaining a house bank relationship and those reporting more than one such relationship. This results in a final sample size of 14,624 household-wave observations.

3.2 Variable Measurement

Our first key variable refers to the house bank relationship. Specifically, we define a dummy variable, $Direct\ Bank$, which takes the value of one if the household maintains an exclusive banking relationship with a direct bank and zero otherwise. Our second set of variables of interest refers to the investment behavior. First, we define the dummy variable $Stock\ Market\ Participation$, which takes the value of one if the household is invested in the

[4] Due to anonymization reasons, the PHF dataset does not include specific variables for some households. For example, for households in wealthy municipalities in the former GDR, the population size of the respective municipality is not given since this might allow to deanonymize households.

stock market, either directly or indirectly, and zero otherwise.[5] Second, we define the variable $Risky\ Asset\ Share$, which is calculated as the proportion of a household's investments in risky financial assets relative to the total financial assets held by the household.

We also introduce several control variables across three dimensions: (i) demographics, (ii) financial literacy, and (iii) attitudes. Please note that most of these control variables refer to the FKP rather than the household. A comprehensive description of these variables is provided in Appendix II-A.

3.3 Summary Statistics

Table II-1 presents descriptive statistics for the variables introduced in Section 3.2. Panel A shows that 5.7% of banking clients maintain an exclusive house bank relationship with a direct bank.[6]

Panel B reports descriptives for the investment behavior. First, a mean of 25.6% for the dummy variable $Stock\ Market\ Participation$ indicates that over a quarter of households in our dataset invest in the stock market.[7] Second, we find that bank clients invest only a relatively low fraction of their financial assets in risky financial assets, such as stocks, bonds, or mutual funds. Specifically, the mean proportion is 8.8%, which is slightly lower than the 10.6% reported by Zhan (2015) in a study on the allocation of risky assets among German households. The distribution is highly skewed, with a median of 0.0% and a 75th percentile of only 2.1%,

[5] For the concrete definition, please see Appendix II-A.

[6] This figure is substantially lower than the 18% of respondents who reported having a current account at a direct bank in 2023, according to a Forsa survey commissioned by the Deutsche Bundesbank (2024). The discrepancy is likely explained by two factors: first, the Forsa survey allowed multiple responses, and second, it was conducted at the individual rather than the household level. Both aspects could have led to higher direct bank usage being reported, as direct banks are more commonly used as secondary accounts and tend to be favored by younger individuals, who are less likely to serve as the FKP in their households. Additionally, the value is further reduced by pooling observations from all waves. In the fifth wave of the panel, 8.4% of households report having a direct bank as their house bank.

[7] Notably, this proportion is higher than figures typically reported for stock market participation in Germany. For instance, the Deutsches Aktien Institut (DAI) (2024) reports that 17.6% of Germans over the age of 14 are invested in the stock market in 2023. Since our analysis focuses on stock market participation at the household level and only for households that maintain a house bank relationship rather than at the individual level for all Germans, we naturally observe higher participation rates.

indicating that the vast majority of households hold negligible or no positions in risky financial assets.

Panel C reveals that the average bank client is 52.9 years old.[8] The 75th percentile is 67.0 years, indicating that more than a quarter of our bank clients are retirees. Our sample is rather balanced between males (54.7%) and females (45.3%). Additionally, 13.4% of bank clients reside in small municipalities with fewer than 5,000 inhabitants. With respect to income and wealth, we find that the average bank client has a household income of €56,900 and household net wealth of €274,400. Finally, 23.3% of our bank clients hold a university degree.

In Panel D, we observe that the FKPs in our dataset exhibit relatively high levels of financial literacy. Specifically, 60.8% of bank clients correctly answered all three questions from the *Big Three* financial literacy test introduced by Lusardi and Mitchell (2008). The average score on this measure is 2.5.[9]

In Panel E, we present statistics on four survey items related to life satisfaction, risk tolerance, trust in others, and impatience. We find that FKPs are generally satisfied with their lives, with the median bank client rating their life satisfaction as an eight on a scale from zero to ten, with zero (ten) indicating the lowest (highest) possible life satisfaction. The other attitudes—risk tolerance, trust, and impatience—exhibit lower median values.

Additionally, Table II-2 provides a correlation matrix of the variables. The maximum absolute correlation between independent variables is 32%, indicating that multicollinearity is not a significant concern in our dataset. Nonetheless, we apply robust standard errors in all our multivariate analyses.

---

[8] Since the age variable is top coded at 90 for anonymity purposes, the true average age is likely to be slightly higher.

[9] In the fourth and fifth waves of the PHF dataset, not all participants received the same three questions. Instead, for a randomly selected half, the classic question on diversification was replaced by another question on diversification. If we restrict our sample to the first three waves and the halves of waves 4 and 5 that received the original “Big Three” questions, the mean value for correct answers remains at 2.5 and the share of bank clients who answered all three questions correctly increases slightly to 63.9%.

Table II-1: Descriptive Statistics

Variable	Mean	SD	Min	25%-Quantile	Median	75%-Quantile	Max	N
Panel A: House Bank								
Direct Bank	5.74%	23.3%	0.00	0.00	0.00	0.00	1.00	14,624
Panel B: Investment Behavior								
Stock Market Participation	25.58%	43.6%	0.00	0.00	0.00	1.00	1.00	14,624
Risky Asset Share	8.82%	20.3%	0.00%	0.00%	0.00%	2.08%	100.00%	14,624
Panel C: Demographics								
Age	52.93	17.60	17.00	39.00	53.00	67.00	90.00	14,624
Male	54.71%	49.78%	0.00	0.00	1.00	1.00	1.00	14,624
Small-Town Resident	13.43%	34.10%	0.00	0.00	0.00	0.00	1.00	14,624
Household Income (yearly)	56.92	69.40	-0.60	25.04	41.81	69.04	4,302.50	14,624
Household Net Wealth	274.41	719.70	-2,440.86	12.42	88.82	311.76	76,305.00	14,624
University Degree	23.25%	42.24%	0.00	0.00	0.00	0.00	1.00	14,624
Panel D: Financial Literacy								
Financial Literacy	2.46	0.78	0.00	2.00	3.00	3.00	3.00	14,624
Financial Literacy = 3	60.78%	48.8%	0.00	0.00	1.00	1.00	1.00	14,624
Panel E: Attitudes								
Satisfaction with Life	7.19	1.79	1.00	6.00	8.00	8.00	10.00	14,624
Risk Tolerance (general)	4.27	1.99	1.00	3.00	4.00	5.00	10.00	14,624
Trust	5.60	1.87	1.00	5.00	5.00	7.00	10.00	14,624
Impatience	4.80	2.35	1.00	3.00	5.00	7.00	10.00	14,624

Notes: This table reports the descriptive statistics for our sample. Definitions for each of the variables can be found in Appendix II-A. The Mean and quantiles are calculated using Rubin rules and the multiple imputations and survey weights included in the dataset.

Table II-2: Correlations

Variable	Age	Male	Small town resident	Household Income (yearly)	Household Net Wealth	University Degree	Financial Literacy	Satisfaction with Life	Risk Tolerance (general)	Trust	Impatience
Age	1.00***	-0.03	0.02	-0.02*	0.11***	-0.08***	-0.10***	0.01**	-0.11***	-0.05***	-0.02**
Male	-0.03*	1.00***	0.00	0.07***	0.04***	0.05***	0.10***	-0.04***	0.13***	-0.03**	-0.05***
Small-Town Resident	0.02	0.00	1.00***	-0.01*	0.03	-0.07***	-0.01**	-0.02**	-0.02**	-0.04***	-0.01**
Household Income (yearly)	-0.02	0.07***	-0.01*	1.00***	0.32***	0.18***	0.11***	0.15***	0.10***	0.06***	0.02**
Household Net Wealth	0.11***	0.04***	0.03	0.32***	1.00***	0.09***	0.06***	0.13***	0.07***	0.02**	0.03**
University Degree	-0.08***	0.05***	-0.07***	0.18***	0.09***	1.00***	0.16***	0.11***	0.08***	0.14***	-0.02**
Financial Literacy	-0.10***	0.10***	-0.01	0.11***	0.06***	0.16***	1.00***	0.09***	0.04***	0.06***	0.01**
Satisfaction with Life	0.01**	-0.04***	-0.02**	0.15***	0.13***	0.11***	0.09***	1.00***	0.11***	0.21***	-0.05***
Risk Tolerance (general)	-0.11***	0.13***	-0.02	0.10***	0.07***	0.08***	0.04***	0.11***	1.00***	0.19***	0.10***
Trust	-0.05***	-0.03**	-0.04***	0.06***	0.02**	0.14***	0.06***	0.21***	0.19***	1.00***	-0.06***
Impatience	-0.02**	-0.05***	-0.01	0.02**	0.03**	-0.02	0.01**	-0.05***	0.10***	-0.06***	1.00***

Notes: This table reports the pairwise correlations for all control variables in our sample. The correlations are calculated using Rubin rules and the multiple imputations and survey weights included in the dataset. Significance levels are denoted as follows: *$p < .1$; **$p < .05$; ***$p < .01$.

4. Empirical Results

4.1 Are Clients of Direct Banks Different?

To examine the characteristics of direct bank clients, we report univariate results in Table II-3 and provide the respective multivariate evidence from logistic regressions in Table II-5.

Table II-3: Differences in Client Characteristics

Variable	Mean Direct Bank	Mean Traditional Bank	Mean Difference
Panel A: Demographics			
Age	43.66	53.49	-9.83***
Male	51.32%	54.92%	-3.60%
Small-Town Resident	11.04%	13.58%	-2.53%
Household Income (yearly)	68.21	56.23	11.98**
Household Net Wealth	187.97	279.67	-91.70***
University Degree	44.18%	21.97%	22.20%***
Panel B: Financial Literacy			
Financial Literacy	2.66	2.44	0.21***
Financial Literacy = 3	72.52%	60.06%	12.45%***
Panel C: Attitudes			
Satisfaction with Life	7.46	7.17	0.29**
Risk Tolerance (general)	4.75	4.24	0.51***
Trust	6.03	5.57	0.45***
Impatience	4.64	4.81	-0.17

Notes: This table reports the mean for direct bank clients and traditional bank clients for all control variables. Mean difference reports the difference between these means. *,**,*** indicate whether the difference between the bank groups is statistically significant at the 10%, 5% and 1% level respectively. The p-values are calculated based on a two-sided t-test for continuous variables and a chi-square test for independence for binary variables. Means and p-values are calculated using Rubin rules and the multiple imputations and survey weights included in the dataset.

In Panel A of Table II-3, we report that direct bank clients differ from traditional bank clients in a number of demographic dimensions. Most notably, clients of direct banks are considerably younger (mean age of 43.7) compared to traditional bank clients (mean age of 53.5). The significant age difference is also evident in Table II-4, which presents the relative proportions of different generations within the customer base of direct and traditional banks, respectively. For example, 36.9% of direct bank clients belong to Generation Y (born in 1981-1996) as opposed to 19.4% for traditional bank clients. This relationship is reversed for the older generations. For example, 32.7% of traditional bank clients belong to the Baby Boomers (born in 1946-1964); the respective fraction for direct bank clients is 17.3%. Notably, almost every sixth client (18.1%) of traditional banks belongs to the Silent Generation (born before 1946). The respective fraction for direct banks is only 3.9%.

Table II-4: Differences between Generations

Generation	Mean Direct Bank	Mean Traditional Bank	Mean Difference
Silent generation (1928–1945)	3.89%	18.06%	-14.17%***
Baby boomer (1946–1964)	17.31%	32.70%	-15.39%***
Generation X (1965–1980)	39.90%	26.72%	13.18%***
Millennials (Generation Y) (1981–1996)	36.87%	19.35%	17.52%***
Generation Z (1997–2012)	2.03%	1.99%	0.04%
Sum	100.0%	100.0%	

Notes: This table reports the mean share that each generation represents among direct bank clients and among traditional bank clients. The Mean Difference column shows the difference between these two means for each generation. *,**,*** indicate whether the difference between the bank groups is statistically significant at the 10%, 5% and 1% level respectively. Means and p-values are calculated using Rubin rules and the multiple imputations and survey weights included in the dataset.

We also find differences with respect to income and wealth. Particularly, direct bank clients are less wealthy than traditional bank clients. The mean household net wealth of direct bank clients is €188,000 as opposed to €279,700 for traditional bank clients. In contrast, direct bank clients seem to earn a higher income (mean of €68,200) compared to traditional bank clients (mean of €56,200). Additionally, we find direct bank clients to more often hold a university degree (44.2%). By comparison, only 22.0% of traditional bank clients completed a university education. We do not document significant differences in terms of living in a small municipality or gender between the two groups.

Panel B of Table II-3 presents information on financial literacy. Direct bank clients exhibit significantly higher financial literacy, as measured by the *Big Three* financial literacy questions. In particular, 72.5% of direct bank clients correctly answered all three questions in the financial literacy survey. For traditional banks, this fraction is 60.1%. The difference in financial literacy levels also translates into a higher mean value for the financial literacy score. The mean value is 2.7 (2.4) for direct bank (traditional bank) clients.[10]

[10] If we once again consider only the first three waves plus the halves of the sample of waves 4 and 5 that received the original Big Three questions, the proportion of clients who answered all three correctly increases to 76.5% for direct bank clients and 63.2% for traditional bank clients. Correspondingly, the average number of correct answers also rises to 2.71 for direct bank clients and 2.49 for traditional bank clients. In both cases, the differences remain statistically significant.

Finally, Panel C of Table II-3 presents survey evidence on clients' attitudes, highlighting substantial differences between the two banking groups. Direct bank clients report greater trust in others (mean of 6.0 compared to 5.6) and are, on average, more willing to take risks than traditional bank clients (mean of 4.8 compared to 4.2). Direct bank clients also indicate a slightly higher satisfaction with life (mean of 7.5 compared to 7.2). No significant difference in impatience is observed between the two groups.

To analyze which characteristics are significantly linked with bank choice in a multivariate setting, we estimated four logistic regression models. In each model, the dependent variable is the binary variable $Direct\ Bank$. Models 1 to 3 include the variables introduced in Panels C to E of Table II-1 separately. Model 4 presents the estimation for the full model, incorporating all independent variables. Robust standard errors, which adjust for imputation error and survey design, and wave fixed effects, accounting for time-specific trends or shocks, are used in all regressions.

Table II-5: Logistic Regression on Direct Bank

	Dependent Variable			
	(1) Direct Bank	(2) Direct Bank	(3) Direct Bank	(4) Direct Bank
Age	-0.0321***			-0.0298***
	(0.0040)			(0.0040)
Male	-0.1798			-0.2632*
	(0.1397)			(0.1445)
Small-Town Resident	-0.0600			-0.0316
	(0.2496)			(0.2503)
Household Income (yearly)	0.0994**			0.0821*
	(0.0428)			(0.0436)
Household Net Wealth	-0.3137***			-0.3740***
	(0.1111)			(0.1222)
University Degree	0.9453***			0.8040***
	(0.1287)			(0.1313)
Financial Literacy		0.4703***		0.3256***
		(0.1048)		(0.1012)
Satisfaction with Life			0.0509	0.0378
			(0.0392)	(0.0437)
Risk Tolerance (general)			0.0996***	0.0881***
			(0.0277)	(0.0318)
Trust			0.0994***	0.0440
			(0.0377)	(0.0378)
Impatience			-0.0339	-0.0388
			(0.0236)	(0.0250)
Constant	-2.1515***	-4.6855***	-4.6815***	-3.7636***
	(0.2649)	(0.3493)	(0.3828)	(0.5334)
Observations	14,624	14,624	14,624	14,624
Adj. Pseudo R^2	0.0879	0.0304	0.0308	0.0982

Notes: This table reports our regression results for a logistic regression with $Direct\ Bank$ as the dependent variable. In column (1) $Direct\ Bank$ is regressed on demographic variables. For interpretability, we z-standardized $Household\ Income$ and $Household\ Net\ Wealth$ (i.e., subtracted their means and divided by their standard deviations) prior to including them in the model. In column (2) $Direct\ Bank$ is regressed on the variable $Financial\ Literacy$, which indicates the score of the FKP in a three-question battery designed to test their financial literacy. In column (3) $Direct\ Bank$ is regressed on a set of variables describing the FKP attitudes. In column (4) all variables from the previous regressions are included in one model. All models include wave fixed effects to account for systematic differences across survey waves. We report the adjusted McFadden Pseudo R^2 for each specification. Standard errors (in parentheses) are adjusted for imputation error, following Rubin's Rules, and for survey design, using replicate weights. The data are weighted using the sampling weights included in the sample. Significance levels are denoted as follows: $^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$.

As shown in Table II-5, the logit results confirm the findings from the group comparison documented in Table II-3. We find a significant negative correlation between age, net wealth, and the likelihood of being a direct bank client. We also confirm that higher income, financial literacy, and possessing a university degree are significantly positively correlated with the likelihood of being a direct bank client. Regarding attitudes, higher levels of risk tolerance are associated with a higher probability of being a direct bank client. In contrast, the relationships observed for life satisfaction and trust do not hold in the full model. The case of trust is particularly illustrative: its significant positive link to $Direct\ Bank$ persists even in a multivariate regression with other attitudes (Model 3), but this association disappears once demographic and financial literacy variables are included (Model 4). This suggests the relationship is likely channeled through these other client characteristics. In sum, most significant characteristics in the univariate group comparison also hold in multivariate estimations, with life satisfaction and trust being the rare exceptions.

4.2 Do Clients of Direct Banks Invest Differently?

We now examine whether direct bank clients exhibit different investment behavior. Table II-6 compares (i) the $Stock\ Market\ Participation$ and (ii) the proportion of risky assets relative to financial assets ($Risky\ Asset\ Share$) between clients of direct and traditional banks.

Notably, we find that direct bank clients exhibit a significantly higher stock market participation rate. Specifically, 40.0% of direct bank clients are invested in the stock market, compared to only 24.7% of traditional bank clients. Additionally, direct bank clients allocate a larger portion of their financial assets to risky financial instruments, such as stocks, bonds, and mutual funds. Concretely, the average direct bank client invests 12.7% of their financial assets in risky assets, whereas this proportion is 8.6% for traditional bank clients.

Table II-6: Differences in Investment Behavior

Variable	Mean Direct Bank	Mean Traditional Bank	Mean difference
Stock Market Participation	39.97%	24.70%	15.27%***
Risky Asset Share	12.68%	8.59%	4.10%***

Notes: This table reports the mean for direct bank clients and traditional bank clients for the investment behavior variables. Mean difference reports the difference between these means. *,**,*** indicate whether the difference between the bank groups is statistically significant at the 10%, 5% and 1% level respectively. The p-values are calculated based on a two-sided t-test for continuous variables and a chi-square test for independence for binary variables. Means and p-values are calculated using Rubin rules and the multiple imputations and survey weights included in the dataset.

Table II-7 examines the link between being a direct bank client and investment behavior within a multivariate framework. Concretely, we estimate the following linear regression model:

$$Investment\ Behavior_i = \beta_0 + \beta_1 Direct\ Bank_i + \gamma c_i + \delta_t + \varepsilon_i \tag{II-1}$$

$Investment\ Behavior_i$ is an indicator variable for either (i) $Stock\ Market\ Participation$ or (ii) $Risky\ Asset\ Share$.

The key explanatory variable is the dummy variable $Direct\ Bank$. c_i represents the vector of control variables, which captures all client characteristics introduced in Table II-1. The term δ_t captures the wave fixed effects, accounting for time-specific trends or shocks that are common to all households within a particular survey wave.

Table II-7: Regression on Investment Variables

	Dependent Variable	
	(1) Stock Market Participation	(2) Risky Asset Share
Direct Bank	0.0878***	0.0296**
	(0.0259)	(0.0132)
Age	0.0009***	0.0015***
	(0.0004)	(0.0002)
Male	0.0532***	0.0174***
	(0.0127)	(0.0058)
Small-Town Resident	0.0104	-0.0089
	(0.0174)	(0.0070)
Household Income (yearly)	0.0431***	0.0078**
	(0.0098)	(0.0038)
Household Net Wealth	0.0453***	0.0222***
	(0.0110)	(0.0052)
University Degree	0.1664***	0.0601***
	(0.0173)	(0.0080)
Financial Literacy	0.0681***	0.0248***
	(0.0059)	(0.0028)
Satisfaction with Life	0.0245***	0.0086***
	(0.0029)	(0.0014)
Risk Tolerance (general)	0.0143***	0.0077***
	(0.0030)	(0.0014)
Trust	0.0023	0.0001
	(0.0030)	(0.0014)
Impatience	0.0021	0.0003
	(0.0023)	(0.0010)
Constant	-0.2869***	-0.1670***
	(0.0395)	(0.0213)
Observations	14,624	14,624
Adj. R^2	0.1271	0.0849
F-Statistic	65.6432***	31.6596***

Notes: This table reports our regression results for two OLS/LPM regressions. In column (1) $Stock\ Market\ Participation$ is regressed on $Direct\ Bank$ and relevant control variables. In column (2) the $Risky\ Asset\ Share$ is regressed on $Direct\ Bank$ and relevant control variables. For interpretability, we z-standardized $Household\ Income$ and $Household\ Net\ Wealth$ (i.e., subtracted their means and divided by their standard deviations) prior to including them in the model. All models include wave fixed effects to account for systematic differences across survey waves. Standard errors are adjusted for imputation error and survey design and reported in parentheses. The data are weighted using the sampling weights included in the sample. Significance levels are denoted as follows: $^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$.

The first model examines the relationship between being a direct bank client and stock market participation. We estimate a linear probability model (LPM), as the mean of the dependent binary variable is 25.6%.[11] Again, we find a strong correlation between stock market

[11] We opt for the linear model over the logistic model because, for the logistic model to provide a better fit, the log odds must be a linear function of the regressors, whereas the probability must not. This condition is met if the relationship between the

participation and being a direct bank client. Specifically, the coefficient for $Direct\ Bank$ is 8.78 percentage points, indicating that direct bank clients exhibit higher stock market participation, even after controlling for a range of client characteristics.

Regarding our control variables, we confirm earlier findings on factors affecting stock market participation.[12] Consistent with previous literature (e.g., Almenberg & Dreber, 2015; Ke, 2021; Sundén & Surette, 1998), we find that men are more likely to invest in the stock market, with a corresponding coefficient of 5.32 percentage points. In addition, we also document a positive association of financial literacy (van Rooij et al., 2011), education (Cole et al., 2014), income (Kaustia et al., 2023), and net wealth (Shum & Faig, 2006) with the likelihood of being invested in the stock market. Also consistent with prior findings in the literature (e.g., Bertaut, 1998; Haliassos & Bertaut, 1995; Shum & Faig, 2006), we document that clients with higher general risk tolerance tend to invest more often in the stock market. Finally, we find a highly significant relationship between higher life satisfaction and $Stock\ Market\ Participation$, aligning with Cui and Cho (2019), who find similar results regarding happiness.

In column (2), we analyze the correlation between being a direct bank client and $Risky\ Asset\ Share$. We find that direct bank clients allocate a larger proportion of their financial assets to risky investments, with the corresponding coefficient being 2.96 percentage points. Given that the average of the dependent variable is 8.82% (see Table II-1), this difference is not only statistically significant but also economically meaningful. Thus, even after controlling for numerous factors, being a direct bank client still corresponds with allocating a greater share of financial assets to risky investments.

---

probability and the log odds is nonlinear. When the probability falls between 0.20 and 0.80, the log odds are nearly a linear function of the probability (e.g., Long & Freese, 2014). The mean of our dependent variable, $Stock\ Market\ Participation$, is 0.256, which falls within the specified range. This suggests that the linear and logistic models provide a comparable fit, and the linear model is preferable due to its ease of interpretation (Hellevik, 2009; Wooldridge, 2010).

[12] Generally, our results are consistent with Bucher-Koenen et al. (2024), who employ the 4th wave of the PHF dataset.

With regard to the control variables, we observe similar patterns as those found for our first dependent variable, $Stock\ Market\ Participation$. Specifically, we also find positive relationships between financial literacy, education, income, net wealth, and the $Risky\ Asset\ Share$. Additionally, older clients tend to invest more in risky financial assets. Furthermore, higher levels of risk tolerance and satisfaction with life are positively associated with a greater allocation to risky investments.

Summing up, we find a very strong association between being a direct bank client and $Stock\ Market\ Participation$. We also observe a positive correlation between $Direct\ Bank$ and the proportion of financial assets invested in risky assets, although this relationship is less pronounced.

4.3 Heterogeneity Analysis

Given that the relationship between being a direct bank client and investment behavior may vary across individuals, we examine whether—and to what extent—this association differs across subgroups. To assess potential heterogeneity, we interact our key explanatory variable, $Direct\ Bank$, with each control variable included in Table II-7.[13] All metric variables are dichotomized using median splits, with the suffix "_high" indicating observations with values above the median for these variables. We estimate the following linear regression model:

$$Investment\ Behavior_i = \beta_0 + \beta_1\ Direct\ Bank_i + \beta_2\ [Indicator\ variable_i] + \beta_3\ Direct\ Bank_i \times [Indicator\ variable_i] + \gamma c_i + \delta_t + \varepsilon_i \quad (2)$$

Table II-8 presents the results, with each row corresponding to a different indicator variable. For instance, in Panel A, β_1 in the first row reports the relationship between being a direct bank client and the likelihood of stock market participation for younger individuals (those below the median age i.e., $Age_{high} = 0$). $\beta_1 + \beta_3$ represents the association for the subsample of the 50%

[13] The methodology is inspired by Brenner et al. (2020), who employ a similar estimation technique to test for heterogeneous treatment effects of consumer fraud on financial well-being.

older individuals, while β_3 indicates the difference in the relationships (i.e., heterogeneity) between younger and older bank clients. Similarly, the second row reports the betas for female clients (β_1), male clients ($\beta_1 + \beta_3$), and the difference between estimated association for the two groups (β_3). Panel B presents the results for $Share\ Risky\ Assets$. It is important to note that the association of being a direct bank client with investment behavior was most pronounced in relation to stock market participation. Therefore, we anticipate the most robust and meaningful results from this specification of investment behavior.

Table II-8: Regression on Investment Variables

Panel A: Dependent Variable: Stock Market Participation					
	β_1	$\beta_1 + \beta_3$	β_3	N	Adj. R^2
Age_high	0.1695^{***}	0.1021^{**}	-0.0674	14,624	0.0079
	(0.0381)	(0.0398)	(0.0536)		
Male	0.0932^{**}	0.1901^{***}	0.0970^{*}	14,624	0.0174
	(0.0393)	(0.0385)	(0.0539)		
Small-Town Resident	0.1253^{***}	0.2547^{***}	0.1294	14,624	0.0080
	(0.0300)	(0.0759)	(0.0814)		
Household Income (yearly)_high	0.1499^{**}	0.1189^{***}	-0.0310	14,624	0.0366
	(0.0604)	(0.0324)	(0.0694)		
Household Net Wealth_high	0.1608^{***}	0.1538^{***}	-0.0070	14,624	0.0744
	(0.0404)	(0.0391)	(0.0576)		
University Degree	0.0942^{***}	0.0842^{**}	-0.0100	14,624	0.0570
	(0.0352)	(0.0429)	(0.0545)		
Financial Literacy_high	0.1570^{***}	0.0931^{**}	-0.0638	14,624	0.0078
	(0.0315)	(0.0458)	(0.0501)		
Satisfaction with Life_high	0.1522^{***}	0.1181^{***}	-0.0342	14,624	0.0095
	(0.0348)	(0.0423)	(0.0515)		
Risk Tolerance (general)_high	0.1213^{***}	0.1424^{***}	0.0211	14,624	0.0120
	(0.0440)	(0.0352)	(0.0542)		
Trust_high	0.1952^{***}	0.1050^{***}	-0.0902	14,624	0.0109
	(0.0471)	(0.0335)	(0.0555)		
Impatience_high	0.1762^{***}	0.1022^{***}	-0.0740	14,624	0.0080
	(0.0386)	(0.0371)	(0.0504)		

Panel B: Dependent Variable: Risky Asset Share					
	β_1	$\beta_1 + \beta_3$	β_3	N	Adj. R^2
Age_high	0.0395***	0.0425*	0.0030	14,624	0.0062
	(0.0133)	(0.0242)	(0.0267)		
Male	0.0092	0.0632***	0.0540**	14,624	0.0075
	(0.0153)	(0.0212)	(0.0259)		
Small-Town Resident	0.0297**	0.0827**	0.0530	14,624	0.0026
	(0.0148)	(0.0331)	(0.0367)		
Household Income (yearly)_high	0.0330	0.0327**	-0.0003	14,624	0.0106
	(0.0247)	(0.0159)	(0.0291)		
Household Net Wealth_high	0.0425**	0.0413**	-0.0012	14,624	0.0390
	(0.0169)	(0.0202)	(0.0268)		
University Degree	0.0317*	0.0010	-0.0307	14,624	0.0291
	(0.0185)	(0.0202)	(0.0273)		
Financial Literacy_high	0.0428***	0.0172	-0.0256	14,624	0.0021
	(0.0153)	(0.0201)	(0.0220)		
Satisfaction with Life_high	0.0499***	0.0143	-0.0357*	14,624	0.0035
	(0.0174)	(0.0169)	(0.0215)		
Risk Tolerance (general)_high	0.0392*	0.0304*	-0.0089	14,624	0.0055
	(0.0203)	(0.0177)	(0.0266)		
Trust_high	0.0833***	0.0107	-0.0726**	14,624	0.0039
	(0.0277)	(0.0129)	(0.0288)		
Impatience_high	0.0527***	0.0184	-0.0343	14,624	0.0021
	(0.0192)	(0.0149)	(0.0214)		

Notes: This table reports coefficient estimates obtained from a linear regression model of the generic form: $Investment\ Behavior_i = \beta_0 + \beta_1\ Direct\ Bank_i + \beta_2\ [Indicator\ variable_i] + \beta_3\ Direct\ Bank_i \times [Indicator\ variable_i] + \gamma c_i + \delta_t + \varepsilon_i$ For each indicator variable (e.g., Age_{high}), β_1 captures the estimated association of $Direct\ Bank$ with the investment behavior for the subgroup with $Age_{high} = 0$ (below-median age). $\beta_1 + \beta_3$ represents the corresponding association for the subgroup with $Age_{high} = 1$ (above-median age), while β_3 indicates the difference in these associations across the two subgroups. All metric variables are split at the median, and the suffix "_high" indicates values above the median. Panel A focuses on $Stock\ Market\ Participation$, whereas Panel B presents results for the $Risky\ Asset\ Share$. All models include wave fixed effects to account for systematic differences across survey waves. Standard errors (in parentheses) are adjusted for imputation error, following Rubin's Rules, and for survey design, using replicate weights. The data are weighted using the sampling weights included in the sample. Significance levels are denoted as follows: $^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$.

As can be inferred from Panel A, our findings regarding stock market participation are highly robust. Specifically, we observe that β_1 and $\beta_1 + \beta_3$ remain statistically significant in all specifications, suggesting a consistent positive relationship between $Direct\ Bank$ and $Stock\ Market\ Participation$ across all subgroups. β_3 is insignificant for almost all subgroups, indicating limited heterogeneity. The sole exception is $Male$, where a significant β_3 indicates a stronger association between being a direct bank client and participation in the stock market for men than for women.

Panel B of Table II-8 replicates the analysis using $Risky\ Asset\ Share$ as the dependent variable. The results are more varied, with approximately 60% of the subsamples showing a significant correlation between being a direct bank client and the $Risky\ Asset\ Share$. While β_1 is slightly more likely to be significant than $\beta_1 + \beta_3$, we only find significant coefficients for β_3 in three subgroups, indicating that the association between being a direct bank client and the proportion invested in risky assets is similar across most subsamples. Given that our baseline estimations (see Table II-7) only detected a weaker association between $Direct\ Bank$ and the share of risky assets, these subgroup findings are not unexpected. Moreover, for subgroups with fewer observations (e.g., bank clients residing in small towns), concerns about statistical power further caution against drawing firm conclusions.

5. Robustness Checks

While our main results consistently indicate that direct bank clients are (i) systematically different from traditional bank clients with regard to their demographic characteristics and financial literacy and (ii) more likely to invest in the stock market and hold a higher fraction of risky assets, several potential methodological concerns warrant further analysis. In this section, we address these concerns through a series of robustness tests. By doing so, we aim to ensure the reliability of our conclusions.

Propensity Score Matching

We address the potential selection bias associated with $Direct\ Bank$, as clients of direct banks may significantly differ from traditional bank clients in terms of observable covariates. These differences suggest varying probabilities of selecting into being a direct bank client. To address this issue, we apply propensity score matching (PSM) to account for the potential selection bias in individuals' likelihood of being direct bank clients. In column (4) of Table II-5, we estimate a logistic regression with $Direct\ Bank$ as the dependent variable. Now, we employ these coefficients to calculate propensity scores. We then construct a control sample of traditional bank clients by matching each direct bank client with a traditional bank client 'twin' based on similar covariates, using a 1:1 nearest-neighbor matching approach without replacement, restricting matches to occur within the same survey wave. This approach results in a matched sample with well-balanced covariates between direct bank clients (treated) and traditional bank clients (controls).[14] Subsequently, we re-estimate the linear probability models on the investment behavior variables of Table II-7 using the matched (balanced) sample.

As shown in Table II-9, our findings for $Stock\ Market\ Participation$ remain robust in the matched sample. However, the coefficient for $Risky\ Asset\ Share$ becomes statistically insignificant. Since the coefficient's magnitude is similar to our main analysis, we cannot conclusively determine if this is due to a genuine lack of association or the reduced statistical power of the smaller sample.

[14] See Appendix II-B for the post-matching diagnostic test.

Table II-9: Regression on Investment Variables with Matched Sample

	Dependent Variable	
	(1) Stock Market Participation	(2) Risky Asset Share
Direct Bank	0.0937***	0.0270
	(0.0339)	(0.0180)
Age	-0.0003	0.0012
	(0.0015)	(0.0008)
Male	0.1053***	0.0416**
	(0.0371)	(0.0169)
Small-Town Resident	0.0154	-0.0025
	(0.0573)	(0.0256)
Household Income (yearly)	0.0299	0.0001
	(0.0297)	(0.0113)
Household Net Wealth	0.0765***	0.0314***
	(0.0290)	(0.0118)
University Degree	0.1610***	0.0595***
	(0.0417)	(0.0188)
Financial Literacy	0.0712**	0.0246**
	(0.0297)	(0.0117)
Satisfaction with Life	0.0237**	0.0042
	(0.0117)	(0.0044)
Risk Tolerance (general)	0.0224**	0.0128**
	(0.0112)	(0.0050)
Trust	-0.0049	-0.0065
	(0.0129)	(0.0062)
Impatience	-0.0022	-0.0011
	(0.0083)	(0.0037)
Constant	-0.1926	-0.0816
	(0.1524)	(0.0780)
Observations	1,694	1,694
Adj. R^2	0.1628	0.1015
F-Statistic	12.6930***	7.1371***

Notes: This table reports our regression results for two OLS/LPM regressions estimated on a matched sample. In column (1) $Stock\ Market\ Participation$ is regressed on $Direct\ Bank$ and relevant control variables. In column (2) the $Risky\ Asset\ Share$ is regressed on $Direct\ Bank$ and relevant control variables. For interpretability, we z-standardized $Household\ Income$ and $Household\ Net\ Wealth$ (i.e., subtracted their means and divided by their standard deviations) prior to including them in the model. All models include wave fixed effects to account for systematic differences across survey waves. Standard errors (in parentheses) are adjusted for imputation error, following Rubin's Rules, and for survey design, using replicate weights. The data are weighted using the sampling weights included in the sample. Significance levels are denoted as follows: $^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$.

Single-Person Households

Our primary analysis is mostly based on the characteristics of the FKP within a household, which may include multiple members. In multi-person households, joint decision-making could complicate the link between demographic variables, bank choice and investment behavior. This can introduce complexity when, for example, both spouses make financial decisions together, but our setup only allows us to capture the demographic information of the FKP. To assess whether household decision-making might distort our results, we restrict our sample to single-person households for this robustness test. In single-person households, there is no ambiguity about which household member makes the financial decision.

In Appendix II-C, we report our results on the link between client characteristics and bank choice. While most associations remain robust, a notable change occurs for net wealth and income. Specifically, the links between being a direct bank client and both household net wealth and income become statistically insignificant.

In Appendix II-D, we investigate the effect of being a direct bank client on the two investment variables. Consistent with our main findings, we again observe a significant relationship between $Direct\ Bank$ and $Stock\ Market\ Participation$. The association between being a direct bank client and the $Risky\ Asset\ Share$ on the other hand becomes insignificant. This is not entirely unexpected, given that the baseline analysis showed a weaker link for this outcome.

Logistic Regression for Stock Market Participation

For ease of interpretation, we hitherto relied on a linear probability model when examining the association between $Stock\ Market\ Participation$ and being a direct bank client in Panel A of Table II-7. However, to ensure that our results are not driven by model choice, we re-estimate the regression using a logistic regression instead of an LPM. The sign, magnitude, and significance of our main coefficient of interest – the indicator for being a direct bank client – remain broadly consistent, validating that our core results do not hinge on using a linear versus nonlinear estimator. See Appendix II-E for more details.

Extended Sample

So far, we have restricted our sample to households that indicate that they have only one house bank. To assess the robustness of our findings, we now also include all other observations for which the dataset is complete for all variables of interest. This includes households that indicate that they have more than one house bank relationship or that indicate that they have no house bank relationship at all. In doing so, we also redefine our $Direct\ Bank$ variable. Previously, it was set to 1 only if a household had an exclusive house bank relationship with a direct bank and 0 if it had an exclusive relationship with a traditional bank; all other households were excluded from the analysis. Under the new definition, $Direct\ Bank$ equals 1 if a household maintains any house bank relationship with a direct bank (even if it also has additional relationships) and 0 otherwise, which means the control group now also includes households without any house bank relationships.

Our topline results remain robust to this new definition of the $Direct\ Bank$ variable, with the exception being the association between $Direct\ Bank$ and $Risky\ Asset\ Share$. The tables can be found in Appendices F and G.

Re-estimation without survey weights

To ensure that our results are not driven by the survey weights included in the data, we re-estimate all regressions using the same data but ignoring the survey weights. While some secondary results become insignificant in this specification (e.g., the association between $Male$ and $Direct\ Bank$ is no longer significant in the full model), our mainline results remain unchanged. We omit these unweighted results from the paper because the survey weighting is integral to the design and renders unweighted estimates less meaningful. However, the unweighted results are available upon request.

Taken together, these exercises indicate that our main conclusions are robust. With few exceptions, the associations between different client characteristics and bank choice remain statistically significant. The positive link between being a direct bank client and $Stock\ Market\ Participation$ persists across all alternative samples and specifications. For the $Risky\ Asset\ Share$, the estimated coefficient remains of similar magnitude but loses statistical significance in several robustness tests, indicating sensitivity in inference rather than a shift in the underlying relationship.

6. Discussion and Implications

6.1 Discussion of Results

In this study, we show that client characteristics and investment behavior differ greatly between direct bank and traditional bank clients. Our findings indicate that direct bank clients are significantly younger, possess higher levels of financial literacy, are better educated, and exhibit higher levels of risk tolerance. Filotto et al. (2021), who study which attributes matter for clients when deciding whether to use a direct bank, also offer a glimpse into the different characteristics of direct bank clients and traditional bank clients. In their sample of Italian bank clients, they find that direct bank clients are younger and better educated. However, they also observe that they are less likely to be male and have lower income than traditional bank clients. Unfortunately, Filotto et al. (2021) do not focus on the characteristics of the clients themselves and do not report the significance of their differences; therefore, it is unclear how well their results generalize to the public.

Since the differences in client characteristics between direct and traditional bank clients have otherwise not been previously analyzed in the literature, these findings cannot be directly compared to prior research. However, we can relate our findings to factors fostering the adoption of digital financial services like internet banking, mobile banking, and other retail finance innovations. Kolodinsky et al. (2004) show that higher-income, better-educated, and relatively younger U.S. clients are significantly more likely to adopt internet banking or to intend to adopt internet banking, thus displaying a similar openness to new technologies as these clients do in our sample. Lassar et al. (2005), on the other hand, do not document a significant effect of age or education on internet banking adoption. They do, however, also find that income is positively related to internet banking adoption.

In the context of mobile banking, Becker et al. (2022) report for a large sample of bank clients in Germany that age is also negatively associated with mobile banking adoption. The authors also document that gender has no significant effect on mobile banking adoption, a result consistent with our findings in this study for direct bank clients. Laukkanen and Pasanen (2008) similarly report a negative effect of age on the adoption of mobile banking but also note that men are significantly overrepresented among adopters, a finding that contradicts the results of Becker et al. (2022) on mobile banking and our results for direct bank clients.

More recently, several studies have looked at the characteristics and investment behavior of neobroker clients. Although neobrokers differ in their service offerings and user interfaces, they share one key feature with direct banks: clients make investment decisions without in-person advice. Regarding client characteristics, Freibauer et al. (2024) find that neobroker users are younger and more risk-tolerant than investors at traditional brokers, consistent with our findings. They do not, however, observe significant differences in financial literacy. Barber et al. (2022) report that neobroker clients tend to be less experienced, with half of them being first-time investors. This pattern may not necessarily reflect lower financial literacy but indicates lower practical investment experience.

In terms of investment behavior, we find strong evidence for direct bank client status being associated with a higher likelihood of stock market participation. Notably, direct bank clients exhibit an 8.78 percentage point higher stock market participation rate, even after controlling for a wide range of variables known to be associated with stock market participation. We observe this positive association between being a direct bank client and stock market participation across nearly all subgroups of bank clients, including both male and female clients.

Furthermore, the association remains robust when applying a nearest neighbor matching estimation.

As no prior study has analyzed this specific relationship, we cannot directly relate our findings on direct bank clients to the literature. However, our results do align well with broader evidence on how digital access influences equity investing. For example, Bogan (2008) shows that internet usage fosters stock market participation, supporting the notion that individuals accustomed to digital banking channels – including direct banks – would be more likely to invest in stocks. Additionally, we can confirm many previously established determinants of stock market participation. From the multivariate analyses of our data on German bank clients, we find that male clients are more likely to invest in the stock market and that income, wealth, and financial literacy are key factors for participation.

Regarding the $Risky\ Asset\ Share$, we report evidence that direct bank clients allocate a larger portion of their financial assets to risky assets, such as stocks, mutual funds, and bonds, even after controlling for a wide range of variables. Specifically, direct bank clients invest 2.96 percentage points more of their financial assets in risky assets, which is substantial given the sample's average of 8.82%. Thus, direct bank clients capture roughly 33% more of the equity premium than clients of traditional banks.

In contrast, research on neobrokers, such as Barber et al. (2022), shows that clients chase attention-grabbing stocks, concentrate trades in recent "top movers," and suffer negative short-term reversals. Whether these patterns extend to direct-bank clients is uncertain: both channels lack in-person financial advice, yet neobrokers overlay this absence with gamified, mobile-first interfaces that can amplify speculative trading incentives.

Our study is, of course, not without limitations. Specifically, the granularity of the survey data is limited. For instance, we have no information on which specific banks clients maintain their relationship with. We are therefore unable to differentiate between neobank and ordinary direct bank clients. Given that neobank clients may differ in meaningful ways (e.g., digital savviness, experience with financial markets) from clients of ordinary direct banks, our findings may mask important heterogeneity. Additionally, our data do not capture how frequently or thoroughly traditional bank clients access their banks' advisory services. Since many traditional banks also offer online brokerage, some traditional customers may effectively behave like direct bank clients. Likewise, direct bank clients may seek advice from third parties (e.g., independent financial planners or online forums). As a result, the distinction between "advised" and "DIY" investors is not as sharp as our data might suggest.

Furthermore, our study does not—and does not intend to—demonstrate a causal relationship between bank choice and investment behavior. Because bank choice is endogenous, any observed correlation is likely influenced by unobservable factors that affect both bank choice and investment behavior. Although our robustness checks address some concerns, fully isolating causality is beyond the scope of our analysis. Future research could, for example, track households that are forced to switch to direct banks due to branch closures; such a quasi-experimental setting would offer stronger evidence on whether bank choice itself drives changes in financial behavior.

6.2 Implications

While our findings are not causal, they are still highly relevant for both consumer protection and bank managers. Regarding the potential vulnerability of direct bank clients, we document that, on average, these clients are relatively well-equipped to navigate financial decision-making in a DIY setting. Specifically, direct bank clients exhibit several positive attributes, such as higher levels of education and financial literacy, which make them less vulnerable compared to the average client of a traditional bank. With regard to investment behavior, we find that direct bank clients, on average, are more likely to participate in the equity risk premium, positioning them favorably for wealth accumulation.

With respect to recommendations for regulatory authorities, we conclude that consumer protection agencies do not have a significant need to implement special regulations for direct bank clients. These clients appear to be less vulnerable compared to traditional bank clients. Furthermore, given their high levels of financial literacy, direct bank clients tend to make sound investment decisions, as shown by their greater participation in the stock market and higher allocation to risky financial assets.

From a bank management perspective, traditional banks should be concerned about the appeal of direct banks to younger, potentially more attractive clients. In particular, traditional banks face challenges due to an aging client base, with a significant portion of their clients being of retirement age. This demographic shift poses a long-term problem if traditional banks cannot retain and attract younger clients, including those currently with direct banks. Given that traditional banks will likely always face a cost disadvantage compared to direct banks, they must emphasize their unique value proposition – namely, personalized financial advice and the human touch in the client-advisor relationship.

7. References

Almenberg, J., & Dreber, A. (2015). Gender, stock market participation and financial literacy. Economics Letters, 137, 140–142. https://doi.org/10.1016/j.econlet.2015.10.009

Balloch, A., Nicolae, A., & Philip, D. (2015). Stock market literacy, trust, and participation. *Review of Finance*, *19*(5), 1925–1963. https://doi.org/10.1093/rof/rfu040

Barber, B. M., Huang, X., Odean, T., & Schwarz, C. (2022). Attention-induced trading and returns: Evidence from Robinhood users. *The Journal of Finance*, *77*(6), 3141–3190. https://doi.org/10.1111/jofi.13183

Barber, B. M., & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics*, *116*(1), 261–292. https://doi.org/10.1162/003355301556400

Becker, M., Stolper, O., & Walter, A. (2022). What drives mobile banking adoption? – An empirical investigation using transaction data. *Zeitschrift Für Bankrecht Und Bankwirtschaft*, *34*(1), 1–11. https://doi.org/10.15375/zbb-2022-0103

Benartzi, S., & Thaler, R. H. (1995). Myopic loss aversion and the equity premium puzzle. *The Quarterly Journal of Economics*, *110*(1), 73–92. https://doi.org/10.2307/2118511

Bertaut, C. C. (1998). Stockholding behavior of U.S. households: Evidence from the 1983-1989 Survey of Consumer Finances. *Review of Economics and Statistics*, *80*(2), 263–275. https://doi.org/10.1162/003465398557500

Black, S. E., Devereux, P. J., Lundborg, P., & Majlesi, K. (2018). Learning to take risks? The effect of education on risk-taking in financial markets. *Review of Finance*, *22*(3), 951–975. https://doi.org/10.1093/rof/rfy005

Bogan, V. (2008). Stock market participation and the internet. *Journal of Financial and Quantitative Analysis*, *43*(1), 191–211. http://www.jstor.org/stable/27647344

Brenner, L., Meyll, T., Stolper, O., & Walter, A. (2020). Consumer fraud victimization and financial well-being. *Journal of Economic Psychology*, *76*, 102243. https://doi.org/10.1016/j.joep.2019.102243

Bucher-Koenen, T., Janssen, B., Knebel, C., & Tzamourani, P. (2024). Financial literacy, stock market participation, and financial wellbeing in Germany. *Journal of Financial Literacy and Wellbeing*, *1*(3), 486–513. https://doi.org/10.1017/flw.2024.5

Chalmers, J., & Reuter, J. (2020). Is conflicted investment advice better than no advice? *Journal of Financial Economics*, *138*(2), 366–387. https://doi.org/10.1016/j.jfineco.2020.05.005

Citterio, A., Marques, B. P., & Tanda, A. (2025). The early days of neobanks in Europe: Identification, performance, and riskiness. *Journal of Financial Services Research*, *68*, 153–199. https://doi.org/10.1007/s10693-024-00433-x

Cole, S., Paulson, A., & Shastry, G. (2014). Smart money? The effect of education on financial outcomes. *Review of Financial Studies*, *27*(7), 2022–2051. https://EconPapers.repec.org/RePEc:oup:rfinst:v:27:y:2014:i:7:p:2022-2051.

Cui, W., & Cho, I. (2019). Household's happiness and financial market participation. *Global Economic Review*, *48*(4), 396–418. https://doi.org/10.1080/1226508X.2019.1640629

Delis, M. D., & Mylonidis, N. (2015). Trust, happiness, and households' financial decisions. *Journal of Financial Stability*, *20*, 82–92. https://doi.org/10.1016/j.jfs.2015.08.002

Deutsche Bundesbank. (2023, July 7). *Changes in bank office statistics in 2022*. https://www.bundesbank.de/en/press/press-releases/changes-in-bank-office-statistics-in-2022-915502

Deutsche Bundesbank. (2024, December 1). *Payment behaviour in Germany in 2023*. https://www.bundesbank.de/en/publications/reports/studies/payment-behaviour-in-germany-in-2023-934896

Deutsche Bundesbank. (2025a). *Panel on Household Finances Wave 1- Wave 5.* https://doi.org/10.12757/BBK.PHF.W1-W5.01.01

Deutsche Bundesbank. (2025b, February 17). *Banking statistics*. https://www.bundesbank.de/en/statistics/banks-and-other-financial-corporations/banks/banking-statistics-816050

Deutsches Aktien Institut. (2024, January 11). *Aktionärszahlen 2023: Erneut über 12 Millionen. Aktiensparer trotzen der Zinswende*. https://www.dai.de/detail/aktionaerszahlen-2023-erneut-ueber-12-millionen-aktiensparer-trotzen-der-zinswende

Federal Statistical Office. (2024, June 14). *Population by territory and average age*. https://www.destatis.de/EN/Themes/Society-Environment/Population/Current-Population/Tables/population-by-territory-and-average-age.html

Filotto, U., Caratelli, M., & Fornezza, F. (2021). Shaping the digital transformation of the retail banking industry. Empirical evidence from Italy. *European Management Journal*, *39*(3), 366–375. https://doi.org/10.1016/j.emj.2020.08.004

Frederick, S., Loewenstein, G., & O'donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature*, *40*(2), 351–401. https://doi.org/10.1257/002205102320161311

Freibauer, J., Grawert, S., & Rieger, M. O. (2024). The effects of trading apps on investment behavior over time. *The European Journal of Finance*, 1–25. https://doi.org/10.1080/1351847X.2024.2401604

Georgarakos, D., & Pasini, G. (2011). Trust, sociability, and stock market participation. *Review of Finance*, *15*(4), 693–725. https://doi.org/10.1093/rof/rfr028

Guiso, L., Sapienza, P., & Zingales, L. (2008). Trusting the stock market. *The Journal of Finance*, *63*(6), 2557–2600. https://doi.org/10.1111/j.1540-6261.2008.01408.x

Haliassos, M., & Bertaut, C. C. (1995). Why do so few hold stocks? *The Economic Journal*, *105*(432), 1110–1129. https://doi.org/10.2307/2235407

Hellevik, O. (2009). Linear versus logistic regression when the dependent variable is a dichotomy. *Quality & Quantity*, *43*(1), 59–74. https://doi.org/10.1007/s11135-007-9077-3

Kaustia, M., Conlin, A., & Luotonen, N. (2023). What drives stock market participation? The role of institutional, traditional, and behavioral factors. *Journal of Banking & Finance*, *148*, 106743. https://doi.org/10.1016/j.jbankfin.2022.106743

Ke, D. A. (2021). Who wears the pants? Gender identity norms and intrahousehold financial decision-making. *The Journal of Finance*, *76*(3), 1389–1425. https://doi.org/10.1111/jofi.13002

Kolodinsky, J. M., Hogarth, J. M., & Hilgert, M. A. (2004). The adoption of electronic banking technologies by US consumers. *International Journal of Bank Marketing*, *22*(4), 238–259. https://doi.org/10.1108/02652320410542536

Lassar, W. M., Manolis, C., & Lassar, S. S. (2005). The relationship between consumer innovativeness, personal characteristics, and online banking adoption. *International Journal of Bank Marketing*, *23*(2), 176–199. https://doi.org/10.1108/02652320510584403

Laukkanen, T., & Pasanen, M. (2008). Mobile banking innovators and early adopters: How they differ from other online users? *Journal of Financial Services Marketing*, *13*(2), 86–94. https://doi.org/10.1057/palgrave.fsm.4760077

Long, J. S., & Freese, J. (2014). *Regression models for categorical dependent variables using Stata* (3rd ed.). *A Stata Press publication*. Stata Press.

Luo, J., & Subrahmanyam, A. (2019). The affect heuristic and stock ownership: A theoretical perspective. *Review of Financial Economics*, *37*(1), 6–37. https://doi.org/10.1002/rfe.1026

Lusardi, A., & Mitchell, O. S. (2008). Planning and financial literacy: How do women fare? *The American Economic Review*, *98*(2), 413–417. https://doi.org/10.1257/aer.98.2.413

Lusardi, A., & Mitchell, O. S. (2014). The economic importance of financial literacy: Theory and evidence. *Journal of Economic Literature*, *52*(1), 5–44. https://doi.org/10.1257/jel.52.1.5

Mankiw, N., & Zeldes, S. P. (1991). The consumption of stockholders and nonstockholders. *Journal of Financial Economics*, *29*(1), 97–112. https://doi.org/10.1016/0304-405X(91)90015-C

Mullainathan, S., Noeth, M., & Schoar, A. (2012). The market for financial advice: An audit study(17929). https://EconPapers.repec.org/RePEc:nbr:nberwo:17929

Patterson, F. M., & Daigler, R. T. (2014). The abnormal psychology of investment performance. *Review of Financial Economics*, *23*(2), 55–63. https://doi.org/10.1016/j.rfe.2013.08.004

Sarel, D., & Marmorstein, H. (2003). Marketing online banking services: The voice of the customer. *Journal of Financial Services Marketing*, *8*(2), 106–118. https://doi.org/10.1057/palgrave.fsm.4770111

Schmidt, T., & Eisele, M. (2013). Oversampling vermögender Haushalte im Rahmen der Studie "Private Haushalte und ihre Finanzen (PHF)". *Deutsche Bundesbank*.

Shum, P., & Faig, M. (2006). What explains household stock holdings? *Journal of Banking & Finance*, *30*(9), 2579–2597. https://doi.org/10.1016/j.jbankfin.2005.11.006

Statista. (2024a). *Germany: Willingness to change primary bank 2024 | Statista*. https://www.statista.com/statistics/1395372/willingness-to-change-primary-bank-in-germany/

Statista. (January 2024b). *Global market volume of neobanks | Statista*. https://www.statista.com/statistics/1228241/neobanks-global-market-size/

Sundén, A. E., & Surette, B. J. (1998). Gender differences in the allocation of assets in retirement savings plans. *The American Economic Review*, *88*(2), 207–211. http://www.jstor.org/stable/116920

van Rooij, M., Lusardi, A., & Alessie, R. (2011). Financial literacy and stock market participation. *Journal of Financial Economics*, *101*(2), 449–472. https://doi.org/10.1016/j.jfineco.2011.03.006

von Kalckreuth, U., Eisele, M., Le Blanc, J., Schmidt, T., & Zhu, J. (2012). The PHF: A comprehensive panel survey on household finances and wealth in Germany. *Deutsche Bundesbank*(13/2012). https://hdl.handle.net/10419/59999 (Bundesbank Discussion Paper).

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). The MIT Press. http://www.jstor.org/stable/j.ctt5hhcfr

Zhan, J. C. (2015). Who holds risky assets and how much? *Empirica*, *42*(2), 323–370. https://doi.org/10.1007/s10663-015-9295-1

8. Appendix

Appendix II-A: Variable Dimensions and Descriptions

Variable	Dimension	Description
Panel A: House Bank		
Direct Bank	1/0	Indicator variable that equals 1 if the household's house bank is a direct bank and 0 if it is a traditional bank
Panel B: Investment Behavior		
Stock Market Participation	1/0	Indicator variable that equals 1 if any member of the household owns either stocks or mutual funds
Risky Asset Share	[0%-100%]	The sum of bonds, stocks, and mutual funds as a percentage of total financial assets
Panel C: Demographics		
Age	[17;90]	Indicates the age of the financially knowledgeable person (FKP) of the household in years
Male	1/0	Indicator variable that equals 1 if the financially knowledgeable person of the household is male
Small-Town Resident	1/0	Indicator variable that equals 1 if the household lives in a municipality with less than 5,000 inhabitants
Household Income (yearly)	€k	Indicates the total annual gross household income in thousand €
Household Net Wealth	€k	Indicates the total net wealth of all household members in thousand €
University Degree	1/0	Indicator variable that equals 1 if the financially knowledgeable person of the household has a university degree
Panel D: Financial Literacy		
Financial Literacy	[0;3]	Indicates the score of the financially knowledgeable person in a three-question battery designed to test their financial literacy.
Financial Literacy = 3	1/0	Indicator variable that equals 1 if the financially knowledgeable person answered all three questions correctly
Panel E: Attitudes		
Satisfaction with Life	[0;10]	Self-rating on a scale from 0 to 10 by the financially knowledgeable person on their satisfaction with life, where 0 equals "completely dissatisfied" and 10 equals "completely satisfied"
Risk Tolerance (general)	[0;10]	Self-rating on a scale from 0 to 10 by the financially knowledgeable person on their willingness to take risks, where 0 equals "Not at all willing to take risks" and 10 equals "Very willing to take risks"
Trust	[0;10]	Self-rating on a scale from 0 to 10 by the financially knowledgeable person on their trust in other people, where 0 equals "do not trust others at all" and 10 equals "I trust others completely"
Impatience	[0;10]	Self-rating on a scale from 0 to 10 by the financially knowledgeable person on their impatience, where 0 equals "Very patient" and 10 equals "Very impatient"

Notes: This table reports the dimensions and definitions of all variables. Dimensions indicate the values the variable can take as well as in which unit they are reported.

Appendix II-B: Post Matching Diagnostic Test

Variable	Mean Direct Bank Post Matching	Mean Traditional Bank Post Matching	Absolute Standardized Mean Difference Post Matching
Panel A: Demographics			
Age	43.66	42.89	0.06
Male	51.32%	59.03%	0.15
Small-Town Resident	11.04%	11.76%	0.02
Household Income (yearly)	68.21	63.39	0.08
Household Net Wealth	187.97	195.67	0.04
University Degree	44.18%	44.87%	0.04
Panel B: Financial Literacy			
Financial Literacy	2.66	2.63	0.08
Financial Literacy = 3	72.52%	70.48%	0.05
Panel C: Attitudes			
Satisfaction with Life	7.46	7.48	0.03
Risk Tolerance (general)	4.75	4.70	0.08
Trust	6.03	5.78	0.13
Impatience	4.64	4.85	0.09

Notes: This table reports the mean for direct bank clients and traditional bank clients for all control variables after propensity score matching. Matching used nearest neighbor matching within each survey wave. The last column reports the absolute standardized mean difference after matching, averaged across imputations. Means are calculated using Rubin rules and the multiple imputations and survey weights included in the dataset.

Appendix II-C: Logistic Regression on Direct Bank for Single-Person Households

	Dependent Variable			
	(1) Direct Bank	(2) Direct Bank	(3) Direct Bank	(4) Direct Bank
Age	-0.0273***			-0.0251***
	(0.0061)			(0.0065)
Male	-0.1213			-0.2137
	(0.2169)			(0.2299)
Small-Town Resident	0.1477			0.1835
	(0.5516)			(0.5569)
Household Income (yearly)	0.0166			-0.0082
	(0.0776)			(0.0802)
Household Net Wealth	-0.2410			-0.2496
	(0.1593)			(0.1640)
University Degree	1.1542***			1.0536***
	(0.2180)			(0.2215)
Financial Literacy		0.5429***		0.2954
		(0.1858)		(0.1928)
Satisfaction with Life			-0.0205	-0.0465
			(0.0569)	(0.0625)
Risk Tolerance (general)			0.1364***	0.1157**
			(0.0454)	(0.0513)
Trust			0.0817	0.0237
			(0.0626)	(0.0655)
Impatience			-0.0133	-0.0234
			(0.0399)	(0.0420)
Constant	-2.5516***	-4.9170***	-4.3796***	-3.5708***
	(0.4800)	(0.5952)	(0.6362)	(0.9205)
Observations	3,915	3,915	3,915	3,915
Adj. Pseudo R^2	0.0887	0.0294	0.0234	0.0932

Notes: This table reports our regression results for a logistic regression with $Direct\ Bank$ as the dependent variable for a sample restricted to single-person households. In column (1) $Direct\ Bank$ is regressed on demographic variables. For interpretability, we z-standardized $Household\ Income$ and $Household\ Net\ Wealth$ (i.e., subtracted their means and divided by their standard deviations) prior to including them in the model. In column (2) $Direct\ Bank$ is regressed on the variable $Financial\ Literacy$, which indicates the score of the FKP in a three-question battery designed to test their financial literacy. In column (3) $Direct\ Bank$ is regressed on a set of variables describing the FKP attitudes. In column (4) all variables from the previous regressions are included in one model. All models include wave fixed effects to account for systematic differences across survey waves. We report the adjusted McFadden Pseudo R^2 for each specification. Standard errors are adjusted for imputation error and survey design and reported in parentheses. The data are weighted using the sampling weights included in the sample. Significance levels are denoted as follows: $^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$.

Appendix II-D: Regression on Investment Variables for Single-Person Households

	Dependent Variable	
	(1) Stock Market Participation	(2) Risky Asset Share
Direct Bank	0.1070**	0.0335
	(0.0485)	(0.0272)
Age	0.0002	0.0011***
	(0.0005)	(0.0003)
Male	0.0344	0.0146
	(0.0224)	(0.0111)
Small-Town Resident	0.0232	-0.0018
	(0.0294)	(0.0135)
Household Income (yearly)	0.0389***	0.0105
	(0.0146)	(0.0077)
Household Net Wealth	0.0587***	0.0361***
	(0.0174)	(0.0113)
University Degree	0.1754***	0.0546***
	(0.0273)	(0.0143)
Financial Literacy	0.0583***	0.0325***
	(0.0099)	(0.0049)
Satisfaction with Life	0.0245***	0.0107***
	(0.0049)	(0.0025)
Risk Tolerance (general)	0.0180***	0.0104***
	(0.0047)	(0.0026)
Trust	-0.0030	-0.0026
	(0.0045)	(0.0026)
Impatience	-0.0007	-0.0011
	(0.0039)	(0.0020)
Constant	-0.1950***	-0.1554***
	(0.0676)	(0.0381)
Observations	3,915	3,915
Adj. R^2	0.1459	0.1083
F-Statistic	21.9203***	13.4578***

Notes: This table reports our regression results for two OLS/LPM regressions for a sample restricted to single-person households. In column (1) *Stock Market Participation* is regressed on *Direct Bank* and relevant control variables. In column (2) the *Risky Asset Share* is regressed on *Direct Bank* and relevant control variables. For interpretability, we z-standardized *Household Income* and *Household Net Wealth* (i.e., subtracted their means and divided by their standard deviations) prior to including them in the model. All models include wave fixed effects to account for systematic differences across survey waves. Standard errors (in parentheses) are adjusted for imputation error, following Rubin's Rules, and for survey design, using replicate weights. The data are weighted using the sampling weights included in the sample. Significance levels are denoted as follows: $^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$.

Appendix II-E: Logistic Regression on Stock Market Participation

	Dependent Variable
	(1) Stock Market Participation
Direct Bank	0.4397***
	(0.1299)
Age	0.0055**
	(0.0024)
Male	0.3196***
	(0.0804)
Small-Town Resident	0.0766
	(0.1064)
Household Income (yearly)	0.2630***
	(0.0717)
Household Net Wealth	0.2985***
	(0.0973)
University Degree	0.8183***
	(0.0854)
Financial Literacy	0.5227***
	(0.0511)
Satisfaction with Life	0.1613***
	(0.0213)
Risk Tolerance (general)	0.0842***
	(0.0185)
Trust	0.0148
	(0.0190)
Impatience	0.0148
	(0.0144)
Constant	-4.8213***
	(0.2979)
Observations	14,624
Adj. Pseudo R^2	0.1280

Notes: This table reports the results of a logistic regression with $Stock\ Market\ Participation$ as the dependent variable. $Stock\ Market\ Participation$ is regressed on $Direct\ Bank$ and relevant control variables. For interpretability, we z-standardized $Household\ Income$ and $Household\ Net\ Wealth$ (i.e., subtracted their means and divided by their standard deviations) prior to including them in the model. All models include wave fixed effects to account for systematic differences across survey waves. We report the adjusted McFadden Pseudo R^2. Standard errors (in parentheses) are adjusted for imputation error, following Rubin's Rules, and for survey design, using replicate weights. The data are weighted using the sampling weights included in the sample. Significance levels are denoted as follows: $^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$.

Appendix II-F: Logistic Regression on Direct Bank for the Extended Sample

	Dependent Variable			
	(1) Direct Bank	(2) Direct Bank	(3) Direct Bank	(4) Direct Bank
Age	-0.0227***			-0.0205***
	(0.0033)			(0.0033)
Male	-0.0850			-0.1896
	(0.1188)			(0.1214)
Small-Town Resident	0.0494			0.0647
	(0.2263)			(0.2209)
Household Income (yearly)	0.0617*			0.0490
	(0.0367)			(0.0375)
Household Net Wealth	-0.1835**			-0.2297***
	(0.0738)			(0.0816)
University Degree	0.8776***			0.7493***
	(0.1097)			(0.1121)
Financial Literacy		0.4853***		0.3476***
		(0.0914)		(0.0901)
Satisfaction with Life			0.0582*	0.0387
			(0.0344)	(0.0380)
Risk Tolerance (general)			0.1030***	0.0952***
			(0.0248)	(0.0277)
Trust			0.0618*	0.0092
			(0.0341)	(0.0335)
Impatience			-0.0430**	-0.0467**
			(0.0211)	(0.0220)
Constant	-2.7096***	-4.8201***	-4.5812***	-4.1630***
	(0.2547)	(0.3168)	(0.3414)	(0.4658)
Observations	17,715	17,715	17,715	17,715
Adj. Pseudo R^2	0.0712	0.0385	0.0363	0.0824

Notes: This table reports our regression results for a logistic regression with $Direct\ Bank$ as the dependent variable. For this analysis $Direct\ Bank$ equals 1 if a household maintains any house bank relationship with a direct bank (even if it also has additional relationships) and 0 otherwise. In column (1) $Direct\ Bank$ is regressed on demographic variables. For interpretability, we z-standardized $Household\ Income$ and $Household\ Net\ Wealth$ (i.e., subtracted their means and divided by their standard deviations) prior to including them in the model. In column (2) $Direct\ Bank$ is regressed on the variable $Financial\ Literacy$, which indicates the score of the FKP in a three-question battery designed to test their financial literacy. In column (3) $Direct\ Bank$ is regressed on a set of variables describing the FKP attitudes. In column (4) all variables from the previous regressions are included in one model. All models include wave fixed effects to account for systematic differences across survey waves. We report the adjusted McFadden Pseudo R^2 for each specification. Standard errors (in parentheses) are adjusted for imputation error, following Rubin's Rules, and for survey design, using replicate weights. The data are weighted using the sampling weights included in the sample. Significance levels are denoted as follows: $^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$.

Appendix II-G: Regression on Investment Variables for the Extended Sample

	Dependent Variable	
	(1) Stock Market Participation	(2) Risky Asset Share
Direct Bank	0.0719***	0.0180
	(0.0221)	(0.0114)
Age	0.0008**	0.0015***
	(0.0003)	(0.0002)
Male	0.0493***	0.0155***
	(0.0120)	(0.0056)
Small-Town Resident	0.0130	-0.0092
	(0.0175)	(0.0070)
Household Income (yearly)	0.0445***	0.0059*
	(0.0083)	(0.0036)
Household Net Wealth	0.0488***	0.0243***
	(0.0103)	(0.0050)
University Degree	0.1705***	0.0648***
	(0.0161)	(0.0074)
Financial Literacy	0.0690***	0.0257***
	(0.0060)	(0.0027)
Satisfaction with Life	0.0260***	0.0095***
	(0.0028)	(0.0014)
Risk Tolerance (general)	0.0137***	0.0087***
	(0.0030)	(0.0014)
Trust	0.0022	0.0005
	(0.0027)	(0.0013)
Impatience	0.0026	0.0002
	(0.0022)	(0.0010)
Constant	-0.2872***	-0.1776***
	(0.0387)	(0.0210)
Observations	17,715	17,715
Adj. R^2	0.1384	0.0950
F-Statistic	83.6037***	37.7333***

Notes: This table reports our regression results for two OLS/LPM regressions. For this analysis, $Direct\ Bank$ equals 1 if a household maintains any house bank relationship with a direct bank (even if it also has additional relationships) and 0 otherwise. In column (1) $Stock\ Market\ Participation$ is regressed on $Direct\ Bank$ and relevant control variables. In column (2) the $Risky\ Asset\ Share$ is regressed on $Direct\ Bank$ and relevant control variables. For interpretability, we z-standardized $Household\ Income$ and $Household\ Net\ Wealth$ (i.e., subtracted their means and divided by their standard deviations) prior to including them in the model. All models include wave fixed effects to account for systematic differences across survey waves. Standard errors (in parentheses) are adjusted for imputation error, following Rubin's Rules, and for survey design, using replicate weights. The data are weighted using the sampling weights included in the sample. Significance levels are denoted as follows: $^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$.

III. Anlageverhalten und Kundenprofile im Vergleich: Unterschiede zwischen Sparkassen, Genossenschaftsbanken und Großbanken

Co-authors:

Andreas Walter

Own share:

60%

This article has been published as:

Becker, M. G., & Walter, A. (2024). Anlageverhalten und Kundenprofile im Vergleich: Unterschiede zwischen Sparkassen, Genossenschaftsbanken und Großbanken. *Zeitschrift für Bankrecht und Bankwirtschaft*, *36*(6), 382-392 https: doi.org/10.15375/zbb-2024-0606

Anlageverhalten und Kundenprofile im Vergleich: Unterschiede zwischen Sparkassen, Genossenschaftsbanken und Großbanken[a]

Martin G. Becker[b] Andreas Walter[c]

Abstract: Traditionell unterhalten Haushalte in Deutschland eine Hausbankbeziehung zu einer Filialbank, die einer der drei Gruppen der Universalbanken angehört: einer Sparkasse, einer Genossenschaftsbank oder einer Großbank. Diese Studie zielt darauf ab, erstmals die strukturellen Unterschiede in der Kundenbasis und im Anlageverhalten zwischen den drei Bankengruppen systematisch zu dokumentieren. Ein besonderes Augenmerk liegt auf der Frage, ob Kunden von Sparkassen und Genossenschaftsbanken im Vergleich zu Großbankkunden konservativer investieren und einen größeren Anteil ihres Finanzvermögens in Bankeinlagen halten. Darüber hinaus wird untersucht, ob die Zugehörigkeit zu einer bestimmten Bankengruppe einen signifikanten Einfluss auf die Aktienmarktpartizipation und das generelle Anlageverhalten besitzt oder ob diese Abweichungen durch Unterschiede in der Kundenstruktur zwischen den Instituten erklärt werden können.

[a] Wir danken der Stiftung für die Wissenschaft für die finanzielle Förderung dieses Projekts.
[b] martin.becker@wirtschaft.uni-giessen.de, Justus-Liebig-Universität Gießen, Professur für Finanzdienstleistungen, Licher Straße 74, 35394 Gießen, Germany
[c] andreas.walter@wirtschaft.uni-giessen.de, Justus-Liebig-Universität Gießen, Professur für Finanzdienstleistungen, Licher Straße 74, 35394 Gießen, Germany

1. Einführung

Der Markt für Retail Banking wird in Deutschland traditionell von drei Bankengruppen dominiert, die ihre Finanzdienstleistungen über ein breites Netz an Filialen anbieten: Sparkassen, Genossenschaftsbanken und Großbanken. Dabei positionieren sich die einzelnen Bankengruppen unterschiedlich gegenüber ihren Kunden und der Öffentlichkeit. Die Deutsche Bank richtet ihre Kampagnen beispielsweise auf ein urbanes und vermögendes Kundenprofil aus, während Sparkassen und Volksbanken bodenständigere Botschaften vermitteln, etwa durch die Unterstützung regionaler Sportvereine. Daher hat sich in der allgemeinen Wahrnehmung verfestigt, dass Großbanken eine andere Kundengruppe bedienen als Sparkassen und Volksbanken. Interessanterweise existieren bislang keine empirischen Studien, die Unterschiede in der Kundenstruktur und im Anlageverhalten der Kunden verschiedener Bankengruppen untersuchen. Es fehlen bislang Antworten auf Fragen wie: Sind Großbankkunden vermögender als Sparkassenkunden? Unterscheiden sich die Kunden einer Großbank systematisch von denen einer Genossenschaftsbank in Bezug auf ihre Risikoaversion? Welche Bankengruppe ist bei Kunden mit Migrationshintergrund besonders beliebt? Verfügen Kunden von Sparkassen über eine geringere Finanzkompetenz als die Kunden von Großbanken? Sind Kunden von Volksbanken besonders selten im Aktienmarkt investiert? Lässt sich eine etwaig niedrige Aktienmarktpartizipation von Sparkassenkunden dadurch erklären, dass ein bestimmter Kundentyp sich bevorzugt für eine Sparkasse entscheidet? Oder ist die Tatsache, Kunde bei einer Sparkasse zu sein – etwa durch eine besonders defensive Anlageberatung bei Sparkassen – entscheidend für die Scheu vor dem Aktienmarkt?

Da bisher keine Antworten auf diese und andere Fragen zu Kundenprofilen und dem Anlageverhalten vorliegen, möchten wir diese Lücke schließen. Konkret nutzen wir das Panel of Household Finances (PHF) der Deutschen Bundesbank aus dem Jahr 2021, um erste Antworten auf die beschriebenen Fragestellungen zu geben. Für knapp 3.000 repräsentativ ausgewählte Haushalte in Deutschland führen wir detaillierte Analysen durch, um einerseits zu untersuchen, inwieweit sich Kunden der Sparkassen, Genossenschaftsbanken und Großbanken hinsichtlich ihrer demographischen Merkmale und ihrer Risikobereitschaft unterscheiden. Andererseits analysieren wir das Anlageverhalten der Haushalte nach Bankengruppe, wobei das Ausmaß der Investitionen in Bankeinlagen und die Partizipation am Aktienmarkt im Vordergrund der Analysen stehen.

Der Rest der Arbeit ist wie folgt aufgebaut. In Abschnitt 2 werden zunächst die institutionellen Hintergründe des deutschen Bankenmarktes knapp erläutert und die zentralen Forschungsfragen der Studie abgeleitet. Abschnitt 3 beschreibt die verwendeten Daten und Variablen. Die Ergebnisse der empirischen Analysen werden in Abschnitt 4 präsentiert. Zunächst analysieren wir dort die Unterschiede in der Kundenstruktur zwischen Sparkassen-, Genossenschaftsbanken- und Großbankkunden. Anschließend untersuchen wir das Anlageverhalten der Kunden in den drei Bankengruppen, insbesondere die Präferenzen für risikoarme Bankeinlagen und die Aktienmarktpartizipation. Schließlich führen wir multivariate Analysen durch, um zu prüfen, inwieweit die Bankzugehörigkeit das Anlageverhalten zu beeinflussen scheint, auch nach der Kontrolle für relevante Kundenmerkmale. Der Beitrag schließt mit einem Fazit, in dem die wesentlichen Ergebnisse der Studie zusammengefasst werden.

2. Institutionelle Rahmenbedingungen und Forschungsfragen

2.1 Institutionelle Rahmenbedingungen

Das deutsche Bankensystem basiert auf dem sog. Drei-Säulen-Modell, das sich in private Geschäftsbanken (Kreditbanken), öffentlich-rechtliche Banken (Sparkassen und Girozentralen) und Genossenschaftsbanken (Genossenschaftsbanken und genossenschaftliche Zentralbanken) unterteilt. Jede dieser Säulen erfüllt spezifische Aufgaben im Bankwesen und bietet unterschiedliche Dienstleistungen für Privat- und Geschäftskunden an.[1] Die sog. Großbanken wie die Deutsche Bank und die Commerzbank, dominieren den Markt der privaten Geschäftsbanken.[2] Diese Großbanken sind international vernetzt und bieten ihre Finanzdienstleistungen im gesamten Bundesgebiet an. Die Sparkassen-Finanzgruppe bildet die öffentlich-rechtliche Säule und konzentriert sich auf die Versorgung der Bevölkerung mit grundlegenden Bankdienstleistungen in regional abgegrenzten Märkten. Die Genossenschaftsbanken, darunter Volks- und Raiffeisenbanken, bilden die dritte Säule und verfolgen einen genossenschaftlichen Ansatz, bei dem die Mitglieder gleichzeitig Eigentümer und Kunden sind. Auch diese Genossenschaftsbanken agieren in der Regel in einem räumlichen begrenzten Geschäftsbereich. Die Kreditinstitute der drei Säulen unterscheiden sich hinsichtlich ihrer Geschäftsausrichtung und Kundenstruktur erheblich. Großbanken haben einen stärkeren Fokus auf das Effektengeschäft und sind stärker im Auslandsgeschäft tätig,[3] während Sparkassen und Genossenschaftsbanken stärker auf regionale Märkte und die Betreuung von Privatkunden sowie kleinen und mittelständischen Unternehmen fokussiert sind.[4] Für unsere Studie sind insbesondere die Unterschiede in der Struktur der Privatkunden relevant, da wir erwarten, dass sich das Anlageverhalten und die Risikobereitschaft der Kunden je nach Bankengruppe deutlich unterscheiden.

Im Weiteren werden wir Institute der öffentlich-rechtlichen Säule als Sparkassen und Volks- und Raiffeisenbanken als Genossenschaftsbanken bezeichnen. Bei den privaten Banken fokussieren wir uns auf die Großbanken, konkret die Deutsche Bank, die Commerzbank, die UniCredit (vormals HypoVereinsbank) und die Postbank.

---

[1] Vgl. Priewasser (2001, S. 144 ff.) und Hartmann-Wendels et al. (2019, S. 29 ff.)
[2] Großbanken halten einen erheblichen Marktanteil an der gesamten Bankenbranche in Deutschland, vgl. Deutsche Bundesbank (2024)
[3] Vgl. Hartmann-Wendels et al. (2019, S. 31 ff.)
[4] Vgl. Hartmann-Wendels et al. (2019, S. 35 ff.)

2.2 Forschungsfragen

Bisher gibt es nach unserem Kenntnisstand keine empirische Studie, die die Charakteristika und das Anlageverhalten der Kunden von Sparkassen, Genossenschaftsbanken und Großbanken systematisch miteinander vergleicht. Unsere Forschungsfragen sind daher explorativ angelegt und basieren nicht auf bestehenden literaturgestützten Hypothesen. Ziel unserer Untersuchung ist es, erstmals empirische Evidenz für die Unterschiede zwischen den Kunden der verschiedenen Bankengruppen in Deutschland zu liefern und damit eine Forschungslücke zu schließen.

Zunächst untersuchen wir, wie sich die Kundenstruktur zwischen Sparkassen, Genossenschaftsbanken und Großbanken unterscheidet. Insbesondere wollen wir klären, inwieweit demografische Merkmale wie das Alter, die Finanzkompetenz, ein Migrationshintergrund oder der Bildungsstand der Kunden zwischen den Bankengruppen variieren. Die Frage, ob sich die Risikobereitschaft und andere finanzielle Charakteristika wie das Haushaltsvermögen und -einkommen in den drei Bankengruppen signifikant unterscheiden, steht dabei zusätzlich im Mittelpunkt unserer Untersuchungen.

Wir wollen ferner in unserer Studie die Frage adressieren, inwieweit sich das Anlageverhalten der Kunden zwischen den drei Bankengruppen unterscheidet. Wir fokussieren uns dabei zum einen auf die Aktienmarktpartizipation der Bankkunden und zum anderen auf den Anteil des Finanzvermögens, der in Bankeinlagen investiert wird. Ziel ist es, herauszufinden, ob Kunden von Sparkassen und Genossenschaftsbanken im Vergleich zu Großbankkunden konservativere Anlageentscheidungen treffen.

Schließlich wollen wir untersuchen, ob die beobachteten Unterschiede im Anlageverhalten primär durch die Bankzugehörigkeit oder durch individuelle Kundenmerkmale erklärt werden können. Dazu führen wir multivariate Analysen durch, um Unterschiede in der Finanzkompetenz, des Einkommens, des Vermögens und der Risikobereitschaft der Bankkunden zu kontrollieren. Wir adressieren dabei die Frage, ob die Wahl der Hausbank weiterhin einen signifikanten Einfluss auf das Anlageverhalten zu haben scheint, nachdem wir für Merkmale der Bankkunden kontrolliert haben.

3. Daten und deskriptive Statistik

3.1 Datensatz

Die Datenbasis für unsere Analyse bildet das „Panel on Household Finances“ (PHF) der Deutschen Bundesbank, welches seit 2010 in mehreren Wellen repräsentative Daten zu den Vermögensverhältnissen, Schulden und dem Finanzverhalten privater Haushalte in Deutschland erfasst.[5] Für die vorliegende Studie verwenden wir die Daten der vierten Welle, die 2021 erhoben wurde und detaillierte Informationen zu 4.119 Haushalten enthält. Die vierte Welle wurde pandemiebedingt überwiegend telefonisch durchgeführt, was jedoch die Vergleichbarkeit mit früheren Erhebungen nicht beeinträchtigt.[6]

In unserer Analyse konzentrieren wir uns auf die 2.926 Haushalte, die eine Beziehung zu einer Hausbank bei einer Sparkasse, einer Genossenschaftsbank oder einer Großbank unterhalten, was etwa 71 % der insgesamt befragten Haushalte entspricht. Haushalte ohne Hausbankbeziehung zu einer Bank der drei genannten Bankengruppen, wie bspw. Kunden von Direktbanken oder Kunden ohne Hausbankbeziehung, bleiben in unserer Untersuchung unberücksichtigt.[7] Die PHF-Daten bieten eine hohe Detailtiefe, bei der bestimmte Informationen, wie das Nettovermögen, auf Haushaltsebene erfasst werden, während andere Variablen, wie etwa die Finanzkompetenz, auf Personenebene erhoben werden.[8] Hierbei wird der "Kompetenzträger Haushaltsfinanzen" (KT) als primär befragte Person des Haushalts

[5] Für Details zum PHF Datensatz siehe Deutsche Bundesbank (2023c), https://www.bundesbank.de/de/bundesbank/forschung/fdsz/forschungsdaten/phf-856346.
[6] Vgl. Deutsche Bundesbank (2023a, S. 26)
[7] Darüber hinweg wurden auch alle Haushalte aus dem Sample entfernt, die angegeben haben, Hausbankbeziehungen mit mehreren Banken zu haben, da diese nicht eindeutig einer einzelnen Bankengruppe zugeordnet werden können.
[8] Der PHF-Befragungsdatensatz verwendet ein stratifiziertes mehrstufiges Sample-Design um ein zufälliges, repräsentatives deutschlandweites Sample zu generieren. Hierbei wird unter anderem gezielt ein Oversampling wohlhabender Haushalte vorgenommen. Dies soll dazu dienen, genauere Angaben zu Vermögensvariablen machen zu können. Für die Auswirkungen des Oversamplings wird in unseren empirischen Analysen korrigiert. Siehe auch Deutsche Bundesbank (2023b, S. 23)

ausgewählt. Diese Person ist in der Regel für die finanziellen Entscheidungen des Haushalts verantwortlich.

3.2 Deskriptive Statistik

Tabelle III-1 berichtet deskriptive Statistiken zur Verteilung der in dieser Studie analysierten Variablen.[9] Panel A zeigt, dass knapp die Hälfte der Hausbankkunden ihre Bankbeziehung bei einer Sparkasse unterhält (49,0 %). Die zweitgrößte Gruppe bilden Genossenschaftsbankkunden mit 30,8 %, gefolgt von Großbankkunden mit 20,3 %. Diese Zahlen korrespondieren eng mit den Marktanteilen bei den Bankeinlagen von Privatpersonen, die vom DSGV für das Jahr 2023 veröffentlicht wurden.[10]

Panel B enthält Informationen zur Verteilung der Kundenmerkmale. Besonders hervorzuheben ist die Variable „Prohibitive Risikoaversion“, die den Wert 1 annimmt, wenn keinerlei Bereitschaft besteht, Risiken einzugehen, um eine höhere Rendite zu erzielen. Bei 61,2 % der Bankkunden trifft dies zu, was einen deutlichen Hinweis auf die ausgeprägte Risikoaversion deutscher Haushalte gibt.[11] Panel C zeigt die Aufteilung der Bankkunden nach Größe der Gemeinde, in der diese Bankkunden leben. Es wird deutlich, dass mehr Kunden in Städten und Gemeinden mit weniger als 20.000 Einwohnern leben (39,4 %) als in Städten mit über 100.000 Einwohnern (32,5 %).

Informationen zum Anlageverhalten finden sich in Panel D, wobei hier die Daten den gesamten Haushalt und nicht nur den "Kompetenzträger Haushaltsfinanzen" (KT) betreffen. Dies erklärt die relativ hohe Aktienmarktpartizipationsrate von 25,5 %.[12] Jedoch bleibt das investierte Volumen am Aktienmarkt gering: Im 75.-Perzentil beträgt das Investitionsvolumen

[9] Eine detaillierte Beschreibung der Variablen findet sich in Appendix A.
[10] Vgl. Deutscher Sparkassen- und Giroverband (2024, S. 43)
[11] Ein besonders starkes Ausmaß der Risikoaversion deutscher Haushalte im internationalen Vergleich dokumentieren auch Rieger et al. (2015) und von Lüde (2013).
[12] Das Deutsche Aktieninstitut (DAI) gibt für das Jahr 2021 eine deutlich niedrigere Partizipationsrate von 18,7% für Einzelpersonen in Deutschland an. Diese Zahl ergibt sich aus 12,1 Millionen Aktiensparen in Deutschland dividiert durch 64,8 Mio. erwachsenen Einwohnern (> 19 Jahren) in Deutschland im Jahr 2021. Vgl. Deutsches Aktieninstitut (2022) und Statistisches Bundesamt (2023)

im Aktienmarkt lediglich 500 €. Die Variablen zum Anlageverhalten, wie das Finanzvermögen bei Banken, weisen eine rechtsschiefe Verteilung auf. Während der Median-Haushalt ein Vermögen bei Banken von 12.500 € hält, liegt der Mittelwert deutlich höher bei 50.249 €. Das hohe Maß an Risikoaversion, das bereits bei der Variablen „Prohibitive Risikoaversion" in Panel B ersichtlich wurde, zeigt sich auch im Anlageverhalten: Der Mittelwert des Anteils der Bankeinlagen am Finanzvermögen liegt bei sehr hohen 87,8 %. Dieser Wert ist auch darauf zurückzuführen, dass der Median-Bankkunde ausschließlich Einlagen bei der Bank hält. Für eine Variablenbeschreibung siehe Appendix A.

Tabelle III-1: Deskriptive Statistiken

Variable	(1) Mittelwert	(2) 25%-Quantil	(3) Median	(4) 75%-Quantil	(5) N
Panel A: Verteilung auf Bankengruppe					
Sparkasse	48,98%	0,00	0,00	1,00	2.926
Genossenschaftsbank	30,75%	0,00	0,00	1,00	2.926
Großbank	20,27%	0,00	0,00	0,00	2.926
Panel B: Kundencharakteristika					
Alter	56,17	42,00	56,00	70,00	2.926
Männlich	0,52	0,00	1,00	1,00	2.926
Migrationshintergrund	0,15	0,00	0,00	0,00	2.926
Haushaltsgröße	1,91	1,00	2,00	2,00	2.926
Haushaltseinkommen (in tsd €)	57,55	26,50	42,50	72,40	2.926
Nettovermögen des Haushalts (in tsd €)	320,29	14,00	113,00	361,00	2.926
Hochschulabschluss	0,21	0,00	0,00	0,00	2.926
Ostdeutschland	0,22	0,00	0,00	0,00	2.926
Finanzkompetenz (Financial Literacy)	2,30	2,00	2,00	3,00	2.926
Prohibitive Risikoaversion	0,61	0,00	1,00	1,00	1.928
Panel C: Klassifizierung nach Gemeindegröße					
[0; 5,000] Einwohner	13,80%	0,00	0,00	0,00	2.893
[5,001; 20,000] Einwohner	25,63%	0,00	0,00	1,00	2.893
[20,001; 100,000] Einwohner	28,11%	0,00	0,00	1,00	2.893
[100,001; 500,000] Einwohner	15,32%	0,00	0,00	0,00	2.893
[> 500,000] Einwohner	17,13%	0,00	0,00	0,00	2.893
Panel D: Anlageverhalten					
Aktienmarktpartizipation	25,53%	0,00	0,00	1,00	2.926
Finanzvermögen bei Banken (in €)	50.248,70	3.000,00	12.500,00	48.000,00	2.926
Bankeinlagen (in €)	30.558,80	2.500,00	10.000,00	31.000,00	2.926
Investition am Aktienmarkt (in €)	17.664,96	0,00	0,00	500,00	2.926
Anteil Einlagen am Finanzvermögen	87,76%	96%	100%	100%	2.926
Anteil Aktien am Finanzvermögen	11,59%	0,00%	0,00%	1,57%	2.926

Anmerkung: Diese Tabelle enthält die deskriptiven Statistiken für unsere Stichprobe. Die Definitionen für die einzelnen Variablen sind in Appendix A zu finden. Der Mittelwert und die Quantile wurden unter Verwendung der in der Stichprobe enthaltenen Sampling-Gewichte berechnet.

4. Empirische Ergebnisse

4.1 Unterschiede in der Kundenstruktur der deutschen Banken

Tabelle III-2 vergleicht die Kundenmerkmale der drei Bankengruppen: Sparkassen, Genossenschaftsbanken und Großbanken. In den Spalten 1 bis 3 sind die Mittelwerte der Variablen für jede Gruppe dargestellt, während in den Spalten 4 bis 6 die Differenzen zwischen den Gruppen aufgeführt sind. Statistisch signifikante Unterschiede werden durch Sternchen markiert, wobei ***, **, und * für signifikante Unterschiede der Gruppen auf dem 1%-, 5%- und 10%-Niveau stehen.

Besonders augenscheinlich sind die Unterschiede in der Risikobereitschaft der Bankkunden. Bei Sparkassen sind gemäß der Variable „Prohibitive Risikoaversion" 67,1 % der Kunden nicht bereit, Risiken einzugehen, um höhere Renditen zu erzielen. Die Anteile bei Genossenschaftsbanken (54,6 %) und Großbanken (58,1 %) sind ebenfalls hoch, aber signifikant niedriger. Sparkassenkunden scheuen demnach das Risiko am deutlichsten. Zudem gibt es Unterschiede im Haushaltseinkommen und im Nettovermögen der Kunden. Großbankkunden haben das höchste mittlere Haushaltseinkommen von 64,9 Tausend €. Das mittlere Haushaltseinkommen der Sparkassenkunden liegt signifikant darunter, bei 52,9 Tausend €. Beim Nettohaushaltsvermögen ergibt sich hingegen ein anderes Bild: Genossenschaftsbankkunden haben hier mit 428,2 Tausend € das höchste mittlere Nettohaushaltsvermögen, deutlich mehr als Kunden von Großbanken (312,6 Tausend €) und Kunden von Sparkassen (255,7 Tausend €).[13]

Eine weitere Auffälligkeit zeigt sich beim Anteil der Kunden mit Migrationshintergrund. Großbanken haben hier mit 23,5 % einen signifikant höheren Anteil als Genossenschaftsbanken, deren Anteil von Kunden mit Migrationshintergrund bei lediglich 9,1 % liegt. Sparkassen nehmen mit einem Anteil von 16,0 % eine Mittelposition ein. Auch beim Bildungsstand gibt es Unterschiede zwischen den Bankengruppen: 26,8 % der Großbankkunden besitzen einen Hochschulabschluss, während dieser Anteil bei Genossenschaftsbanken nur 19,0 % beträgt. Deutliche Unterschiede finden sich auch bei der geografischen Verteilung der Kunden. So leben nur 10,6 % der Genossenschaftsbankkunden in

[13] Der Median des Nettovermögens von Genossenschaftsbankkunden liegt bei 176 Tausend €, etwa doppelt so hoch wie der Median der Sparkassenkunden mit 81 Tausend €.

den neuen Bundesländern, während dieser Anteil bei Großbanken mit 29,1 % und bei Sparkassen mit 25,4 % signifikant höher ist.[14]

Es ist auch aufschlussreich, die Kundenmerkmale zu betrachten, bei denen keine Unterschiede zwischen den Bankengruppen bestehen. So gibt es beispielsweise keine signifikanten Unterschiede in der Alters- und Geschlechterstruktur sowie in der durchschnittlichen Haushaltsgröße. Ebenfalls zeigen sich bei der Finanzkompetenz, gemessen durch die „Big Three" Finanzkompetenzfragen,[15] keine signifikanten Unterschiede zwischen den Kunden der Bankengruppen. Obwohl Sparkassenkunden im Durchschnitt etwas weniger Fragen zur Finanzkompetenz richtig beantworten (Mittelwert 2,27) als Genossenschaftsbankkunden (Mittelwert 2,33) und Großbankkunden (Mittelwert 2,32), ist dieser Unterschied statistisch nicht signifikant.

Panel B zeigt eine Klassifikation in fünf Gemeindegrößen-Cluster. Dabei wird deutlich, dass Genossenschaftsbanken den Großteil ihrer Kunden in sehr kleinen Gemeinden haben, während die Kunden der Großbanken überwiegend in größeren Gemeinden leben. So wohnt etwa jeder fünfte Genossenschaftsbankkunde in einer Gemeinde mit weniger als 5.000 Einwohnern, während bei Großbanken nur etwa jeder zwanzigste Kunde in solch einer kleinen Gemeinden ansässig ist. Im Gegensatz dazu lebt etwa jeder vierte Großbankkunde in einer Stadt mit mehr als 500.000 Einwohnern, wohingegen dies bei Genossenschaftsbankkunden nur auf jeden zwölften zutrifft. Die geographische Verteilung der Sparkassenkunden ist ausgeglichener, wobei ihr Profil dem der Genossenschaftsbanken stärker ähnelt als dem der Großbanken.

[14] Genossenschaftsbanken waren nach der Wende in Ostdeutschland schwach vertreten, da ihre geringe Zahl und eingeschränkte Finanzierungsmöglichkeiten sie strukturell benachteiligten, Vgl. Hahn (1996) und Schildbach (2019)

[15] Die Fragen wurden in Lusardi und Mitchell (2008) eingeführt und sind international für die Messung der Finanzkompetenz sehr stark verbreitet. Ein Selbsttest kann man bspw. unter https://gflec.org/education/big-three/ durchführen. In der vierten Welle des PHF-Datensatzes haben nicht mehr alle Teilnehmer die gleichen drei Fragen erhalten. Stattdessen wurde für eine zufällig ausgewählte Hälfte die klassische Frage zu Diversifikation durch eine andere Frage zur Diversifikation ersetzt. Wenn wir unser Sample nur auf die Hälfte beschränken, welche die Original „Big Three" Fragen erhalten hat, steigt der Mittelwert bei allen drei Bankgruppen (Sparkasse: 2,42; Genossenschaftsbanken: 2,50; Großbanken: 2,53). Die Unterschiede zwischen den Bankengruppen bleiben jedoch weiterhin insignifikant.

Tabelle III-2: Charakteristika nach Bankengruppe

Variable	(1) Mittelwert Sparkasse	(2) Mittelwert Genossenschaftsbank	(3) Mittelwert Großbank	(4) Differenz Sparkasse-Genossenschaftsbank	(5) Differenz Sparkasse-Großbank	(6) Differenz Genossenschaftsbank -Großbank
Panel A: Kundencharakteristika						
Alter	55,69	56,56	56,72	-0,87	-1,02	-0,16
Männlich	0,51	0,51	0,57	0,00	-0,06	-0,06
Migrationshintergrund	16,02%	9,12%	23,52%	6,90%***	-7,50%**	-14,40%***
Haushaltsgröße	1,87	1,92	1,98	-0,04	-0,11	-0,06
Haushaltseinkommen (in tsd €)	52,88	60,13	64,93	-7,25**	-12,05***	-4,80
Nettovermögen des Haushalts (in tds €)	255,69	428,22	312,62	-172,53***	-56,94	115,59**
Hochschulabschluss	18,98%	21,64%	26,76%	-2,66%	-7,78%**	-5,12%
Ostdeutschland	25,42%	10,65%	29,13%	14,77%***	-3,71%	-18,48%***
Finanzkompetenz (Financial Literacy)	2,27	2,33	2,34	-0,06	-0,07	-0,01
Prohibitive Risikoaversion	67,10%	54,57%	58,09%	12,53%***	9,01%*	-3,52%
Panel B: Klassifizierung nach Gemeindegröße						
[0; 5.000] Einwohner	12,55%	21,13%	5,65%	-8,58%***	6,90%***	15,48%***
[5.001; 20.000] Einwohner	26,39%	29,37%	18,10%	-2,98%	8,30%**	11,28%***
[20.001; 100.000] Einwohner	28,83%	27,81%	26,84%	1,02%	1,99%	0,98%
[100.001; 500.000] Einwohner	13,96%	13,85%	20,85%	0,12%	-6,88%**	-7,00%**
[> 500.000] Einwohner	18,26%	7,83%	28,56%	10,42%***	-10,31%***	-20,73%***

Anmerkung: Die Spalten 1 – 3 dieser Tabelle enthalten die Mittelwerte der demografischen Merkmale nach Bankengruppen (Panel A). Die Spalten 4 – 6 zeigen die Differenzen zwischen den Mittelwerten der einzelnen Bankengruppen. Panel B gibt die Klassifizierung nach Gemeindegröße an, in der die Bankkunden wohnhaft sind. Bspw. steht der Wert von 12,55 % in Spalte (1) dafür, dass 12,55 % der Sparkassenkunden in einer Gemeinde wohnen, die unter 5.000 Einwohner besitzt. *,**,*** gibt an, ob der Unterschied zwischen den Bankengruppen auf dem 10 %, 5 % und 1 % Niveau statistisch signifikant ist. Die P-Werte hierfür werden auf Basis eines zweiseitigen t-Tests bzw. bei binären Variablen auf Basis eines zweiseitigen Wilcoxon-Mann-Whitney-Tests berechnet. Die Mittelwerte sowie die P-Werte werden unter Berücksichtigung der Sampling- Gewichte berechnet.

4.2 Unterschiede im Anlageverhalten zwischen Bankengruppen

Im Folgenden analysieren wir, inwieweit sich die Kunden der Bankengruppen in ihrem Anlageverhalten unterscheiden. Tabelle III-3 vergleicht dazu Variablen des Anlegerverhaltens über die drei Bankengruppen hinweg. In den Spalten 1 bis 3 sind wiederum die Mittelwerte der Variablen für jede Gruppe dargestellt, während die Spalten 4 bis 6 die Differenzen zwischen den Gruppen angeben. Statistisch signifikante Unterschiede werden erneut durch Sternchen gekennzeichnet, wobei ***, ** und * Signifikanzniveaus von 1 %, 5 % und 10 % angeben.

Bezüglich der Aktienmarktpartizipation zeigen sich vor allem signifikante Unterschiede zwischen den Kunden der Sparkassen und den Kunden der Großbanken. So haben 30,3 % der Großbankkunden entweder in Aktien oder Fonds investiert und partizipieren damit am Aktienmarkt. Bei Sparkassenkunden liegt dieser Wert signifikant niedriger bei 23,1 %, während er bei Genossenschaftsbanken mit 26,2 % ebenfalls deutlich, aber nicht signifikant geringer als bei Großbanken ausfällt. In Abschnitt 4.3. werden wir untersuchen, ob dieser Unterschied zwischen den Bankengruppen auch dann bestehen bleibt, wenn wir für Kundenmerkmale kontrollieren.

Auch beim Finanzvermögen zeigen sich deutliche Unterschiede. Sparkassenkunden haben im Durchschnitt 42,4 Tausend € bei ihrem Kreditinstitut investiert, deutlich weniger als Genossenschaftsbankkunden (58,4 Tausend €) oder Großbankkunden (56,9 Tausend €). Betrachtet man die relative Aufteilung des Finanzvermögens auf Bankeinlagen einerseits und Kapitalmarktinvestitionen andererseits, werden ebenfalls signifikante Unterschiede zwischen den Bankengruppen deutlich. Sparkassenkunden halten mit 89,6 % einen deutlich höheren Anteil ihres Finanzvermögens in Bankeinlagen als Großbankkunden (84,6 %). Auch Genossenschaftsbankkunden halten mit 86,9 % einen geringeren Anteil ihres Finanzvermögens in Bankeinlagen als Sparkassenkunden. Ein ähnliches Muster zeigt sich beim Anteil des Finanzvermögens, das in Aktien oder Fonds investiert ist: Bei Großbankkunden liegt dieser Anteil bei 14,8 %, deutlich höher als bei Sparkassenkunden mit 9,9 %. Ob diese Unterschiede unter Berücksichtigung der Kundenmerkmale bestehen bleiben, untersuchen wir ebenfalls im nachfolgenden Abschnitt 4.3.

Tabelle III-3: Anlage nach Bankengruppe

Variable	(1) Mittelwert Sparkasse	(2) Mittelwert Genossenschaftsbank	(3) Mittelwert Großbank	(4) Differenz Sparkasse-Genossenschaftsbank	(5) Differenz Sparkasse-Großbank	(6) Differenz Genossenschaftsbank -Großbank
Aktienmarktpartizipation	23,14%	26,21%	30,30%	-3,07%	-7,16%**	-4,09%
Finanzvermögen bei Banken (in €)	42.383,82	58.365,94	56.939,33	-15.982,12	-14.555,50**	1.426,61
Bankeinlagen (in €)	27.704,86	33.000,99	33.750,40	-5.296,13	-6.045,54	-749,41
Investition am Aktienmarkt (in €)	13.738,48	21.339,82	21.578,19	-7.601,34	-7.839,72**	-238,37
Anteil Einlagen am Finanzvermögen	89,63%	86,89%	84,58%	2,73%*	5,05%***	2,32%
Anteil Aktien am Finanzvermögen	9,95%	12,09%	14,81%	-2,14%	-4,86%**	-2,72%

Anmerkung: Die Spalten 1 – 3 dieser Tabelle enthalten die Mittelwerte der Investmentmerkmale nach Bankengruppen. Die Spalten 4 – 6 zeigen die Differenzen zwischen den Mittelwerten der einzelnen Bankengruppen. *,**,*** gibt an, ob der Unterschied zwischen den Bankengruppen auf dem 10 %, 5 % und 1 % Niveau statistisch signifikant ist. Die P-Werte hierfür werden auf Basis eines zweiseitigen t-Tests bzw. bei binären Variablen auf Basis eines zweiseitigen Wilcoxon-Mann-Whitney-Tests berechnet. Die Mittelwerte sowie die P-Werte werden unter Berücksichtigung der Sampling-Gewichte berechnet.

4.3 Multivariate Analysen zum Anlageverhalten

In diesem Abschnitt gehen wir der Frage nach, ob die Unterschiede im Anlegerverhalten durch die Finanzberatung der Bankengruppen mitbeeinflusst werden oder ob diese Unterschiede allein durch individuelle Kundenmerkmale erklärt werden können. Konkret untersuchen wir dabei (i) die Aktienmarktpartizipation, (ii) den Anteil der Bankeinlagen am Finanzvermögen und (iii) den Anteil der Aktien und Fonds am Finanzvermögen als zu erklärende Variablen. In den Tabellen III-4 bis 6 schätzen wir jeweils drei Modelle. Im ersten Modell werden lediglich Dummy-Variablen für die Bankengruppen Sparkasse und Genossenschaftsbank inkludiert, um die Ergebnisse der Gruppenvergleiche aus Abschnitt 4.2. zu replizieren. Im zweiten Modell ergänzen wir eine Reihe von Kontrollvariablen, um Kundenmerkmale zu berücksichtigen. Im dritten Modell wird schließlich die Variable „Prohibitive Risikoaversion" hinzugefügt.[16]

[16] Zu beachten ist, dass die Datenabdeckung für die Variable „Prohibitive Risikoaversion" geringer ist und wir für das Modell in Spalte (3) lediglich 1.928 anstatt 2.926 Beobachtungen verwenden können.

Aktienmarktpartizipation

Tabelle III-4 zeigt die Ergebnisse des linearen Regressionsmodells (LPM) zur Erklärung der Aktienmarktpartizipation.[17] Zunächst bestätigen die Ergebnisse aus Modell (1), dass Sparkassenkunden ohne Berücksichtigung der Kundenmerkmale seltener in Aktien und Fonds investieren als Großbankkunden (Effektstärke -7,16%)[18]. Berücksichtigt man jedoch alle Kundencharakteristika in Modell (3), verringert sich dieser Effekt (Effektstäke -3,44%) und verliert seine statistische Signifikanz. Dies deutet darauf hin, dass die geringe Aktienmarktpartizipation von Sparkassenkunden nicht signifikant durch die Beratung in diesen Instituten verursacht wird, sondern vielmehr maßgeblich auf spezifische Eigenschaften der Sparkassenkunden zurückzuführen ist, die sie weniger häufig zu Aktiensparern machen.

Bezüglich der Kontrollvariablen finden wir einige interessante Ergebnisse. In Modell (2) zeigt sich ein deutlicher negativer Effekt eines Migrationshintergrunds: Kunden mit Migrationshintergrund haben eine um 10,5 % geringere Partizipationsrate am Aktienmarkt, und dies unter Berücksichtigung diverser anderer relevanter Kundenmerkmale. Gemessen an einer durchschnittlichen Aktienmarktpartizipation von 25,5 % in der gesamten Stichprobe, ist dieser Effekt auch ökonomisch hoch signifikant. Ein ähnlich starker Zusammenhang besteht auch bei Kunden aus den neuen Bundesländern, die eine um 10,2 % geringere Partizipationsrate als Westdeutsche aufweisen, ebenfalls nach Kontrolle diverser anderen Kundenmerkmale.[19] In Modell (3) zeigt sich zudem der starke Effekt der Variablen „Prohibitive Risikoaversion“. Nicht nur ist der Koeffizient hoch signifikant, sondern die Erklärungskraft des Modells nimmt durch die Hinzunahme dieser Variable ebenfalls deutlich zu. Zudem bestätigen wir die in der Literatur bekannten Befunde, dass Männer[20], Personen mit höherer Finanzkompetenz[21] und höherem

[17] Da die Aktienmarktpartizipation über 20% liegt, verwenden wir ein lineares Regressionsmodell anstatt logistische Verfahren. Siehe zur Begründung auch Stolper und Walter (2019).

[18] Aufgrund der statistisch insignifikanten F-Statistik bei Modell (1) sind die Ergebnisse dieses Modells mit Vorsicht zu interpretieren.

[19]Vgl. Laudenbach et al. (2024)

[20] Vgl. Almenberg und Dreber (2015); Fey et al. (2020) und van Rooij et al. (2011)

[21] Vgl. Bucher-Koenen et al. (2023) und van Rooij et al. (2011)

Bildungsabschluss[22] häufiger am Aktienmarkt investieren. Auch der positive Zusammenhang zwischen Einkommen[23] und Vermögen[24] mit der Aktienmarktpartizipation wird durch unsere Analysen bestätigt.

[22] Vgl. Guiso et al. (2003); Bertaut (1998); Cole et al. (2014) und Thomas und Spataro (2018)
[23] Vgl. Haliassos und Bertaut (1995) und Kaustia et al. (2023)
[24] Vgl. Guiso et al. (2003); Fagereng et al. (2017); Oehler et al. (2024) und Campbell (2006)

Tabelle III-4: Aktienmarktpartizipation

	(1) Aktienmarktpartizipation	(2) Aktienmarktpartizipation	(3) Aktienmarktpartizipation
Sparkasse	-0,0716**	-0,0494	-0,0344
	(0,0338)	(0,0314)	(0,0398)
Genossenschaftsbank	-0,0409	-0,0673*	-0,0619
	(0,0368)	(0,0351)	(0,0440)
Alter		0,0003	0,0019*
		(0,0007)	(0,0011)
Männlich		0,0661***	0,0373
		(0,0235)	(0,0298)
Migrationshintergrund		-0,1054***	-0,1032**
		(0,0332)	(0,0459)
Haushaltsgröße		-0,0313***	-0,0450***
		(0,0104)	(0,0154)
Haushaltseinkommen (in tsd €)		0,0012***	0,0006***
		(0,0002)	(0,0002)
Nettovermögen des Haushalts (in tds €)		0,0001***	0,00002
		(0,00002)	(0,00002)
Hochschulabschluss		0,1344***	0,0568
		(0,0293)	(0,0384)
Ostdeutschland		-0,1020***	-0,0753**
		(0,0246)	(0,0315)
Finanzkompetenz (Financial Literacy)		0,0701***	0,0378**
		(0,0123)	(0,0161)
Prohibitive Risikoaversion			-0,4051***
			(0,0350)
Konstante	0,3030***	0,0663	0,4342***
	(0,0289)	(0,0651)	(0,1043)
Beobachtungen	2.926	2.926	1.928
R^2	0,0040	0,1445	0,3094
F-Statistik	2,3301	26,7546***	41,9020***

Anmerkung: Die Tabelle zeigt die Ergebnisse von Regressionen, in denen die Aktienmarktpartizipation auf verschiedene Bankengruppen regressiert wird. In Spalte 1 wird die Aktienmarktpartizipation auf eine Dummy-Variable für Sparkassen (1 = Sparkasse als Hausbank, 0 = andere Bank) und eine Dummy-Variable für Genossenschaftsbank (1 = Genossenschaftsbank als Hausbank, 0 = andere Bank) regressiert, wobei Großbanken als Referenzgruppe dienen. In Spalte 2 wird die Regression mit demografischen Variablen als Kontrollvariablen wiederholt. In Spalte 3 wird zusätzlich noch die Prohibitive Risikoaversion als Kontrollvariable ergänzt und die Regression mit verkleinertem Sample wiederholt. In der Regression werden die Sampling-Gewichte berücksichtigt und die robusten Standardfehler geclustert berechnet.

Anteil der Bankeinlagen

In Tabelle III-5 analysieren wir, inwieweit die hohe Quote an Bankeinlagen am Finanzvermögen von Sparkassenkunden durch deren Eigenschaften erklärt werden kann. Im Gegensatz zur Analyse der Aktienmarktpartizipation bleibt der Effekt, dass Sparkassenkunden einen höheren Anteil ihres Finanzvermögens in Bankeinlagen investieren, bestehen, auch wenn das Signifikanzniveau und die Höhe des Koeffizienten deutlich abnehmen. Der Basiseffekt verringert sich von 5,05 % im unkontrollierten Modell (1)[25] auf 3,59 % in Modell (3). Vergleicht man diesen Effekt mit dem der Genossenschaftsbankkunden, zeigt sich, dass die Effektstärke (3,89 %) im kontrollierten Modell ähnlich hoch ist wie bei den Sparkassenkunden, jedoch ist sie bei den Genossenschaftsbankkunden statistisch nicht signifikant. Dies deutet darauf hin, dass vor allem Sparkassenkunden auch nach Berücksichtigung von Kundeneigenschaften mehr in Bankeinlagen investieren als Kunden von Großbanken – der Referenzkategorie. Ob dieses Ergebnis durch eine Bankberatung beeinflusst wird oder ob unser Modell wichtige Kontrollvariablen nicht berücksichtigt (*omitted variable bias*), können wir in dieser Studie jedoch nicht abschließend beantworten.

Da die Faktoren, die mit der Bankeinlagenquote in Verbindung stehen, nach unserem Kenntnisstand bisher noch nicht untersucht wurden, gehen wir näher auf die Koeffizienten der Kontrollvariablen ein. In Modell (3), das auch die Risikoaversion berücksichtigt, zeigt sich, dass Kunden mit Migrationshintergrund 5,62 % mehr ihres Finanzvermögens in Bankeinlagen investieren als andere Bankkunden. Ein noch stärkerer Zusammenhang besteht bei der Dummy-Variable „Prohibitive Risikoaversion“ (Effektstärke 18,06 %). Vergleicht man Modell (2) und Modell (3), wird ein weiterer interessanter Zusammenhang deutlich: In Modell (2) investieren Personen aus den neuen Bundesländern 3,72 % mehr ihres Finanzvermögens in Bankeinlagen. Dieser Effekt verliert jedoch an Signifikanz, wenn in Modell (3) die Risikoaversion als zusätzliche Variable berücksichtigt wird (Effektstärke 2,51 %). Dies mag dadurch begründet werden, dass Bankkunden aus den neuen Bundesländern im Durchschnitt risikoaverser im Hinblick auf Kapitalanlagen sind.[26] Bezüglich der anderen Kontrollvariablen zeigen sich

[25] Aufgrund der statistisch insignifikanten F-Statistik bei Modell (1) sind die Ergebnisse dieses Modells wiederum vorsichtig zu interpretieren.
[26] Vgl. Laudenbach et al. (2024)

ähnliche Einflussfaktoren wie bei der Analyse der Aktienmarktpartizipation, jedoch mit umgekehrten Vorzeichen: Einkommen, Vermögen, Bildungsniveau und Finanzkompetenz stehen in einem negativen Zusammenhang zum Anteil des Finanzvermögens in Bankeinlagen.

Tabelle III-5: Einlagenquote

	(1) Einlagenquote	(2) Einlagenquote	(3) Einlagenquote
Sparkasse	0,0505***	0,0399**	0,0359*
	(0,0193)	(0,0183)	(0,0217)
Genossenschaftsbank	0,0232	0,0362*	0,0389
	(0,0212)	(0,0206)	(0,0254)
Alter		-0,0004	-0,0014**
		(0,0004)	(0,0006)
Männlich		-0,0211	-0,0127
		(0,0135)	(0,0164)
Migrationshintergrund		0,0538***	0,0562**
		(0,0172)	(0,0220)
Haushaltsgröße		0,0226***	0,0215***
		(0,0061)	(0,0076)
Haushaltseinkommen (in tsd €)		-0,0006***	-0,0004***
		(0,0001)	(0,0001)
Nettovermögen des Haushalts (in tds €)		-0,00004***	-0,00001
		(0,00001)	(0,00001)
Hochschulabschluss		-0,0639***	-0,0364*
		(0,0171)	(0,0197)
Ostdeutschland		0,0372***	0,0251
		(0,0142)	(0,0166)
Finanzkompetenz (Financial Literacy)		-0,0401***	-0,0202**
		(0,0069)	(0,0087)
Prohibitive Risikoaversion			0,1806***
			(0,0188)
Konstante	0,8458***	0,9732***	0,8358***
	(0,0166)	(0,0392)	(0,0540)
Beobachtungen	2.926	2.926	1.928
R^2	0,0061	0,1124	0,2335
F-Statistik	3,8484	17,9089***	22,3419***

Anmerkung: Die Tabelle zeigt die Ergebnisse von Regressionen, in denen die Einlagenquote auf verschiedene Bankengruppen regressiert wird. In Spalte 1 wird die Aktienmarktpartizipation auf eine Dummy-Variable für Sparkassen (1 = Sparkasse als Hausbank, 0 = andere Bank) und eine Dummy-Variable für Genossenschaftsbank (1 = Genossenschaftsbank als Hausbank, 0 = andere Bank) regressiert, wobei Großbanken als Referenzgruppe dienen. In Spalte 2 wird die Regression mit demografischen Variablen als Kontrollvariablen wiederholt. In Spalte 3 wird zusätzlich noch die Prohibitive Risikoaversion als Kontrollvariable ergänzt und die Regression mit verkleinertem Sample wiederholt. In der Regression werden die Sampling-Gewichte berücksichtigt und die robusten Standardfehler geclustert berechnet.

Anteil der Aktienmarktinvestments (Aktienquote)

Abschließend untersuchen wir, ob der geringere Anteil des Vermögens, den Sparkassenkunden am Aktienmarkt investieren, auch in einem multivariaten Kontext bestehen bleibt. Tabelle III-6 zeigt die Schätzungen zu den Faktoren der Aktienquote am Finanzvermögen. Dabei zeigt sich, dass die Effektstärke der Dummy-Variable Sparkasse abnimmt, sobald Kontrollvariablen einbezogen werden. Der Effekt sinkt von -4,86 % in Modell (1)[27] auf -3,85 % in Modell (2) und schließlich auf -3,62 % in Modell (3). Auch das Signifikanzniveau sinkt im letzten Modell auf lediglich 10 %. Die Effektstärke und das Signifikanzniveau ähneln damit den Werten für Genossenschaftsbankkunden (Effektstärke -4,51 % in Modell (3)). Analog zu den Ergebnissen zur Bankeinlagenquote lässt sich der Einfluss der Zugehörigkeit zu Sparkassen oder Genossenschaftsbanken auf die Aktienquote nicht vollständig erklären. Unsere Studie kann erneut nicht abschließend klären, ob die Anlageberatung dieser beiden Bankengruppen weniger auf Alternativen zu Bankeinlagen hinweist als die von Großbanken. Es könnten auch unbeobachtbare Faktoren in den Kundeneigenschaften (*omitted variable bias*) die geringe Investitionsquote in Aktien erklären.

Bezüglich der Kontrollvariablen erhalten wir analoge Ergebnisse zu den Modellschätzungen zur Aktienmarktpartizipation (siehe Tabelle III-4). So steht wiederum die Variable „Prohibitive Risikoaversion“ in einem stark negativen Zusammenhang mit der Aktienquote (Effektstärke -14,2 %). Unsere Analysen zeigen ferner, dass Männer, Personen mit höherer Finanzkompetenz und höherem Bildungsabschluss einen größeren Anteil ihres Finanzvermögens am Aktienmarkt investieren. Ebenso bestätigen unsere Ergebnisse einen positiven Zusammenhang zwischen Einkommen und Vermögen mit der Aktienquote.

[27] Aufgrund der statistisch insignifikanten F-Statistik bei Modell (1) sind die Ergebnisse dieses Modells mit Vorsicht zu genießen.

Tabelle III-6: Aktienquote

	(1) Aktienquote	(2) Aktienquote	(3) Aktienquote
Sparkasse	-0,0486**	-0,0385**	-0,0362*
	(0,0190)	(0,0181)	(0,0214)
Genossenschaftsbank	-0,0272	-0,0394*	-0,0451*
	(0,0206)	(0,0201)	(0,0243)
Alter		0,0004	0,0015***
		(0,0004)	(0,0005)
Männlich		0,0214	0,0128
		(0,0130)	(0,0158)
Migrationshintergrund		-0,0481***	-0,0484**
		(0,0171)	(0,0216)
Haushaltsgröße		-0,0231***	-0,0197***
		(0,0059)	(0,0074)
Haushaltseinkommen (in tsd €)		0,0005***	0,0004***
		(0,0001)	(0,0001)
Nettovermögen des Haushalts (in tsd €)		0,00004***	0,00001
		(0,00001)	(0,00001)
Hochschulabschluss		0,0609***	0,0344*
		(0,0167)	(0,0192)
Ostdeutschland		-0,0372***	-0,0222
		(0,0141)	(0,0163)
Finanzkompetenz (Financial Literacy)		0,0391***	0,0203**
		(0,0068)	(0,0086)
Prohibitive Risikoaversion			-0,1745***
			(0,0178)
Konstante	0,1481***	0,0280	0,1417***
	(0,0164)	(0,0383)	(0,0493)
Beobachtungen	2.926	2.926	1.928
R^2	0,0057	0,1087	0,2354
F-Statistik	3,4433	17,3761***	23,9571***

Anmerkung: Die Tabelle zeigt die Ergebnisse von Regressionen, in denen die Aktienquote auf verschiedene Bankengruppen regressiert wird. In Spalte 1 wird die Aktienmarktpartizipation auf eine Dummy-Variable für Sparkassen (1 = Sparkasse als Hausbank, 0 = andere Bank) und eine Dummy-Variable für Genossenschaftsbank (1 = Genossenschaftsbank als Hausbank, 0 = andere Bank) regressiert, wobei Großbanken als Referenzgruppe dienen. In Spalte 2 wird die Regression mit demografischen Variablen als Kontrollvariablen wiederholt. In Spalte 3 wird zusätzlich noch die Prohibitive Risikoaversion als Kontrollvariable ergänzt und die Regression mit verkleinertem Sample wiederholt. In der Regression werden die Sampling-Gewichte berücksichtigt und die robusten Standardfehler geclustert berechnet.

5. Fazit

Die vorliegende Studie analysiert die Kundenstruktur und das Anlageverhalten von Sparkassen-, Genossenschaftsbanken- und Großbankkunden in Deutschland. Unsere explorativen Analysen dokumentieren signifikante Unterschiede zwischen den Kunden der Bankengruppen, insbesondere im Hinblick auf Risikoeinstellung, Migrationshintergrund, geografische Verteilung sowie der Einkommens- und Vermögenssituation. Während Sparkassenkunden im Durchschnitt risikoaverser sind und einen höheren Anteil ihres Finanzvermögens in Bankeinlagen halten, investieren Großbankkunden häufiger in Aktien und Fonds. Multivariate Analysen deuten darauf hin, dass diese Unterschiede nicht alleine durch Unterschiede in der Bankberatung in den Instituten erklärt werden können, sondern auch auf spezifische Kundenmerkmale zurückzuführen sind.

6. Literaturverzeichnis

Almenberg, J. & Dreber, A. (2015). Gender, stock market participation and financial literacy. *Economics Letters*, *137*, 140–142. https://doi.org/10.1016/j.econlet.2015.10.009

Bertaut, C. C. (1998). Stockholding behavior of U.S. households: Evidence from the 1983-1989 Survey Of Consumer Finances. *The Review of Economics and Statistics*, *80*(2), 263–275. https://doi.org/10.1162/003465398557500

Bucher-Koenen, T., Janssen, B., Knebel, C. & Tzamourani, P. (2023). Financial literacy, stock market participation, and financial wellbeing in Germany. *Journal of Financial Literacy and Wellbeing*, *1*(3), 486–513. https://doi.org/10.1017/flw.2024.5

Campbell, J. Y. (2006). Household finance. *The Journal of Finance*, *61*(4), 1553–1604. https://doi.org/10.1111/j.1540-6261.2006.00883.x

Cole, S., Paulson, A. & Shastry, G. K. (2014). Smart money? The effect of education on financial outcomes. *The Review of Financial Studies*, *27*(7), 2022–2051. https://doi.org/10.1093/rfs/hhu012

Deutsche Bundesbank. (2023a). *Monthly Report April 2023*. https://www.bundesbank.de/resource/blob/744244/f073835ff84728760f2e1b007f5af54a/472B63F073F071307366337C94F8C870/2023-04-monatsbericht-data.pdf

Deutsche Bundesbank. (2023b). *PHF User Guide Version 4.0*.

Deutsche Bundesbank. (2023c). *Studie zur wirtschaftlichen Lage privater Haushalte (PHF)*. Abgerufen am 27.11.2024 von https://www.bundesbank.de/de/bundesbank/forschung/fdsz/forschungsdaten/phf-856346.

Deutsche Bundesbank. (2024). *Marktanteile der Bankengruppen an der gesamten Bilanzsumme der Bankenbranche in Deutschland im November 2023 [Graph]*. Abgerufen am 14.11.2024 von https://de.statista.com/statistik/daten/studie/166006/umfrage/marktanteile-der-bankengruppen-in-deutschland/.

Deutscher Sparkassen- und Giroverband. (2024). *Finanzbericht 2023 der Sparkassen-Finanzgruppe*. https://www.dsgv.de/sparkassen-finanzgruppe/publikationen/Finanzbericht_2023.html

Deutsches Aktieninstitut. (2022). *Deutschland und die Aktie*. Abgerufen am 27.11.2024 von https://www.dai.de/detail/aktionaerszahlen-2021-weiter-auf-hohem-niveau/.

Fagereng, A., Gottlieb, C. & Guiso, L. (2017). Asset market participation and portfolio choice over the life-cycle. *The Journal of Finance*, *72*(2), 705–750. https://doi.org/10.1111/jofi.12484

Fey, J.-C., Lerbs, O., Schmidt, C. & Weber, M. (2020). *Risk attitude and capital market participation: Is there a gender investment gap in Germany?* (20-080). ZEW - Leibniz Centre for European Economic Research. https://EconPapers.repec.org/RePEc:zbw:zewdip:20080

Guiso, L., Haliassos, M. & Jappelli, T. (2003). Household stockholding in Europe: Where do we stand and where do we go? *Economic Policy*, *18*(36), 123–170. https://doi.org/10.1111/1468-0327.00104

Hahn, O. (1996). Die genossenschaftliche Bankwirtschaft der Bundesrepublik Deutschland. *Zeitschrift für das gesamte Genossenschaftswesen*, *46*(1), 164–178. https://doi.org/10.1515/zfgg-1996-0125

Haliassos, M. & Bertaut, C. C. (1995). Why do so few hold stocks? *The Economic Journal*, *105*(432), 1110–1129. https://doi.org/10.2307/2235407

Hartmann-Wendels, T., Pfingsten, A. & Weber, M. (2019). *Bankbetriebslehre* (7. Aufl.). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-58290-9

Kaustia, M., Conlin, A. & Luotonen, N. (2023). What drives stock market participation? The role of institutional, traditional, and behavioral factors. *Journal of Banking & Finance*, *148*, 106743. https://doi.org/10.1016/j.jbankfin.2022.106743

Laudenbach, C., Malmendier, U. & Niessen-Ruenzi, A. (2024). *The long-lasting effects of experiencing communism on attitudes towards financial markets*. SAFE Working Paper (Nr. 429). https://doi.org/10.2139/ssrn.3526926

Lusardi, A. & Mitchell, O. S. (2008). Planning and financial literacy: How do women fare? *American Economic Review*, *98*(2), 413–417. https://doi.org/10.1257/aer.98.2.413

Oehler, A., Horn, M. & Wendt, S. (2024). Investment in risky assets and participation in the financial market: Does financial literacy matter? *International Review of Economics*, *71*(1), 19–45. https://doi.org/10.1007/s12232-023-00432-9

Priewasser, E. (2001). *Bankbetriebslehre* (7., erw.u. vollst. überarb. Aufl.). Oldenbourg.

Rieger, M. O., Wang, M. & Hens, T. (2015). Risk preferences around the world. *Management Science*, *61*(3), 637–648. https://doi.org/10.1287/mnsc.2013.1869

Schildbach, J. (2019). *Private Haushalte in Deutschland: Regionale Unterschiede im Bankgeschäft* (Deutschland-Monitor). http://hdl.handle.net/11159/4082

Statistisches Bundesamt. (2023). *Bevölkerung nach Altersgruppen in Deutschland*. Abgerufen am 27.11.2024 von https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Bevoelkerungsstand/Tabellen/bevoelkerung-altersgruppen-deutschland.html.

Stolper, O. & Walter, A. (2019). Birds of a feather: The impact of homophily on the propensity to follow financial advice. *The Review of Financial Studies*, *32*(2), 524–563. https://doi.org/10.1093/rfs/hhy082

Thomas, A. & Spataro, L. (2018). Financial literacy, human capital and stock market participation in Europe. *Journal of Family and Economic Issues*, *39*(4), 532–550. https://doi.org/10.1007/s10834-018-9576-5

van Rooij, M., Lusardi, A. & Alessie, R. (2011). Financial literacy and stock market participation. *Journal of Financial Economics*, *101*(2), 449–472. https://doi.org/10.1016/j.jfineco.2011.03.006

von Lüde, R. (2013). Anlageverhalten auf Finanzmärkten. *Wirtschaftsdienst*, *93*(5), 328–336. https://doi.org/10.1007/s10273-013-1529-x

7. Appendix

Appendix III-A: Variablen Definition und Ausprägung

Variable	Ausprägung	Definition
Panel A: Hausbank		
Sparkasse	1/0	Indikatorvariable, die den Wert 1 annimmt, wenn die Hausbank des Haushalts eine Sparkasse ist, und sonst den Wert 0
Genossenschaftsbank	1/0	Indikatorvariable, die den Wert 1 annimmt, wenn die Hausbank des Haushalts eine Genossenschaftsbank ist, und sonst den Wert 0
Großbank	1/0	Indikatorvariable, die den Wert 1 annimmt, wenn die Hausbank des Haushalts eine Großbank ist, und sonst den Wert 0
Panel B: Kundencharakteristika		
Alter	[18; 100]	Gibt das Alter des Kompetenzträger Haushaltsfinanzen des Haushalts in Jahren an
Männlich	1/0	Indikatorvariable, die den Wert 1 annimmt, wenn der Kompetenzträger Haushaltsfinanzen des Haushalts männlich ist, und sonst den Wert 0
Migrationshintergrund	1/0	Indikatorvariable, die den Wert 1 annimmt, wenn der Kompetenzträger Haushaltsfinanzen des Haushalts einen Migrationshintergrund hat, und sonst den Wert 0
Haushaltsgröße	[1; 8]	Gibt die Anzahl der im Haushalt lebenden Personen an
Haushaltseinkommen (in tsd €)	k€	Gibt das gesamte jährliche Brutto-Haushaltseinkommen in Tausend € an
Nettovermögen des Haushalts (in tsd €)	k€	Gibt das gesamte Nettovermögen aller Haushaltsmitglieder in Tausend € an
Ostdeutschland	1/0	Indikatorvariable, die den Wert 1 annimmt, wenn der Haushalt in einem der neuen Bundesländer liegt, und sonst den Wert 0
Finanzkompetenz (Financial Literacy)	[0;3]	Gibt den Wert an, den der Kompetenzträger Haushaltsfinanzen bei einem 3-teiligen Fragenset zur Finanzkompetenz erzielt hat
Prohibitive Risikoaversion	1/0	Indikatorvariable, die den Wert 1 annimmt, wenn der Haushalt bei der Befragung angegeben hat, dass er nicht bereit ist irgendein Risiko einzugehen, um eine höhere Rendite zu erzielen
Panel C: Klassifizierung nach Gemeindegröße		
[0; 5.000] Einwohner	1/0	Indikatorvariable, die den Wert 1 annimmt, wenn der Haushalt in einer Gemeinde mit bis zu 5000 Einwohnern liegt, und sonst den Wert 0
[5.001; 20.000] Einwohner	1/0	Indikatorvariable, die den Wert 1 annimmt, wenn der Haushalt in einer Gemeinde mit 5.001 bis 20.000 Einwohnern liegt, und sonst den Wert 0
[20.001; 100.000] Einwohner	1/0	Indikatorvariable, die den Wert 1 annimmt, wenn der Haushalt in einer Gemeinde mit 20.001 bis 100.000 Einwohnern liegt, und sonst den Wert 0
[100.001; 500.000] Einwohner	1/0	Indikatorvariable, die den Wert 1 annimmt, wenn der Haushalt in einer Gemeinde mit 100.001 bis 500.000 Einwohnern liegt, und sonst den Wert 0
[> 500.000] Einwohner	1/0	Indikatorvariable, die den Wert 1 annimmt, wenn der Haushalt in einer Gemeinde mit mehr als 500.000 Einwohnern liegt, und sonst den Wert 0
Panel D: Anlageverhalten		
Aktienmarkt-partizipation	1/0	Indikatorvariable die den Wert 1 annimmt, wenn der Haushalt am Aktienmarkt partizipiert (d.h. Aktien oder Investmentfonds besitzt), und sonst den Wert 0
Finanzvermögen bei Banken (in €)	€	Gibt das gesamte liquidierbare Finanzvermögen an, das der Haushalt bei Banken und Finanzdienstleistern hält. Dies beinhaltet neben den Bankeinlagen und Aktienmarktinvestitionen auch Unternehmen- und Staatsanleihen.
Bankeinlagen (in €)	€	Gibt die Summe der Bankeinlagen des Haushalts an
Investition am Aktienmarkt (in €)	€	Gibt die Summe der Aktienmarktinvestitionen des Haushaltes an. Dies beinhaltet neben Aktien auch Investmentfonds.
Anteil Einlagen am Finanzvermögen	[0; 100%]	Stellt den Anteil der Bankeinlagen am Finanzvermögen bei Banken da
Anteil Aktien am Finanzvermögen	[0; 100%]	Stellt den Anteil der Aktienmarktinvestitionen am Finanzvermögen bei Banken da

Anmerkung: Diese Tabelle enthält die Ausprägungen und Definitionen aller Variablen. Die Ausprägungen geben die Werte an, die die Variable annehmen kann, sowie die Einheit, in der sie angegeben wird.

IV. Financial Support Among Siblings: The Relevance of Personal and Family Characteristics

Co-author:

Charlotte Clara Becker

Own share:

60%

This article is currently under review at:

Journal of Marriage and Family

Previous versions of this paper have been presented at the following conferences and workshops:

- 119th Annual Meeting of the American Sociological Association (ASA), 2024, Montreal, Canada
- European Consortium for Sociological Research (ECSR) Conference, 2024, Barcelona, Spain
- 44th Sunbelt Conference, 2024, Edinburgh, Great Britain
- Fifth Research in Behavioral Finance Conference (RBFC), 2024, Amsterdam, Netherlands

Financial Support Among Siblings:
The Relevance of Personal and Family Characteristics

Charlotte Clara Becker[a] Martin G. Becker[b]

Abstract

Objective: This study investigates the prevalence and drivers of financial solidarity among siblings in young adulthood, expanding the focus of family support research beyond the parent-child dyad.
Background: Drawing on the intergenerational solidarity model, the study tests need-based and ability-based perspectives and further analyzes how relationship characteristics and family context condition financial support.
Method: Using data from the US KINMATRIX sample (N=3,124; 5,556 dyads), the study employs a dual analytical strategy. The analysis combines theory-driven linear probability models with algorithmically determined Lasso regressions. This approach not only tests specific hypotheses but also rigorously isolates the most robust predictors of support.
Results: 16% of young adults received financial support from at least one sibling. Results partially support an ability-based perspective: transfers were driven by the provider's capacity rather than recipient hardship. Family context was also critical; support was congruent with parental transfers and significantly more likely for those with an immigration background. Although an individual's gender did not predict their overall likelihood of receiving support, gender moderated pathways: for women, emotional closeness was linked to the receipt and provision, whereas for men, parenthood and immigration background were particularly relevant.
Conclusion: Siblings constitute a vital secondary financial safety net. However, this support is less a response to disadvantages and more a function of the provider's economic capacity and the broader family culture of support.

Keywords: siblings, financial support, young adulthood, intragenerational solidarity, family, gender differences

[a] charlotte.becker@gesis.de, GESIS – Leibniz Institute for the Social Sciences, Unter Sachsenhausen 6-8, 50670 Köln, Germany
[b] martin.becker@wirtschaft.uni-giessen.de, Justus Liebig University Gießen, Department of Financial Services, Licher Straße 74, 35394 Gießen, Germany

1. Introduction

The family serves as a fundamental safety and support network across the life course. Family members provide each other with a variety of different forms of support, such as emotional assistance and practical help. One form of support that becomes especially relevant in times of unforeseen hardships, such as illness, unemployment, or economic crisis, is financial support. The importance of this private financial safety net is underscored by the fact that a significant portion of the population lacks sufficient personal savings to sustain themselves through prolonged periods of financial difficulty (Lusardi et al., 2011). However, financial transfers within families are not exclusively focused on alleviating immediate hardships. Instead, they are also employed to improve long-term socioeconomic outcomes, such as funding education (Kuperberg, 2023), supporting entrepreneurial aspirations (Basu & Parker, 2001; Lee & Persson, 2016), or achieving home ownership (Cook, 2025).

This duality of financial support – alleviating hardships and investing in the future – might be particularly relevant during young adulthood. Young adults often carry significant student loan debt (Board of Governors of the Federal Reserve System, 2024; Fry & Cilluffo, 2024) while being employed in initially low-paying jobs (Abel & Deitz, 2015; Bengali & Hobijn, 2014). At the same time, they are attempting to establish independent households (Houle & Berger, 2015; Mezza et al., 2020) and start their own families (Addo, 2014; Bozick & Estacion, 2014). These competing financial pressures likely create a high demand for financial support from family members.

So far, previous research on financial support within the family, independent of the receivers' age, has focused mainly on parents as the providers of financial assistance (Fingerman et al., 2012; McGarry & Schoeni, 1997; Silverstein & Bengtson, 1997). This "vertical" perspective centering on intergenerational solidarity, although valuable, overlooks the broader family network and the potential for "lateral" support structures such as siblings.

Siblings share a unique, enduring bond, often outlasting relationships with parents, yet their role as economic actors within family networks remains understudied. To gain a deeper understanding of family dynamics, it is essential to move beyond the parent-child dyad and extend the intergenerational solidarity model by exploring how siblings contribute to the family's role as a private safety net. The present study fills this gap by investigating the prevalence of financial support among young adult siblings and identifying the factors that drive the likelihood of receiving such assistance.

To determine these drivers, we examined several distinct theoretical mechanisms regarding family solidarity. First, we examined a need-based perspective and an ability-based perspective, investigating whether sibling transfers are primarily driven by the recipient's economic needs or by the provider's financial capacity. Second, we explored the relevance of the sibling relationship and the broader family context. We analyzed whether support is conditioned by the quality of the specific sibling bond and how the wider family setting influences these transfers – for example whether siblings step in to compensate for the lack of parental support or if sibling support acts congruently with parent behavior. Finally, we examined gendered pathways to solidarity. Rather than simply asking whether women receive and provide more support than men, we investigated how gender conditions the drivers of support.

For our analyses, we utilized egocentric network data from the US subsample of the KINMATRIX data set. Focusing on young adults, aged 25 to 35, KINMATRIX provides detailed information on the existence, characteristics, and support behaviors of broader kinship networks. This allows for a granular analysis of how personal characteristics, relationship attributes, and family context drive financial solidarity among young adult siblings. We employed linear probability models to identify primary determinants of sibling financial solidarity. However, to distinguish fundamental drivers from statistical noise, we explicitly contrasted these theory-driven models with a machine learning approach using Lasso regressions. This allowed us to validate our variable selection through a strict, algorithmic filter that identified the most robust predictors by penalizing less relevant variables – effectively removing them from the model. This ensured our findings captured true structural patterns rather than artifacts of model specification. Finally, to capture distinct pathways to financial support, we further disaggregated our analysis by gender, examining how determinants differ for men and women. Before elaborating further on the data used and methods applied, the following sections introduce the empirical background and theoretical considerations on which the study is based.

2. Theoretical and Empirical Background

Family members support each other through various means. A common model categorizing family support and family relationships is Bengtson and Roberts' (1991) intergenerational solidarity model (ISM). The original model distinguishes six dimensions of solidarity between adult children and their parents or grandparents: *structural solidarity* (number of family members and their health and location), *associative solidarity* (frequency of contact), *affectual solidarity* (emotional closeness), *normative solidarity* (family values and expectations), *consensual solidarity* (shared values and attitudes), and *functional solidarity* (exchange of various forms of support). Although initially designed to characterize intergenerational solidarity within the family (Bengtson & Roberts, 1991), we extend the model's application to include intragenerational relationships within the family, specifically sibling relationships. We use the ISM model as a base for our theoretical considerations and expectations. Our study focuses on one specific form of functional solidarity, financial support. However, other dimensions and concepts associated with the model are also applicable to our research question, particularly in their potential relevance as predictors of sibling financial support.

Currently, research on financial solidarity among siblings is scarce, especially when focusing on young adulthood. Therefore, in the following literature overview, we will not only present studies on this exact issue but rather broaden the scope and also look at studies from adjacent fields that yield relevance for the research question.

One particularly promising stream of research is financial support between other family members, particularly between parents and their (adult) children. These studies reveal various factors influencing whether and to what extent parents provide financial support to their children. Many of these factors could also be relevant when considering financial support among siblings in young adulthood.

Firstly, parental support appeared to be driven by personal characteristics of the receiving child. Younger adult children were more likely to receive financial support from their parents, and when they received it, they received more than older children (Berry, 2008; Björnberg & Latta, 2007; McGarry, 2016). Besides age, some studies further suggested that parents also take their children's gender into consideration, with daughters being more likely to receive financial support but not receiving larger amounts (McGarry, 2016). Other studies, however, do not find any differences between daughters and sons (Berry, 2008; Björnberg & Latta, 2007). Further, the children's socio-economic situation appeared relevant. Parents were more likely to support a lower-income child than a high earner, and also when support was given, the amount appeared to be higher (Björnberg & Latta, 2007; Dunn & Phillips, 1997). Additionally, parents appeared to consider their children's current life situations, with children who were in school, just graduated, currently unemployed, recently purchased a house, and having children themselves being more likely to receive financial support and receiving larger amounts (Berry, 2008; Dunn & Phillips, 1997; Fingerman et al., 2009; Kim et al., 2012; McGarry, 2016; McGarry & Schoeni, 1997). Grandparents, on the other hand, did not increase the frequency of financial support to a grandchild based on the characteristics of the grandchild or their current financial situation (Huo et al., 2018).

Secondly, parents' personal characteristics seem to influence their children’s likelihood of receiving financial support from them and the amount received. With higher education, increasing income, and greater wealth in general, parents were more likely to provide financial support to their children, and they provided larger amounts (Berry, 2008; Emery, 2013; McGarry, 2016). This was also the case for parents who were married or in a partnership rather than single (Berry, 2008; Emery, 2013; Fomby & Kravitz-Wirtz, 2019). However, conflicting findings exist regarding the significance of relationship status (Kim et al., 2012; McGarry & Schoeni, 1994). Additionally, paternal health affected the probability of support, with parents having a disability being less likely to provide financial support and those in poor health providing smaller amounts (Berry, 2008; Kim et al., 2012; McGarry, 2016).

Thirdly, multiple studies found links between certain dimensions of the ISM and the likelihood and extent of financial transfers. Higher affectual solidarity, so more positive relationships and closeness in the dyad, for example, increased the likelihood of transfers as well as the amount transferred from parents, especially mothers, to their children (Kim et al., 2012; Swartz et al., 2011). Additionally, children who shared higher associational solidarity with their parents, i.e., had more regular contact with their parents, also received larger transfers (Kim et al., 2012). Similarly, structural solidarity, specifically the geographical proximity between parents and their children, appeared to be relevant. Children living close to their parents were significantly more likely to receive financial support from their parents and received more than those living farther away (Berry, 2008; Kim et al., 2012; McGarry & Schoeni, 1994). However, studies on grandparental financial support do not indicate a relevance of these solidarity dimensions. Neither the closeness of the relationship between the grandparent and the grandchild, nor the relationship between the grandparent and the parent, nor the geographic proximity to the grandchild was significantly linked to the amount or frequency of financial support (Huo et al., 2018; Michalski & Shackelford, 2005).

Fourthly, the overall family setting was identified as a relevant factor across studies on financial support within the family. The family setting describes characteristics within the family that go beyond the two individuals involved in the financial transfer. One such aspect that appeared particularly relevant was the overall number of children in the family, the so-called sibship size. An increase in the number of children reduced the likelihood and amount of financial transfers from parents to an individual adult child (Berry, 2008; McGarry, 2016). Therefore, only children were far more likely to receive financial support than children with (multiple) siblings (Emery, 2013). Further, having siblings who are (still) co-residing with the parents decreased the likelihood and amount received (Berry, 2008). The number of grandchildren a grandparent had, however, did not affect the frequency of financial support to an individual grandchild (Huo et al., 2018). Another relevant factor regarding the family was the family's immigration background. Previous research on financial support between parents and their children revealed parents with an immigration background to be less likely to financially support their children, but their children to be more likely to support them (Klaus & Baykara-Krumme, 2017).

Although studies on the financial support between parents, grandparents, and adult children provide an interesting base for the discussion of financial support between siblings, studies examining the relationships between siblings themselves should also be considered. Research on financial support between siblings is currently scarce. However, insight might be gained from the broader field of sibling solidarity and particularly from studies on functional solidarity between siblings. The mechanisms and pathways identified in this literature may also shape expectations regarding both the prevalence of financial support among siblings and the conditions under which it is most likely to occur.

Indeed, when examining studies on functional solidarity among siblings, factors similar to those identified in research on parental financial support were evident. As in parent-child transfers, the personal characteristics of the person receiving solidarity appeared to be important. For example, siblings increased contact with their brothers and sisters who recently separated from their partners or started parenthood (Buyukkececi & Çineli, 2024), and those severely impaired by an illness were more likely to receive help around the house (Mulder & van der Meer, 2009). Similarly, siblings in a partnership were less likely than unpartnered siblings to receive solidarity in the form of help around the house and emotional support (Voorpostel et al., 2007). Regarding material help, measured as a combination of giving and lending money or goods and sending gifts, flowers, or cards, however, the relationship and parenthood status of the receiving sibling did not appear to be relevant (Eriksen & Gerstel, 2002). Surprisingly, the receiving sibling's employment status also appeared to be negligible (Eriksen & Gerstel, 2002; Mulder & van der Meer, 2009).

Regarding the potential provider's personal characteristics, results from previous studies appear somewhat inconsistent. Whereas some suggest that the health and income of the potential provider do not affect their likelihood to provide practical and emotional support to a sibling (Voorpostel et al., 2007), others indicate that an increase in family income also increases the total amount of practical, emotional, and material help provided to the siblings within the family (Eriksen & Gerstel, 2002) and that individuals with higher incomes are more likely to be part of a general sibling exchange network either as a provider or a recipient (White & Riedmann, 1992). Higher education of the potential provider was associated with an increased likelihood of providing help with odd jobs around the house as well as emotional support and an increase in the amount of support provided (Voorpostel et al., 2007; White & Riedmann, 1992). Age, on the other hand, appeared to negatively impact the functional solidarity among siblings. With increasing age, individuals were less likely to be involved in the exchange of support, both regarding the provision and receipt of support (White, 2001; White & Riedmann, 1992).

In addition to the characteristics of the potential recipient and provider, the relationship between the siblings involved appears to be a key indicator for the likelihood of support in the sibling dyad. As described above for the parent-child dyad, different dimensions of the ISM were linked to sibling support. Higher affectual solidarity increased the likelihood of being involved in the exchange of functional solidarity within the sibling group in general and increased the provision of practical, personal, and material help and its likelihood within the individual sibling dyad (Eriksen & Gerstel, 2002; Voorpostel et al., 2007; Weaver et al., 2003; White & Riedmann, 1992). Similarly, increased contact (associational solidarity) and increased geographical proximity (structural solidarity) also led to an increased likelihood of receiving and providing functional solidarity in general (White, 2001; White & Riedmann, 1992). However, the effect of geographical proximity highly depended on the type of support considered. Whereas geographical proximity appeared to be relevant for practical help, typically requiring physical presence, like help around the house (Eriksen & Gerstel, 2002; Mulder & van der Meer, 2009; Voorpostel et al., 2007), it was not linked to the provision of emotional and material support (Eriksen & Gerstel, 2002; Voorpostel et al., 2007), which can also be provided over larger distances. This stands in contrast to the increase in financial support found in parent-child dyads. Another relationship aspect that appears to be relevant in the provision of functional support within siblings is the gender composition of the dyad. Women were found to be more likely to both receive and provide help in sibling dyads (Mulder & van der Meer, 2009; Voorpostel et al., 2007; White, 2001; White & Riedmann, 1992). Especially sister-sister dyads exhibited a great amount of functional solidarity (Spitze & Trent, 2006; Weaver et al., 2003). However, first indications suggest that for material support, gender might not be as relevant (Eriksen & Gerstel, 2002).

Lastly, the family setting appeared important to functional solidarity among siblings. For example, an increase in sibship size also increased an individual's likelihood of involvement in a sibling exchange network, either as a provider or a recipient of support (White & Riedmann, 1992). However, although an increase in the sibship size increased an individual's likelihood to receive and provide functional solidarity to/from at least one sibling (White, 2001), it also led to a decreased likelihood of support provision in the individual sibling dyads (Eriksen & Gerstel, 2002; Voorpostel et al., 2007). Regarding the above-discussed immigration background, only a few studies on functional solidarity included this aspect of the family setting, and no consistent pattern could be identified. Although some studies (Mulder & van der Meer, 2009; White, 2001) did not indicate a significant effect on the general and household support received and provided, another study (Spitze & Trent, 2006) suggests that immigration background positively affects some but not all types of practical support. Regarding financial or material support, unfortunately, this aspect has not yet been examined.

Taken together, existing research on financial support between parents and children, as well as studies of sibling functional solidarity more broadly, points to a set of recurring factors that affect the likelihood of support exchanges. Across both strands of literature, personal characteristics of the potential recipient and of the potential provider, the relationship between them, and the broader family setting emerged as particularly relevant. However, although these factors have been examined in the context of intergenerational financial support and in relation to non-financial forms of functional solidarity among siblings, their relevance for financial support between siblings remains largely unexplored. Building on the literature review, the following sections develop theoretical expectations regarding the likelihood of receiving financial support from a sibling, structured around the four domains identified.

The first group of factors identified in family solidarity research was personal characteristics of the potential recipient. Family members appear to provide support based on the needs of the potential recipient, meaning the family acts as a safety net to fall back on in times of hardship (Eggebeen & Davey, 1998). The need for financial support could be indicated by various factors, for example, age, with younger individuals in early stages of their careers and family planning phases potentially needing more financial support. But also, through a low socioeconomic status in general, or through poor health, having children, or being in education due to the high costs associated with these statuses (Looney & Yannelis, 2024; Malik, 2019; Richard et al., 2018). Existing research on financial solidarity provided by parents to their adult children supports the conceptualization of the family as a safety net and the application of a need-based allocation system (Berry, 2008; Dunn & Phillips, 1997; Fingerman et al., 2009; McGarry, 2016; McGarry & Schoeni, 1997). Therefore, we expect the personal characteristics of the potential recipient to also be relevant for financial support from siblings, with those indicating greater needs being more likely to receive support.

Not only do we expect the financial solidarity among siblings to be need-based, the literature also introduces the notion of an ability-based provision of solidarity. This means that family members who are more able to provide support are also more likely to do so (Berry, 2008; Emery, 2013; McGarry, 2016). We apply this notion to financial solidarity among siblings and expect individuals to be more likely to receive financial support from a sibling if that sibling has a greater ability to provide it. Different personal characteristics may serve as indicators of this ability. Firstly, the socioeconomic status of the sibling in general could provide information on their ability, given that those who have higher socioeconomic statuses tend to have greater financial means. Additionally, education and age were found to be relevant factors in the provision of other forms of sibling support (Descartes, 2007; White, 2001). They might also be relevant for financial support among young adult siblings, because individuals' incomes and savings tend to increase with education and over the lifespan (Francis-Devine, 2024; Guzman et al., 2023; Niehues & Stockhausen, 2020), giving better-educated and older siblings a greater ability to provide financial support. Similar to when needs are evaluated, the health of a sibling should be considered a potential limitation on their ability to provide support. Siblings in poor health may, for example, face financial constraints themselves, leaving them with fewer resources to support other family members. Another sibling characteristic that should be considered with regard to the ability-based approach is gender. Because women tend to earn less, accumulate less wealth, and have less access to personal spending money (Cantillon, 2013; Drechsel-Grau et al., 2022; Kassenboehmer & Sinning, 2014; Ruel & Hauser, 2013; Sierminska et al., 2010), they might also be less able to financially support their siblings. This, in turn, leads to the expectation that individuals will be less likely to receive financial support from a sister than from a brother.

Besides the needs of the potential recipient and the abilities of the potential provider, factors related to the sibling relationship should also be considered. As indicated above, relationship aspects such as emotional closeness, regular contact, and geographic proximity appeared to be positively linked to financial transfers between parents and their children (Lennartsson et al., 2010; Motel & Szydlik, 1999) as well as other types of solidarity between siblings (Eriksen & Gerstel, 2002; Spitze & Trent, 2006; Voorpostel et al., 2007). It could be assumed that this also holds for financial transfers between siblings. Individuals might feel a stronger obligation to support those siblings with whom they have regular contact and are emotionally close. Further, individuals might also be more likely to ask these siblings for financial support rather than asking siblings with whom they are not close and with whom they do not regularly interact. We therefore expect individuals to be more likely to receive financial support from a sibling if they share a closer relationship, live in proximity to one another, and have more regular contact. Another important aspect regarding the sibling relationships is the gender composition of the sibling dyad. Across studies on various forms of support, women were found to provide and receive more solidarity from their siblings than men (Gilligan et al., 2020; White, 2001), to have larger family support networks overall (Descartes, 2007), to take over more caregiving responsibilities, and to be more family-oriented in general (Gerstel, 2011; Szydlik, 2016).

Especially sister-sister relationships were found to exhibit more support behavior than other sibling combinations (Descartes, 2007; Spitze & Trent, 2006; Voorpostel et al., 2007; Weaver et al., 2003). Therefore, women might also be more likely to receive financial assistance from their siblings, especially their sisters.

In addition to the relationship factors, we also assume the overall family setting in which the siblings are embedded to be of relevance. The behaviors and abilities of one family member can influence those of others. Previous literature introduced the congruence hypothesis, which suggests that behaviors and affection modeled in the parent-child relationship are mirrored in sibling relationships (Boer et al., 1992; Derkman et al., 2011; Noller, 2005; Seginer, 1998). Meaning more positive relationships and supportive behavior between children and their parents might also lead to more positive relationships and support among the family's children. In contrast to the congruence hypothesis, the compensation hypothesis suggests that conflict-ridden or nonexistent relationships with certain family members may lead to closer and more positive bonds with other family members (Voorpostel & Blieszner, 2008; Whiteman et al., 2011). Previous research on parent-child and sibling relationships generally provided more support for the congruence hypothesis, with siblings modeling their relationships after the relationships that they have with their parents (Brody et al., 1994; Derkman et al., 2011; Noller, 2005; Seginer, 1998) rather than compensating for a challenging parent-child bond with a closer sibling relation (Voorpostel & Blieszner, 2008). The notion of the congruence hypothesis is further supported by literature on financial socialization, with the financial socialization of children primarily happening within the family context (Gudmunson & Danes, 2011; Shim et al., 2010). Parental financial socialization and behavior influence individuals' financial literacy, financial well-being, and financial behavior (Shim et al., 2010; Zhao & Zhang, 2020), particularly regarding their investments in family members (LeBaron, 2019). Their socialization, therefore, is likely to also affect their need for financial support, as well as their ability and willingness to provide such support for a sibling. Applied to the financial solidarity between young adult siblings, both the congruent hypothesis and available literature on financial socialization suggest that an individual might be more likely to receive financial support from a sibling if their parents also provide support, rather than if the parents do not.

Another family factor that needs to be considered when discussing financial transfers within the family is the number of siblings present. Children in larger families are less likely to receive financial assistance from their parents (Emery, 2013; McGarry, 2016). This might also be the case with siblings, as an increased family size can limit individuals' capacity for strong ties with each other (White, 1994), as well as siblings' ability to provide financial support in every sibling dyad. However, a larger family could also lead to more family-oriented values and closer ties (White, 1994), and thereby to an increased likelihood for an individual to receive support from a sibling. One last family characteristic is the immigration background. Although the immigration background is a shared family characteristic, it is also linked to the previously mentioned considerations of needs, abilities, and socialization. Individuals with an immigration background might have greater financial needs because they tend to earn less, accumulate less wealth, and, due to discrimination, might have less access to forms of financial support outside the family realm, such as loans and credit cards (Algan et al., 2010; Campbell & Kaufman, 2006; Cookson et al., 2025; Hao, 2004). Therefore, their siblings might provide them with a financial safety net, beyond what families without an immigration background might offer their members. On the other hand, these exact factors might also make it more difficult for a sibling in an immigrant family to provide financial support. Lastly, similar to family size, the immigration background of a family may also affect an individual's likelihood of receiving financial support from a sibling through family values. In Western societies, family solidarity norms tend to be stronger among immigrants than non-immigrants (Arends-Tóth & Van De Vijver, 2008; Carnein & Baykara-Krumme, 2013). If individuals perceive a greater obligation to provide support to a sibling due to their socialization and the family values they share, they might also be more likely to provide that support. The expectations regarding the total effect of an immigration background are unclear, as decreased abilities to provide financial support could be outweighed by potentially greater needs and increased family solidarity norms.

Overall, we expect all four areas identified to affect an individual's likelihood of receiving financial support from their sibling: The needs of the potential recipient, the ability of the potential provider, the relationship shared between the siblings, and the family context in which they are embedded. Although expectations are straightforward for some factors across these four groupings, others, like gender, might influence the likelihood of receiving financial support through multiple paths, and the expectation of the overall effect might therefore be less clear.

3. Methods

3.1 Data

For our analyses, we used the US-subsample of the KINMATRIX data set, a cross-sectional data set collected during the winters of 2022/2023 and 2023/2024 (Leopold, Becker, et al., 2025). Comprising large-scale, ego-centric family network data from adults aged 25 to 35, this data set offers detailed insights into their family compositions and relationships. In the survey, the respondents answered questions on the existence of family members, including siblings, on the relationship they share with each family member, and on personal characteristics of themselves and all their family members.

Because the data was collected via online panel and soft sampling quotas, it is not strictly representative of the age group in the US. However, previous publications have indicated a high alignment with information from traditional probability surveys and official statistics (Leopold, Raab, et al., 2025).

3.2 Sample Selection

The initial US sample included 5,005 respondents with 7,945 respondent-sibling dyads consisting of full biological siblings sharing the exact same set of parents. We focused exclusively on full biological siblings, in the following simply called siblings, because half- and stepsiblings might differ systematically in years of co-residence, age spacing, and exposure to stepfamily resource regimes, which would conflate kin type with unobserved socialization factors and undermine comparability.

From this sample, our selection followed a multi-step procedure. First, we restricted the sample to respondent-sibling dyads, effectively excluding respondents without siblings. Further, we excluded dyads where the respondent reported a sibling to be deceased or did not know whether the sibling was alive or not. This restriction ensured that age remained a meaningful variable in the analysis and, crucially, that all siblings in the sample had equal opportunity to provide or receive financial support at the time of the survey. Additionally, we removed dyads where the respondent did not know the sibling's name. This criterion served as a proxy for a minimal level of social contact and relationship recognition; we posit that financial interactions are highly implausible in family settings where basic identifying information is unknown. Furthermore, we excluded sibling dyads based on two age criteria: we removed siblings younger than 14 due to the presumed lack of financial autonomy among children and young teenagers. We also excluded those with a respondent-sibling age gap exceeding 35 years, because such age gaps were likely data errors given the biological limits of the mother's reproductive window. Lastly, we excluded sibling dyads as well as whole respondents with missing values regarding respondent, sibling, relationship, or family characteristics. The final sample used included 3,124 participants, with a total of 5,556 siblings. Appendix IV- A indicates the number of respondents and dyads remaining at each step of the sample selection process.

3.3 Variables

Our main variable of interest is $Financial\ Support$ received in a specific respondent-sibling dyad, which was measured by the question *"Who has ever given or loaned you a larger amount of money?"*. The participants received a list of all family members they mentioned earlier in the survey, including siblings. They were asked to indicate all family members from whom they had received money. $Financial\ Support$ was coded as 0 if the respondent had never received money from the sibling and as 1 if they had.

We analyzed the determinants of financial support using four categories of independent variables covering the needs of the potential recipient (respondent characteristics), the ability of the potential provider (sibling characteristics), the characteristics of the relationship, and the broader family context. All variables, including sibling characteristics, are based on the respondent's reports.

First, we operationalized the respondent's characteristics. $Female$ is a dummy variable taking the value of 1 for women and 0 for men based on the self-reported gender of the respondent. Respondents who answered *"Prefer not to say"* or *"Other gender or no gender"* were coded as missings and are therefore excluded from the analysis. This choice was made due to low case counts. Age is included as a continuous variable in years. $Subjective\ Social\ Status$ (SSS) captures the respondent's perception of their own socioeconomic standing using the MacArthur Scale. Respondents placed themselves on a 'social ladder' relative to society at large, ranging from 1 (least money, education, and respected jobs) to 10 (most money, education, and respected jobs). For the analysis, we recoded this variable to a 0–9 scale so that the lowest rank serves as a meaningful reference point (0) in the regression models. Finally, $Bad\ Health$ is a dummy variable capturing the respondent's health status, derived from the question, *"How is your health in general? Would you say it is…?"*.

The variable takes a value of 0 if the respondent replied, "*very good*" or "*good*," and 1 if they indicated *"fair," "bad,"* or *"very bad"* health.

Regarding siblings' characteristics, we employed similar concepts. $Sibling\ Female$ is a dummy variable coded as 1 for sisters and 0 for brothers. $Sibling\ Age$ indicates the age of the sibling in years. We also include the $Sibling\ SSS$, based on the respondent's estimation of the sibling's position on the same social ladder. $Sibling\ Bad\ Health$ is a dummy variable indicating the health status of the sibling. Unlike the respondent measure, this variable is derived from the question, *"Please indicate for these persons if they have any long-term mental or physical health problems."* It is coded as 1 if the respondent indicates the presence of such problems and 0 otherwise.

Next, we included variables describing the relationship within the sibling dyad. $Sister - Sister\ Dyad$ is a dummy variable which equals 1 if both the respondent and the sibling are female and 0 otherwise. To compare the socioeconomic standing within the dyad, we included $Sibling\ SSS\ >\ Respondent\ SSS$, a dummy variable indicating whether the estimation of the respondent's sibling's status is higher than the respondent's own reported status. We also included $Sibling\ Older\ than\ Respondent$, a dummy variable coded as 1 if the sibling is older than the respondent and 0 otherwise. Geographical proximity was captured by the dummy variable $Distance\ >\ 1\ Hour$, set to 1 when the travel time between the respondent and their sibling exceeded one hour. Respondents originally selected from five categories ranging from *'Lives in the same building'* to *'1 hour or longer'*. We recoded the first four categories (all under 60 minutes) as 0 to distinguish local from long-distance ties.

We accounted for the interpersonal dynamics within the sibling dyad through two specific indicators. Firstly, $High\ Emotional\ Closeness$, a dummy variable built on the respondent's answer to the following question, *"How close do you feel emotionally to each of these persons today?"*. The answer provided was recoded as 1 if the respondent indicated they are either *"pretty close"* or *"very close"* to their sibling and 0 if they indicated *"somewhat close", "not too close"*, or *"not at all close"*. Secondly, $Monthly\ Contact$, capturing interaction frequency, based on the question, *"How often are you in contact with each of these persons, adding up all visits, letters, phone calls, etc.?"*. It was coded as 1 if personal contact occurs at least once a month and 0 otherwise.

Finally, we included a set of variables characterizing the broader family context. $Sibship\ Size$ is a continuous variable indicating the number of full biological siblings (including the respondent) in the family. To capture the respondent's own family formation, we included a dummy variable, $Respondent\ has\ Children$, indicating whether the respondent has children or not. We also account for intergenerational transfers using $Parental\ Support\ Received$. As with the dependent variable, this measure was derived from the question asking respondents to identify all family members who had ever given or loaned them a larger amount of money. The variable takes a value of 1 if the respondent selected one or both parents and 0 otherwise. Lastly $Immigration\ Background$ is a dummy variable indicating whether the family has a recent immigration history; the variable equals 1 if the respondent or at least one parent was born outside the US and 0 otherwise, thereby covering first- and second-generation immigrants. Descriptive statistics for all variables used in the analysis are provided in Appendix IV-B.

3.4 Analytical Approach

To capture the prevalence of support, we first examined the data at two distinct levels: the respondent level, to capture the overall prevalence of support from any sibling, and the dyadic level, to examine support from specific siblings. To contextualize the findings, we also used a parent sample to estimate the share of respondents receiving financial support from at least one parent and the share of parents providing support overall, applying the above-described exclusion criteria also to the parent-child sample.

Next, we applied multivariate analyses, using the relationship (dyad) between the respondent and a single sibling as the unit of observation. We started our multivariate analyses with linear probability models (LPM) to examine the potential determinants of financial support. We regressed our binary indicator, $Financial\ Support$, on four sets of attributes. In Model 1 (M1) we only included the personal characteristics of the respondent, in Model 2 (M2) we solely focused on the characteristics of the sibling, in Model 3 (M3) we inspected the relevance of various relationship indicators, and in Model 4 (M4) we directed our attention to family characteristics. Finally, in Model 5 (M5), all respondent, sibling, relationship, and family characteristics were considered together. For all models (M1–M5), we included robust standard errors clustered at the respondent level in order to account for individuals with multiple siblings and therefore multiple respondent-sibling dyads in the data set.

Although we favor LPMs for the direct interpretability of their marginal effects, we acknowledge that the mean of the dependent variable falls outside the range where LPM typically yields the most consistent fit (Long & Freese, 2014). To ensure our results are robust to the choice of model, we re-estimated our analyses using a set of logistic regressions.

We complemented our theory-driven LPMs with an algorithmic variable selection approach using linear Lasso regressions. This step allowed us to rigorously isolate the most robust predictors without relying solely on pre-specified theoretical assumptions. Whereas standard regression models risk overfitting and rely on the researchers' selection choices, Lasso addresses these limitations by applying a penalty term that systematically shrinks the coefficients of less informative variables to zero. This effectively distinguishes robust structural determinants from statistical noise (Tibshirani, 1996).

We trained the models on a random 50% split of our sample (clustered by respondent) and utilized a three-step selection strategy to test the stability of our predictors under increasing levels of strictness. First, we used 10-fold cross-validation (CV) (again clustered by respondent) to minimize out-of-sample prediction error (Hastie et al., 2009). This served as our permissive baseline, retaining a broad set of variables that contribute predictive power, even if their individual predictive contributions are small.

Second, we employed the Adaptive Lasso (Zou, 2006). This two-step procedure applies varying weights to the penalty term – penalizing smaller coefficients more heavily – to correct for bias and ensure selection consistency. Finally, we used the Bayesian Information Criterion (BIC), which imposes a stricter penalty on model complexity to isolate the most robust determinants (Zou et al., 2007).

We then evaluated the performance of these models on the remaining 50% holdout sample. This progression from predictive accuracy to consistent selection and finally to strict parsimony allowed us to distinguish between variables that are merely helpful for prediction versus those that act as fundamental drivers of support (Hastie et al., 2009).

After identifying the key predictors of support through Lasso regressions, we extended the original LPM analyses to capture gender-specific dynamics. Although our primary analysis tested for a direct effect of gender, this does not necessarily capture the full picture. Instead, it is possible that the dynamics of support differ depending on gender – meaning different factors might be relevant for men compared to women. To this end, we ran four additional regressions. Model 6a and Model 6b (M6a and M6b) provide separate analyses for the respondents' gender, with M6a only including male respondents and M6b only considering female respondents. This allows us to inspect whether the same characteristics are relevant for men and women regarding receiving financial support from their siblings. Model 7a and Model 7b (M7a and M7b), on the other hand, differentiate between the gender of the sibling, with M7a focusing on respondents' relationships with their brothers and M7b focusing on their relationships with sisters. This allows us to inspect whether the same characteristics are relevant for receiving financial support from brothers and from sisters. Again, we included robust standard errors clustered at the respondent level. Also, just like for M1-M5, we performed robustness checks using logistic regressions.

4. Results

In our sample, 11% of the respondents' siblings provided them with financial support, and 16% of respondents reported having received such support from at least one sibling. Although these numbers were substantially lower than the respective numbers for parents in the sample (49% provided support to the respondents, 61% of respondents have received support from at least one parent), these findings underscore that siblings are an important group to consider when studying family financial support.

4.1 Determinants of Financial Support

An inspection of the block-wise models revealed that respondent characteristics alone (M1) held negligible explanatory power, with key indicators of need failing to reach significance even in isolation. In contrast, factors related to the sibling's capacity, relational cohesion, and family context appeared relevant across specifications. As M5 incorporates all variable blocks simultaneously and demonstrated the best fit, we centered our interpretation on these estimates, referencing the partial models only where the inclusion of the other variable blocks led to noteworthy changes. Figure IV-1 visualizes the coefficients from the fully specified M5; full results for all models are available in Appendix IV-C.

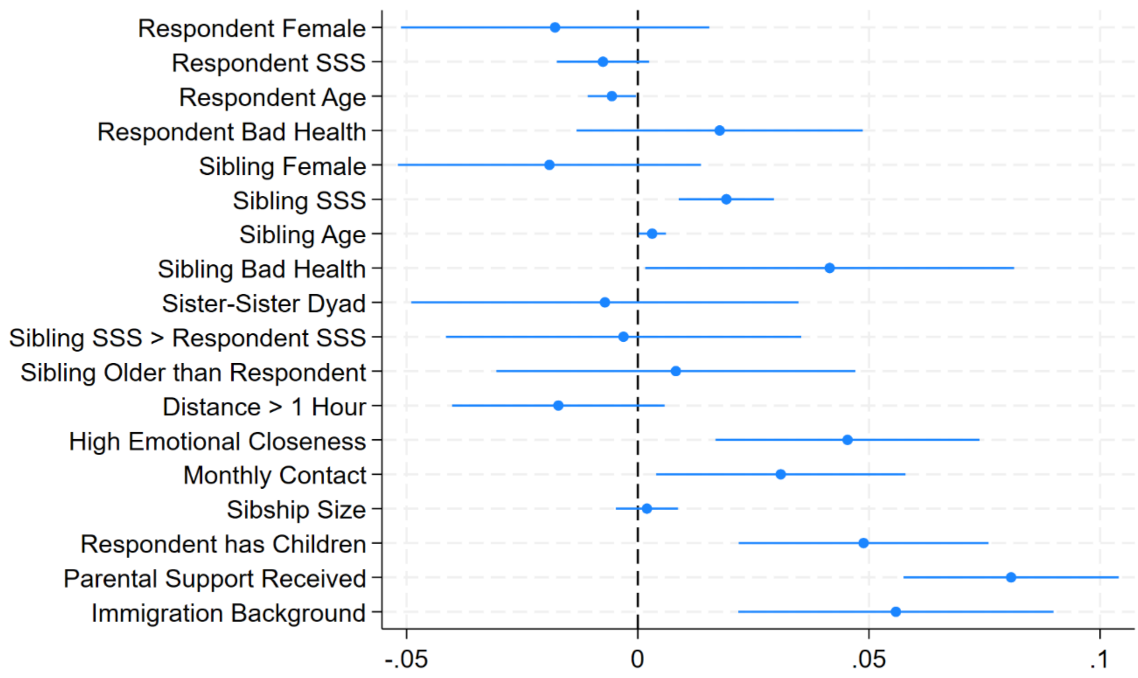

Figure IV-1: Determinants of Sibling Financial Support (LPM Estimates)

Notes: This figure reports results from the full LPM regression (Model 5) predicting the receipt of financial support. The model includes respondent, sibling, relationship, and family characteristics simultaneously. Point estimates are shown with 95% confidence intervals based on robust standard errors clustered at the respondent level.

Consistent with the block-wise analysis, the fully specified model offered little support for a strictly need-based perspective. Respondent characteristics such as gender, social status, and health remained statistically insignificant. Only Age emerged as a significant predictor, with younger respondents being significantly more likely to have received support. However, given its insignificance in M1 and because the dependent variable indicates ever having received support rather than having received support in, for example, the last year, this association could also be a period effect and should be treated with caution.

Results regarding sibling characteristics somewhat supported the ability-based perspective. Respondents were significantly more likely to have received financial support from siblings with higher SSS and higher Age, suggesting that economic capacity was a prerequisite for support. Additionally, in the simpler M2, $Sibling\ Female$ was significantly negatively related to the likelihood of receiving financial support, indicating that brothers were more likely to provide support than sisters. The loss of significance in the full M5 appears to be driven by multicollinearity with the sister-sister dyad indicator; in untabulated robustness checks where we re-estimated M5 excluding the $Sister - Sister\ Dyad$ variable, $Sibling\ Female$ returned to statistical significance. In contrast to these indicators of capacity, our findings regarding sibling health status contradicted this 'ability' perspective. Counterintuitively, poor health in a sibling was associated with a higher likelihood of receiving support from them, although this association only reached statistical significance in the fully specified model.

In terms of relationship characteristics, gender composition did not show a significant effect; sister-sister dyads were not significantly more or less likely to engage in financial support than other sibling combinations. Further, in the restricted M3, respondents were significantly more likely to have received support when their sibling’s social status exceeded their own. However, this relationship lost statistical significance following the inclusion of predictors besides the relationship characteristics. Similarly, whereas respondents younger than their siblings were significantly more likely to have received support in Model 3, this association was no longer significant once absolute age levels were controlled for. It is important to note that this loss of significance again appears to be driven by multicollinearity rather than a lack of association. In untabulated robustness checks where we re-estimated M5 excluding the absolute levels of Age and SSS , for both the respondent and the sibling, the coefficients for both relative difference indicators remained statistically significant.

Living over an hour away from a sibling did not appear to be significantly linked to the likelihood of receiving financial support from them. We did, however, observe that $High\ Emotional\ Closeness$ and $Monthly\ Contact$ both were linked to significantly higher likelihoods of receiving financial support from a sibling. Respondents who describe a sibling as 'pretty' or 'very' close emotionally were roughly 4.5 percentage points more likely to have received financial support from that sibling. Similarly, when a sibling was in contact with the respondent at least once a month, the likelihood of receiving financial support was 3.1 percentage points higher.

Regarding the family context, several factors turned out to be relevant. Whereas the total number of siblings was unrelated to the likelihood of having received financial support from a specific sibling, we found that respondents who had children were significantly more likely to have received support. And even though this variable represents a change in family structure, functionally it aligned with the need-based perspective, suggesting that siblings are more likely to step in to help with the increased financial pressures associated with raising a family. Furthermore, respondents who received financial support from their parents were significantly more likely to also receive support from a sibling. In fact, parental support constituted the strongest predictor in our model, increasing the likelihood of sibling support by 8.1 percentage points. This strong link between parental and sibling support lends credence to the congruence hypothesis mentioned earlier, suggesting that support behaviors are often consistent across different family relationships. Additionally, individuals with an immigration background exhibited a significantly higher probability of receiving financial support from a sibling. With an estimated increase of 5.6 percentage points, $Immigration\ Background$ represented the second strongest predictor in the model.

To ensure our results are robust to the choice of model, we reran all five models using a set of logistic regressions. These regressions yielded results which are virtually unchanged from our main results (see Appendix IV-E), confirming that the identified drivers of sibling support are consistent regardless of modeling approach.

4.2 Algorithmic Variable Selection (Lasso Results)

Whereas the LPM and logistic regressions focused on estimating the associations between specific characteristics and financial support, our Lasso regressions employed an algorithmic variable selection strategy. By applying increasingly strict penalties, the Lasso procedure isolated the most robust predictors of support, excluding variables that contribute little to predictive accuracy. Table IV-1 shows that this algorithmic variable selection corroborates the findings discussed above.

First, consistent with our theoretical framework, the Lasso procedure using cross-validation retained almost all predictors included in our full LPM specification (17 of 18). This suggests that although certain variables may lack individual statistical significance in the LPM, they collectively contribute to minimizing out-of-sample prediction error. However, given that CV-Lasso can be overly permissive, this result highlights the necessity of the stricter Adaptive and BIC approaches to isolate the most robust predictors.

Second, the Adaptive Lasso refined the model to 12 predictors. Crucially, the algorithm systematically eliminated most indicators of the respondent's needs (only $Bad\,Health$ remained), as well as $Sister - Sister\,Dyad$, $Monthly\,Contact$ and $Sibship\,Size$. Conversely, it retained all indicators of sibling capacity and most variables related to relationship characteristics and family context, the areas also identified as particularly relevant in the LPM.

Third, the even stricter BIC selection reduced the model to 6 'core' predictors. Notably, this algorithmic selection process again prioritized sibling capacity (SSS, Age, and being the older sibling), $High\,Emotional\,Closeness$ and family context ($Immigration\,Background$ and $Parental\,Support\,Received$), whereas it systematically dropped all indicators of the respondent's needs, including $Bad\,Health$.

Overall, the variables prioritized by the algorithms mirror the areas identified as most relevant for support in the LPM. The selection pattern, therefore, provides independent evidence for the ability-based perspective over the need-based perspective and highlights the relevance of the provider's economic capacity and the broader family culture of support. The stability of these predictors across increasingly strict selection penalties – with R^2 remaining relatively stable from the full model (0.0695; in sample) to the BIC model (0.0507; out of sample) – further validates the robustness of our findings.

Table IV-1: Robustness of Variable Selection for Predictors of Received Sibling Support

	Full Model (M5)	CV-Lasso	Adaptive-Lasso	BIC-Lasso
Respondent				
Female	✓	✓		
Subjective Social Status (SSS)	✓			
Age	✓	✓		
Bad Health	✓	✓	✓	
Sibling				
Female	✓	✓	✓	
SSS	✓	✓	✓	✓
Age	✓	✓	✓	✓
Bad Health	✓	✓	✓	
Relationship				
Sister-Sister Dyad	✓	✓		
Sibling SSS > Respondent SSS	✓	✓	✓	
Sibling Older than Respondent	✓	✓	✓	✓
Distance > 1 Hour	✓	✓	✓	
High Emotional Closeness	✓	✓	✓	✓
Monthly Contact	✓	✓		
Family				
Sibship Size	✓	✓		
Respondent has Children	✓	✓	✓	
Parental Support Received	✓	✓	✓	✓
Immigration Background	✓	✓	✓	✓
Observations (Sibling Dyads)	5,556	5,556	5,556	5,556
Number of Predictors	18	17	12	6
R^2	0.0695	0.0536	0.0530	0.0507

Notes: This table compares variable selection across the full theoretical specification (Model 5) and three Lasso selection criteria. CV-Lasso uses 10-fold cross-validation; Adaptive Lasso employs a two-step consistent selection procedure; and BIC-Lasso utilizes the Bayesian Information Criterion to prioritize model parsimony. A checkmark (✓) indicates that a variable was included in the final model. Number of predictors indicates the count of variables included in the final model. R^2 represents the proportion of variance explained; for Lasso models, this is the out-of-sample prediction R^2.

4.3 Gendered Pathways of Solidarity

Although the LPMs only identified sibling gender to be a significant predictor of financial support, and neither the respondent nor the siblings' gender was among the remaining predictors in the strictest Lasso specification, our stratified analyses revealed that both respondent and sibling gender conditioned specific pathways of support (Figure IV-2; see Appendix IV-D for full regression results).

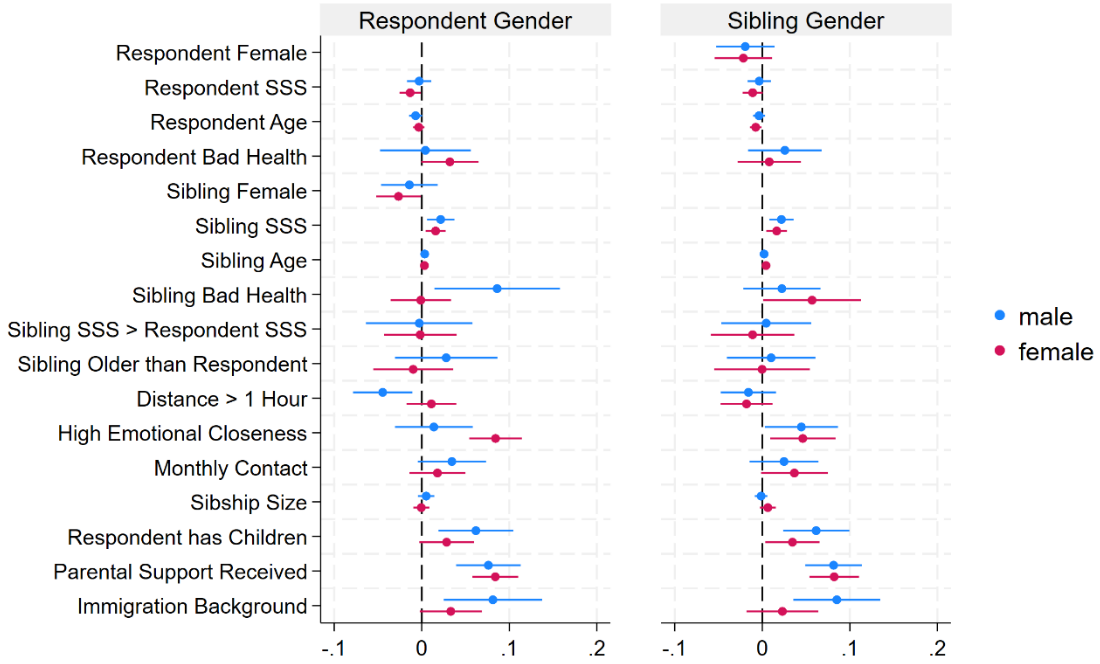

Figure IV-2: Determinants of Sibling Financial Support by Gender (LPM Estimates)

Notes: This figure displays coefficients from LPM regressions predicting the receipt of financial support. The left panel presents results stratified by the respondent's gender (Male vs. Female), and the right panel presents results stratified by the sibling's gender (Brother vs. Sister). Point estimates are shown with 95% confidence intervals based on robust standard errors clustered at the respondent level.

First, regarding respondent characteristics, the stratified analysis revealed distinct patterns that were masked in the aggregate model. Although *Respondent SSS* was insignificant in the main model, it emerged as a significant negative predictor for female recipients, meaning women with higher SSS were less likely to have received support from their siblings than women with lower SSS. Similarly, while *Respondent Age* was significant in the main model, our results indicate it was only significant for female providers, meaning respondents were more likely to have received financial support from a sister if they were younger. *Respondent Female* and *Respondent Bad Health* remained insignificant in all specifications.

Second, regarding sibling characteristics, we observed that the provider's gender was a significant predictor for female recipients. Women were significantly less likely to have received financial support from sisters than from brothers, whereas for male recipients, the sibling's gender was not significant. Furthermore, the counterintuitive finding regarding poor provider health appeared to be driven by distinct gender combinations. Specifically, the coefficient was significant only for male recipients and female providers, indicating that brothers were more likely to have received support from a sick sibling, and sisters were more likely to provide support despite their own poor health.

Third, regarding relationship characteristics, we observed a distinct divergence in the drivers of support. Geographical proximity emerged as a significant predictor only for male respondents. Living at a greater distance reduced the likelihood of support for male recipients, whereas for women and potential providers, regardless of gender, distance played no significant role. Conversely, for women, receipt of financial support was highly contingent on emotional closeness; the positive association was strong and significant for female respondents, whereas for men, the association was considerably weaker and not statistically different from zero. However, when examining potential providers, the association was statistically significant regardless of gender.

Finally, concerning the family context, we observed a gendered divergence regarding the role of parenthood and immigration background. The financial pressures of raising a family acted as a stronger driver for men: male respondents who had children were significantly more likely to have received support, whereas for female respondents, this association was not statistically significant. When examining providers, however, both brothers and sisters appeared responsive to the 'need' signal sent by a sibling raising a family, although the association was weaker and less significant for sisters compared to brothers. Similarly, the link between immigration background and support was driven largely by male family members. Men in families with immigration backgrounds were significantly more likely to have received financial support. Regarding the source of support, immigration background significantly increased the likelihood of having received support from brothers, whereas it was not a significant predictor of having received support from sisters.

Across these stratified findings, it should be noted that while the set of significant predictors differed clearly both by the gender of the potential recipient and provider, the coefficients were not always statistically distinguishable. This suggests that gender influences which characteristics are the most salient drivers of support rather than creating fundamentally opposing behaviors.

As with the main analysis, we re-estimated these stratified models using logistic regressions. The results remained virtually unchanged and are reported in Appendix IV-F.

5. Discussion

This study set out to examine the prevalence and determinants of financial solidarity among siblings in young adulthood – a life phase characterized by significant economic transitions. Using data from the US subsample of the KINMATRIX data set, we analyzed 5,556 sibling dyads to examine the circumstances under which siblings act as a source of financial support. By drawing on established theories explaining financial support from parents and non-financial assistance among siblings, we analyzed whether these mechanisms also hold true for financial transfers between siblings. Our descriptive results confirm that siblings are a relevant, albeit secondary, source of financial solidarity, with 16% of young adults reporting having received financial support from at least one sibling.

Regarding the determinants of support, our results offer a nuanced evaluation of the need-based and ability-based theoretical perspectives. Contrary to the expectation that the family acts as a safety net, directing resources to those with the greatest needs (Berry, 2008; Eggebeen & Davey, 1998; Kim et al., 2012; McGarry, 2016), we find that personal hardships such as poor health and low social status did not trigger support from siblings. The notable exception to this pattern is parenthood, with respondents that have children being significantly more likely to have received support. This suggests that sibling support in young adulthood may be less about alleviating general disadvantages and more about facilitating specific life stages, a pattern previously observed where childless siblings provide resources to those raising families (Voorpostel et al., 2007).

Though we find little support for the need-based perspective, our results somewhat support the ability-based perspective (Cook, 2025; Emery, 2013): respondents were significantly more likely to have received financial support from siblings who were male, older, and had higher subjective social status. We did, however, observe a counterintuitive nuance in our fully specified model: siblings in poor health were more likely to have provided support. Although this appears to contradict the ability perspective, it may reflect 'altruism born of suffering' (Staub & Vollhardt, 2008), where personal challenges heighten empathy even amid resource constraints. Overall, however, the pattern indicates that financial solidarity between siblings is less a response to disadvantages and more a function of the provider's economic ability.

In terms of relationship characteristics, our analysis supports the notion that stronger interpersonal bonds – specifically associational and affectual solidarity – facilitate functional solidarity. Consistent with the literature on parental financial support (Lennartsson et al., 2010; Motel & Szydlik, 1999), emotional closeness and monthly contact were among the strongest predictors of financial transfers.

Regarding the broader family context, our results clearly favor the congruence hypothesis over the compensation hypothesis (Derkman et al., 2011; Hank & Steinbach, 2018; Voorpostel & Blieszner, 2008). We find that financial support from parents is a strong positive predictor of support from siblings. Rather than siblings stepping in to compensate for absent parental resources, we find that families that provide support across generations also tend to provide it within generations. Furthermore, our analysis indicates that sibship size does not constrain sibling-to-sibling solidarity. This stands in distinct contrast to intergenerational transfers, where empirical studies find that a larger number of children reduces the support available to each (Emery, 2013; Roksa, 2019).

Additionally, we find that individuals with an immigration background were significantly more likely to have received financial support from a sibling. Theoretically, we posited two competing mechanisms here: a lower ability to provide due to structural barriers versus a higher need for support often coupled with normative obligations (Arends-Tóth & Van De Vijver, 2008; Carnein & Baykara-Krumme, 2013). The fact that this link remains significant even when controlling for aspects of need and ability reinforces the interpretation that a form of normative solidarity – specifically cultural family values – rather than purely economic disparities, is at play.

Our conclusions were further supported by the algorithmic variable selection using Lasso regressions, which isolated the most robust structural determinants through increasingly strict penalization. While the CV-Lasso retained almost all variables – confirming the collective predictive value of our theoretical framework – the stricter algorithms provided independent empirical grounds for challenging the need-based perspective. The Adaptive-Lasso excluded most measures of recipient hardship, whereas it retained all indicators of sibling capacity and most relational and family context characteristics. The even more restrictive BIC-Lasso distilled the model even further, dropping all respondent characteristics and focusing exclusively on aspects of provider capacity ($Sibling\ SSS$, $Sibling\ Age$ and being the older sibling), $High\ Emotional\ Closeness$, and family context ($Parental\ Support\ Received$ and $Immigration\ Background$). This replicated the inference of M5, showing that most of the variables with the strongest and significant statistical associations in the LPM were indeed the most robust predictors in the data.

Finally, our stratified analyses focused on gendered aspects in the receipt of financial support from siblings. However, whereas previous literature simply focused on whether women give and receive more support than men (Descartes, 2007; Voorpostel et al., 2007; White, 2001), we investigated differences not only in the likelihood of receipt but also in its respective drivers.

And although our main model suggests that respondents' gender per se does not predict the likelihood of receiving financial support, the stratified results revealed that gender fundamentally structures the pathways to solidarity. Emotional closeness, for example, increased the likelihood of support regardless of provider gender. However, it was a statistically significant predictor only for female and not for male recipients. Similarly, while geographical proximity was not a significant predictor for female recipients, men were less likely to have received support from a sibling living more than an hour away. Additionally, men responded more strongly to tangible family circumstances. For men – both as recipients and providers – support was strongly linked to parenthood and immigration background. Especially immigration background was a strictly male-driven factor: it was a significant predictor for men both as recipients and providers but remained insignificant for women. This indicates that in families with an immigration background, financial support among siblings is heavily skewed from and towards men. These stratified findings contribute to the literature by demonstrating that distinct gendered pathways can exist even when aggregated outcomes appear equal, highlighting the need to examine gender-specific pathways in family sociology (Fingerman et al., 2020; Suitor et al., 2006).

Our study is, of course, not without limitations. First, our analysis relies exclusively on the respondent's perspective. Although we account for sibling and relationship characteristics – such as their health, their social status, and the emotional closeness between the siblings– these measures are based entirely on the respondent's reporting rather than on data directly from the siblings themselves. This approach may introduce measurement error, particularly if respondents are unaware of their sibling's true financial situation or if their perception of the relationship quality differs from their sibling's views.

Second, our measure of financial support is broad. The survey question asked, "Who has ever given or loaned you a larger amount of money?", conflating gifts with loans and leaving the interpretation of what constitutes a "large" sum up to the respondent. Although the literature suggests that the distinction between giving and lending is often blurred in close family networks (Heath & Calvert, 2013), the reliance on a subjective threshold introduces potential measurement error. For instance, the perception of what constitutes a 'large amount' may vary systematically depending on the respondent's socioeconomic context.

Finally, our data is cross-sectional, limiting our ability to make causal claims. Although we observe associations between relationship quality and the receipt of financial support, we cannot rule out reverse causality – for instance, that receiving support improves the relationship rather than the relationship driving the support. Furthermore, although the survey asks who has "ever" provided support, retrospective questions are subject to recall bias (Dex, 1995). Respondents may be more likely to report recent transfers or particularly memorable instances of support, potentially underreporting older or less salient exchanges.

In conclusion, this study highlights that siblings constitute a vital, yet often overlooked, component of the private financial safety net in young adulthood. Our results indicate that financial solidarity among siblings is less a response to individuals' hardship and more a function of the provider's economic capacity, the sibling relationship, and the broader family context. Furthermore, our analysis uncovered that although gender does not dictate the likelihood of support, it fundamentally shapes the determinants of support.

Ultimately, the results showed that incorporating siblings and gender into the analyses of families as private financial safety nets is essential for a comprehensive understanding of how financial support is organized within young adulthood and within families in general.

6. References

Abel, J. R., & Deitz, R. (2015). Underemployment in the early careers of college graduates following the Great Recession (Staff Report No. 749). Federal Reserve Bank of New York. https://hdl.handle.net/10419/130641

Addo, F. R. (2014). Debt, cohabitation, and marriage in young adulthood. Demography, 51(5), 1677–1701. https://doi.org/10.1007/s13524-014-0333-6

Algan, Y., Dustmann, C., Glitz, A., & Manning, A. (2010). The economic situation of first and second-generation immigrants in France, Germany and the United Kingdom. The Economic Journal, 120(542), F4–F30. https://doi.org/10.1111/j.1468-0297.2009.02338.x

Arends-Tóth, J., & Van De Vijver, F. J. R. (2008). Family relationships among immigrants and majority members in the Netherlands: The role of acculturation. Applied Psychology, 57(3), 466–487. https://doi.org/10.1111/j.1464-0597.2008.00331.x

Basu, A., & Parker, S. C. (2001). Family finance and new business start-ups. Oxford Bulletin of Economics and Statistics, 63(3), 333–358. https://doi.org/10.1111/1468-0084.00224

Bengali, L., & Hobijn, B. (2014). The wage growth gap for recent college grads. FRBSF Economic Letter. https://ideas.repec.org//a/fip/fedfel/00024.html

Bengtson, V. L., & Roberts, R. E. L. (1991). Intergenerational solidarity in aging families: An example of formal theory construction. Journal of Marriage and the Family, 53(4), 856. https://doi.org/10.2307/352993

Berry, B. (2008). Financial transfers from living parents to adult children: Who is helped and why? The American Journal of Economics and Sociology, 67(2), 207–239. https://doi.org/10.1111/j.1536-7150.2008.00568.x

Björnberg, U., & Latta, M. (2007). The roles of the family and the welfare state: The relationship between public and private financial support in Sweden. Current Sociology, 55(3), 415–445. https://doi.org/10.1177/0011392107076083

Board of Governors of the Federal Reserve System. (2024). Economic well-being of U.S. households in 2023. Board of Governors of the Federal Reserve System. https://doi.org/10.17016/8960.1

Boer, F., Goedhart, A., & Treffers, P. D. A. (1992). Siblings and their parents. In F. Boer & J. Dunn (Eds.). Children's sibling relationships: Developmental and clinical issues (pp. 41-53). Lawrence Erlbaum Associates Publishers.

Bozick, R., & Estacion, A. (2014). Do student loans delay marriage? Debt repayment and family formation in young adulthood. Demographic Research, 30, 1865–1891. https://doi.org/10.4054/DemRes.2014.30.69

Brody, G. H., Stoneman, Z., & McCoy, J. K. (1994). Contributions of family relationships and child temperaments to longitudinal variations in sibling relationship quality and sibling relationship styles. Journal of Family Psychology, 8(3), 274–286. https://doi.org/10.1037/0893-3200.8.3.274

Buyukkececi, Z., & Çineli, B. (2024). Adult sibling relationships: The impact of cohabitation, marriage, separation, and childbearing. Journal of Marriage and Family, 86(3), 698–717. https://doi.org/10.1111/jomf.12949

Campbell, L. A., & Kaufman, R. L. (2006). Racial differences in household wealth: Beyond Black and White. Research in Social Stratification and Mobility, 24(2), 131–152. https://doi.org/10.1016/j.rssm.2005.06.001

Cantillon, S. (2013). Measuring differences in living standards within households. Journal of Marriage and Family, 75(3), 598–610. https://doi.org/10.1111/jomf.12023

Carnein, M., & Baykara-Krumme, H. (2013). Einstellungen zur familialen Solidarität im Alter: Eine vergleichende Analyse mit türkischen Migranten und Deutschen. Journal of Family Research, 25(1), 29–52. https://doi.org/10.20377/jfr-162

Cook, J. (2025). Lateral financial assistance with home ownership: Understanding the role of siblings. Sociology, 59(1), 55–71. https://doi.org/10.1177/00380385241261038

Cookson, J. A., Guttman-Kenney, B., & Mullins, W. (2025). Immigration and credit in America (SSRN Scholarly Paper No. 5187062). Social Science Research Network. https://doi.org/10.2139/ssrn.5187062

Derkman, M. M. S., Engels, R. C. M. E., Kuntsche, E., van der Vorst, H., & Scholte, R. H. J. (2011). Bidirectional associations between sibling relationships and parental support during adolescence. Journal of Youth and Adolescence, 40(4), 490–501. https://doi.org/10.1007/s10964-010-9576-8

Descartes, L. (2007). Exchange in adult sibling relationships: Support and tension. International Journal of Sociology of the Family, 33(2), 303–325.

Dex, S. (1995). The reliability of recall data: A literature review. Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique, 49(1), 58–89. https://doi.org/10.1177/075910639504900105

Drechsel-Grau, M., Peichl, A., Schmid, K. D., Schmieder, J. F., Walz, H., & Wolter, S. (2022). Inequality and income dynamics in Germany. Quantitative Economics, 13(4), 1593–1635. https://doi.org/10.3982/QE1912

Dunn, T. A., & Phillips, J. W. (1997). The timing and division of parental transfers to children. Economics Letters, 54(2), 135–137. https://doi.org/10.1016/S0165-1765(97)00018-9

Eggebeen, D. J., & Davey, A. (1998). Do safety nets work? The role of anticipated help in times of need. Journal of Marriage and the Family, 60(4), 939–950. https://doi.org/10.2307/353636

Emery, T. (2013). Intergenerational transfers and European families: Does the number of siblings matter? Demographic Research, 29, 247–274. https://doi.org/10.4054/DemRes.2013.29.10

Eriksen, S., & Gerstel, N. (2002). A labor of love or labor itself: Care work among adult brothers and sisters. Journal of Family Issues, 23(7), 836–856. https://doi.org/10.1177/019251302236597

Fingerman, K. L., Cheng, Y.-P., Tighe, L., Birditt, K. S., & Zarit, S. (2012). Relationships between young adults and their parents. In A. Booth, S. L. Brown, N. S. Landale, W. D. Manning, & S. M. McHale (Eds.), Early Adulthood in a Family Context (pp. 59–85). Springer. https://doi.org/10.1007/978-1-4614-1436-0_5

Fingerman, K. L., Huo, M., & Birditt, K. S. (2020). Mothers, fathers, daughters, and sons: Gender differences in adults' intergenerational ties. Journal of Family Issues, 41(9), 1597–1625. https://doi.org/10.1177/0192513X19894369

Fingerman, K. L., Miller, L., Birditt, K., & Zarit, S. (2009). Giving to the good and the needy: Parental support of grown children. Journal of Marriage and Family, 71(5), 1220–1233. https://doi.org/10.1111/j.1741-3737.2009.00665.x

Fomby, P., & Kravitz-Wirtz, N. (2019). Family systems and parents' financial support for education in early adulthood. Demography, 56(5), 1875–1897. https://doi.org/10.1007/s13524-019-00807-0

Francis-Devine, B. (2024). Average earnings by age and region. https://commonslibrary.parliament.uk/research-briefings/cbp-8456/

Fry, R., & Cilluffo, A. (2024, September 18). 5 facts about student loans. Pew Research Center. https://www.pewresearch.org/short-reads/2024/09/18/facts-about-student-loans/

Gerstel, N. (2011). Rethinking families and community: The color, class, and centrality of extended kin ties. Sociological Forum, 26(1), 1–20. https://doi.org/10.1111/j.1573-7861.2010.01222.x

Gilligan, M., Stocker, C. M., & Jewsbury Conger, K. (2020). Sibling relationships in adulthood: Research findings and new frontiers. Journal of Family Theory & Review, 12(3), 305–320. https://doi.org/10.1111/jftr.12385

Gudmunson, C. G., & Danes, S. M. (2011). Family financial socialization: Theory and critical review. Journal of Family and Economic Issues, 32(4), 644–667. https://doi.org/10.1007/s10834-011-9275-y

Guzman, G., Kollar, M., & U.S. Census Bureau. (2023). Income in the United States: 2022-current population reports (Nos. P60-279; Current Population Reports). U.S. Census Bureau. https://www.census.gov/library/visualizations/2023/demo/p60-279.html

Hank, K., & Steinbach, A. (2018). Intergenerational solidarity and intragenerational relations between adult siblings. Social Science Research, 76, 55–64. https://doi.org/10.1016/j.ssresearch.2018.08.003

Hao, L. (2004). Wealth of immigrant and native-born Americans. International Migration Review, 38(2), 518–546. https://doi.org/10.1111/j.1747-7379.2004.tb00208.x

Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). The elements of statistical learning: Data mining, inference, and prediction (Second edition). Springer.

Heath, S., & Calvert, E. (2013). Gifts, loans and intergenerational support for young adults. Sociology, 47(6), 1120–1135. https://doi.org/10.1177/0038038512455736

Houle, J. N., & Berger, L. (2015). Is student loan debt discouraging homeownership among young adults? Social Service Review, 89(4), 589–621. https://doi.org/10.1086/684587

Huo, M., Kim, K., Zarit, S. H., & Fingerman, K. L. (2018). Support grandparents give to their adult grandchildren. The Journals of Gerontology: Series B, 73(6), 1006–1015. https://doi.org/10.1093/geronb/gbw208

Kassenboehmer, S. C., & Sinning, M. G. (2014). Distributional changes in the gender wage gap. ILR Review, 67(2), 335–361. https://doi.org/10.1177/001979391406700203

Kim, J., Choi, S., Chatterjee, S., & Kim, H. (2012). The motivation for intergenerational time and financial transfers. Family and Consumer Sciences Research Journal, 40(4), 373–397. https://doi.org/10.1111/j.1552-3934.2012.02118.x

Klaus, D., & Baykara-Krumme, H. (2017). Die Lebenssituationen von Personen in der zweiten Lebenshälfte mit und ohne Migrationshintergrund. In K. Mahne, J. K. Wolff, J. Simonson, & C. Tesch-Römer (Eds.), Altern im Wandel: Zwei Jahrzehnte Deutscher Alterssurvey (DEAS) (pp. 359–379). Springer Fachmedien. https://doi.org/10.1007/978-3-658-12502-8_24

Kuperberg, A. (2023). Families, relationships and paying for college. Journal of Student Financial Aid, 52(2). https://doi.org/10.55504/0884-9153.1768

LeBaron, A. B. (2019). The socialization of financial giving: A multigenerational exploration. Journal of Family and Economic Issues, 40(4), 633–646. https://doi.org/10.1007/s10834-019-09629-z

Lee, S., & Persson, P. (2016). Financing from family and friends. The Review of Financial Studies, 29(9), 2341–2386. https://doi.org/10.1093/rfs/hhw031

Lennartsson, C., Silverstein, M., & Fritzell, J. (2010). Time-for-money exchanges between older and younger generations in Swedish families. Journal of Family Issues, 31(2), 189–210. https://doi.org/10.1177/0192513X09344158

Leopold, T., Becker, C. C., Buyukkececi, Z., Çineli, B., & Raab, M. (2025). KINMATRIX - Uncovering the kinship matrix (Version 1.0.0) [Dataset]. GESIS. https://doi.org/10.4232/1.14380

Leopold, T., Raab, M., Becker, C. C., Buyukkececi, Z., & Çineli, B. (2025). Mapping modern kinship networks: First results from the KINMATRIX survey. Journal of Marriage and Family, 87(2), 478–504. https://doi.org/10.1111/jomf.13049

Long, J. S., & Freese, J. (2014). Regression models for categorical dependent variables using Stata (3rd edition). Stata Press.

Looney, A., & Yannelis, C. (2024). What went wrong with federal student loans? Journal of Economic Perspectives, 38(3), 209–236. https://doi.org/10.1257/jep.38.3.209

Lusardi, A., Schneider, D. J., & Tufano, P. (2011). Financially fragile households: Evidence and implications (Working Paper No. 17072). National Bureau of Economic Research. https://doi.org/10.3386/w17072

Malik, R. (2019). Working families are spending big money on child care. https://www.americanprogress.org/wp-content/uploads/sites/2/2019/06/Working-Families-SpendingBRIEF.pdf

McGarry, K. (2016). Dynamic aspects of family transfers. Journal of Public Economics, 137, 1–13. https://doi.org/10.1016/j.jpubeco.2016.03.008

McGarry, K., & Schoeni, R. F. (1994). Transfer behavior: Measurement and the redistribution of resources within the family (Working Paper No. 4607). National Bureau of Economic Research. https://doi.org/10.3386/w4607

McGarry, K., & Schoeni, R. F. (1997). Transfer behavior within the family: Results from the asset and health dynamics study. The Journals of Gerontology: Series B, 52B(Special_Issue), 82–92. https://doi.org/10.1093/geronb/52B.Special_Issue.82

Mezza, A., Ringo, D., Sherlund, S., & Sommer, K. (2020). Student loans and homeownership. Journal of Labor Economics, 38(1), 215–260. https://doi.org/10.1086/704609

Michalski, R. L., & Shackelford, T. K. (2005). Grandparental investment as a function of relational uncertainty and emotional closeness with parents. Human Nature, 16(3), 293–305. https://doi.org/10.1007/s12110-005-1012-5

Motel, A., & Szydlik, M. (1999). Private Transfers zwischen den Generationen. Zeitschrift Für Soziologie, 28(1), 3–22. https://doi.org/10.1515/zfsoz-1999-0101

Mulder, C. H., & van der Meer, M. J. (2009). Geographical distances and support from family members. Population, Space and Place, 15(4), 381–399. https://doi.org/10.1002/psp.557

Niehues, J., & Stockhausen, M. (2020). Vermögensgrenzen: Große gruppenspezifische Unterschiede. Niehues IW-Kurzbericht, (105). https://www.iwkoeln.de/studien/judith-niehues-maximilian-stockhausen-grosse-gruppenspezifische-unterschiede-487547.html

Noller, P. (2005). Sibling relationships in adolescence: Learning and growing together. Personal Relationships, 12(1), 1–22. https://doi.org/10.1111/j.1350-4126.2005.00099.x

Richard, P., Walker, R., & Alexandre, P. (2018). The burden of out of pocket costs and medical debt faced by households with chronic health conditions in the United States. PLOS ONE, 13(6), e0199598. https://doi.org/10.1371/journal.pone.0199598

Roksa, J. (2019). Intergenerational exchange of support in low-income families: Understanding resource dilution and increased contribution. Journal of Marriage and Family, 81(3), 601–615. https://doi.org/10.1111/jomf.12558

Ruel, E., & Hauser, R. M. (2013). Explaining the gender wealth gap. Demography, 50(4), 1155–1176. https://doi.org/10.1007/s13524-012-0182-0

Seginer, R. (1998). Adolescents' perceptions of relationships with older sibling in the context of other close relationships. Journal of Research on Adolescence, 8(3), 287–308. https://doi.org/10.1207/s15327795jra0803_1

Shim, S., Barber, B. L., Card, N. A., Xiao, J. J., & Serido, J. (2010). Financial socialization of first-year college students: The roles of parents, work, and education. Journal of Youth and Adolescence, 39(12), 1457–1470. https://doi.org/10.1007/s10964-009-9432-x
Sierminska, E. M., Frick, J. R., & Grabka, M. M. (2010). Examining the gender wealth gap. Oxford Economic Papers, 62(4), 669–690. https://doi.org/10.1093/oep/gpq007
Silverstein, M., & Bengtson, V. L. (1997). Intergenerational solidarity and the structure of adult child–parent relationships in American families. American Journal of Sociology, 103(2), 429–460. https://doi.org/10.1086/231213
Spitze, G., & Trent, K. (2006). Gender differences in adult sibling relations in two-child families. Journal of Marriage and Family, 68(4), 977–992. https://doi.org/10.1111/j.1741-3737.2006.00308.x
Staub, E., & Vollhardt, J. (2008). Altruism born of suffering: The roots of caring and helping after victimization and other trauma. American Journal of Orthopsychiatry, 78(3), 267–280. https://doi.org/10.1037/a0014223
Suitor, J. J., Pillemer, K., & Sechrist, J. (2006). Within-family differences in mothers' support to adult children. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 61(1), S10–S17. https://doi.org/10.1093/geronb/61.1.S10
Swartz, T. T., Kim, M., Uno, M., Mortimer, J., & O'Brien, K. B. (2011). Safety nets and scaffolds: Parental support in the transition to adulthood. Journal of Marriage and Family, 73(2), 414–429. https://doi.org/10.1111/j.1741-3737.2010.00815.x
Szydlik, M. (2016). Sharing lives: Adult children and parents. Routledge.
Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society Series B: Statistical Methodology, 58(1), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x
Voorpostel, M., & Blieszner, R. (2008). Intergenerational solidarity and support between adult siblings. Journal of Marriage and Family, 70(1), 157–167. https://doi.org/10.1111/j.1741-3737.2007.00468.x
Voorpostel, M., van der Lippe, T., Dykstra, P. A., & Flap, H. (2007). Similar or different? The importance of similarities and differences for support between siblings. Journal of Family Issues, 28(8), 1026–1053. https://doi.org/10.1177/0192513X07300713
Weaver, S. E., Coleman, M., & Ganong, L. H. (2003). The sibling relationship in young adulthood: Sibling functions and relationship perceptions as influenced by sibling pair composition. Journal of Family Issues, 24(2), 245–263. https://doi.org/10.1177/0192513X02250098
White, L. K. (1994). Coresidence and leaving home: Young adults and their parents. Annual Review of Sociology, 20(1), 81–102. https://doi.org/10.1146/annurev.so.20.080194.000501
White, L. K. (2001). Sibling relationships over the life course: A panel analysis. Journal of Marriage and Family, 63(2), 555–568. https://doi.org/10.1111/j.1741-3737.2001.00555.x
White, L. K., & Riedmann, A. (1992). Ties among adult siblings. Social Forces, 71(1), 85–102.
Whiteman, S. D., McHale, S. M., & Soli, A. (2011). Theoretical perspectives on sibling relationships. Journal of Family Theory & Review, 3(2), 124–139. https://doi.org/10.1111/j.1756-2589.2011.00087.x
Zhao, H., & Zhang, L. (2020). Talking money at home: The value of family financial socialization. International Journal of Bank Marketing, 38(7), 1617–1634. https://doi.org/10.1108/IJBM-04-2020-0174

Zou, H. (2006). The adaptive lasso and its oracle properties. Journal of the American Statistical Association, 101(476), 1418–1429. https://doi.org/10.1198/016214506000000735

Zou, H., Hastie, T., & Tibshirani, R. (2007). On the “degrees of freedom” of the lasso. https://doi.org/10.48550/ARXIV.0712.0881

7. Appendix

Appendix IV-A: Sample Selection Process

Removal Step	Respondents Remain	Sibling Dyads Remain
Full US-Sample	5,005	7,945
Restriction to Respondents with at Least One Sibling	4,025	7,945
Exclusion of Deceased Siblings or Unknown Vital Status	3,986	7,709
Exclusion of Siblings with Unknown Names	3,928	7,539
Exclusion of Siblings Based on Age Criteria	3,891	7,327
Exclusion of Siblings with Missing Values	3,124	5,556
Final Analytical Sample (N)	3,124	5,556

Notes: This table displays the sequential exclusion criteria applied to the initial US sample to derive the final analytical sample. The first column indicates the removal step, with Columns 2 and 3 indicating the number of respondents and sibling dyads, respectively, remaining in the sample after each step.

Appendix IV-B: Summary Statistics

	N	Mean	SD	Min	Median	Max
Support Indicators						
Any Sibling Financial Support	3,124	0.16	0.37	0	0	1
Financial Support (Specific Sibling)	5,556	0.11	0.31	0	0	1
Respondent						
Female	3,124	0.49	0.50	0	0	1
Subjective Social Status (SSS)	3,124	4.29	2.18	0	4	9
Age	3,124	30.37	3.18	25	31	35
Bad Health	3,124	0.30	0.46	0	0	1
Sibling						
Female	5,556	0.51	0.50	0	1	1
SSS	5,556	4.78	2.30	0	5	9
Age	5,556	30.41	7.11	14	30	68
Bad Health	5,556	0.17	0.37	0	0	1
Relationship						
Sister-Sister Dyad	5,556	0.25	0.43	0	0	1
Sibling SSS > Respondent SSS	5,556	0.43	0.49	0	0	1
Sibling Older than Respondent	5,556	0.48	0.50	0	0	1
Distance > 1 Hour	5,556	0.44	0.50	0	0	1
High Emotional Closeness	5,556	0.69	0.46	0	1	1
Monthly Contact	5,556	0.83	0.38	0	1	1
Family						
Sibship Size	3,124	3.10	1.46	2	3	13
Respondent has Children	3,124	0.41	0.49	0	0	1
Parental Support Received	3,124	0.61	0.49	0	1	1
Immigration Background	3,124	0.19	0.40	0	0	1

Notes: This table reports descriptive statistics for the analytical sample. Variables regarding the respondent and family context are reported at the respondent level (N=3,124), whereas variables regarding the sibling and relationship are reported at the dyad level (N=5,556). $Any\ Sibling\ Financial\ Support$ indicates if the respondent received financial support from any sibling, whereas $Financial\ Support\ (Specific\ Sibling)$ indicates if a specific sibling provided financial support within a dyad.

Appendix IV-C: LPM Regressions Predicting the Receipt of Financial Support from a Sibling

	M1	M2	M3	M4	M5
Respondent					
Female	-0.0106				-0.0179
	(0.0132)				(0.0170)
Subjective Social Status (SSS)	0.0059				-0.0075
	(0.0036)				(0.0051)
Age	-0.0012				-0.0056*
	(0.0022)				(0.0027)
Bad Health	0.0110				0.0177
	(0.0166)				(0.0158)
Sibling					
Female		-0.0231*			-0.0191
		(0.0108)			(0.0167)
SSS		0.0198***			0.0192***
		(0.0030)			(0.0052)
Age		0.0025**			0.0031*
		(0.0008)			(0.0015)
Bad Health		0.0376			0.0415*
		(0.0207)			(0.0203)
Relationship					
Sister-Sister Dyad			-0.0219		-0.0071
			(0.0119)		(0.0214)
Sibling SSS > Respondent SSS			0.0370**		-0.0031
			(0.0133)		(0.0196)
Sibling Older than Respondent			0.0470***		0.0082
			(0.0120)		(0.0198)
Distance > 1 Hour			-0.0136		-0.0172
			(0.0122)		(0.0117)
High Emotional Closeness			0.0480**		0.0454**
			(0.0146)		(0.0146)
Monthly Contact			0.0483***		0.0309*
			(0.0141)		(0.0138)
Family					
Sibship Size				-0.0019	0.0020
				(0.0032)	(0.0034)
Respondent has Children				0.0360**	0.0488***
				(0.0134)	(0.0138)
Parental Support Received				0.0892***	0.0807***
				(0.0123)	(0.0119)
Immigration Background				0.0624***	0.0558**
				(0.0179)	(0.0174)
Constant	0.1197	-0.0590*	0.0057	0.0317*	-0.0097
	(0.0685)	(0.0301)	(0.0146)	(0.0155)	(0.0752)
Observations (Sibling Dyads)	5,556	5,556	5,556	5,556	5,556
R^2	0.0022	0.0292	0.0238	0.0310	0.0695
F-Statistic	0.85	13.01***	11.66***	18.29***	7.29***

Notes: This table reports results from LPM regressions predicting the receipt of financial support. Models 1–4 include respondent, sibling, relationship, and family characteristics separately, whereas Model 5 includes all covariates simultaneously. Robust standard errors clustered at the respondent level are reported in parentheses. Significance levels are denoted as follows: $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

Appendix IV-D: LPM Regressions Predicting the Receipt of Financial Support from a Sibling by Gender

	Respondent		Sibling	
	Male (M6a)	Female (M6b)	Male (M7a)	Female (M7b)
Respondent				
Female	-	-	-0.0195	-0.0216
			(0.0170)	(0.0167)
Subjective Social Status (SSS)	-0.0029	-0.0133*	-0.0035	-0.0110
	(0.0071)	(0.0062)	(0.0068)	(0.0058)
Age	-0.0070	-0.0034	-0.0037	-0.0075*
	(0.0038)	(0.0033)	(0.0035)	(0.0033)
Bad Health	0.0042	0.0322	0.0258	0.0079
	(0.0264)	(0.0167)	(0.0214)	(0.0185)
Sibling				
Female	-0.0141	-0.0266*	-	-
	(0.0165)	(0.0130)		
SSS	0.0217**	0.0159**	0.0219**	0.0164**
	(0.0080)	(0.0058)	(0.0071)	(0.0060)
Age	0.0034	0.0031	0.0021	0.0043
	(0.0023)	(0.0016)	(0.0017)	(0.0022)
Bad Health	0.0862*	-0.0011	0.0224	0.0569*
	(0.0365)	(0.0176)	(0.0225)	(0.0285)
Relationship				
Sibling SSS > Respondent SSS	-0.0028	-0.0018	0.0045	-0.0111
	(0.0310)	(0.0211)	(0.0262)	(0.0243)
Sibling Older than Respondent	0.0280	-0.0098	0.0101	-0.0002
	(0.0298)	(0.0233)	(0.0258)	(0.0278)
Distance > 1 Hour	-0.0446**	0.0110	-0.0159	-0.0180
	(0.0173)	(0.0145)	(0.0161)	(0.0151)
High Emotional Closeness	0.0138	0.0843***	0.0446*	0.0463*
	(0.0226)	(0.0153)	(0.0213)	(0.0191)
Monthly Contact	0.0344	0.0179	0.0248	0.0367
	(0.0199)	(0.0163)	(0.0201)	(0.0194)
Family				
Sibship Size	0.0051	-0.0005	-0.0014	0.0064
	(0.0048)	(0.0046)	(0.0037)	(0.0046)
Respondent has Children	0.0619**	0.0284	0.0616**	0.0345*
	(0.0218)	(0.0159)	(0.0193)	(0.0158)
Parental Support Received	0.0762***	0.0841***	0.0815***	0.0821***
	(0.0188)	(0.0133)	(0.0165)	(0.0144)
Immigration Background	0.0813**	0.0332	0.0852***	0.0230
	(0.0286)	(0.0181)	(0.0254)	(0.0210)
Constant	-0.0102	-0.0497	-0.0614	0.0182
	(0.1063)	(0.0979)	(0.0970)	(0.0928)
Observations (Sibling Dyads)	2,302	3,254	2,760	2,796
R^2	0.0939	0.0667	0.0806	0.0659
F-Statistic	4.12***	8.21***	6.33***	5.06***

Notes: This table reports results from LPM regressions predicting the receipt of financial support stratified by gender. Models 6a and 6b provide estimates for male and female respondents separately, and Models 7a and 7b provide estimates for brothers and sisters separately. Robust standard errors clustered at the respondent level are reported in parentheses. Significance levels are denoted as follows: *p < .05; **p < .01; ***p < .001.

Appendix IV-E: Logistic Regressions Predicting the Receipt of Financial Support from a Sibling

	M1	M2	M3	M4	M5
Respondent					
Female	0.8948				0.8338
	(0.1234)				(0.1465)
Subjective Social Status (SSS)	1.0637				0.9220
	(0.0404)				(0.0433)
Age	0.9873				0.9409*
	(0.0222)				(0.0273)
Bad Health	1.1226				1.2519
	(0.1948)				(0.2069)
Sibling					
Female		0.7789*			0.8235
		(0.0919)			(0.1511)
SSS		1.2450***			1.2353***
		(0.0413)			(0.0608)
Age		1.0262**			1.0341*
		(0.0084)			(0.0146)
Bad Health		1.4438			1.5193*
		(0.2810)			(0.2995)
Relationship					
Sister-Sister Dyad			0.7756		0.8792
			(0.1037)		(0.2141)
Sibling SSS > Respondent SSS			1.4777**		0.9968
			(0.2038)		(0.2050)
Sibling Older than Respondent			1.6605***		1.1476
			(0.2159)		(0.2404)
Distance > 1 Hour			0.8689		0.8363
			(0.1135)		(0.1092)
High Emotional Closeness			1.7639**		1.7740**
			(0.3500)		(0.3611)
Monthly Contact			2.2114**		1.8888*
			(0.5572)		(0.4790)
Family					
Sibship Size				0.9775	1.0042
				(0.0392)	(0.0423)
Respondent has Children				1.4977**	1.7753***
				(0.2145)	(0.2606)
Parental Support Received				3.0797***	2.9771***
				(0.5428)	(0.5269)
Immigration Background				1.8322***	1.7670***
				(0.2795)	(0.2735)
Observations (Sibling Dyads)	5,556	5,556	5,556	5,556	5,556
AIC	3660.30	3512.13	3534.79	3494.09	3290.00
BIC	3693.41	3545.24	3581.15	3527.21	3415.83

Notes: This table reports odds ratios from logistic regression models predicting the receipt of financial support. Models 1–4 include respondent, sibling, relationship, and family characteristics separately, whereas Model 5 includes all covariates simultaneously. Robust standard errors clustered at the respondent level are reported in parentheses. Significance levels are denoted as follows: $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

Appendix IV-F: Logistic Regressions Predicting the Receipt of Financial Support from a Sibling by Gender

	Respondent		Sibling	
	Male (M6a)	Female (M6b)	Male (M7a)	Female (M7b)
Respondent				
Female	-	-	0.8248	0.7375
			(0.1434)	(0.1496)
Subjective Social Status (SSS)	0.9744	0.8538*	0.9640	0.8827*
	(0.0582)	(0.0563)	(0.0634)	(0.0477)
Age	0.9318	0.9439	0.9664	0.9163*
	(0.0384)	(0.0350)	(0.0355)	(0.0341)
Bad Health	1.0727	1.4612*	1.3629	1.1282
	(0.3086)	(0.2652)	(0.2780)	(0.2463)
Sibling				
Female	0.8560	0.7225*	-	-
	(0.1565)	(0.1097)		
SSS	1.2570**	1.2174**	1.2538***	1.2063**
	(0.0893)	(0.0744)	(0.0851)	(0.0737)
Age	1.0316	1.0442*	1.0254	1.0419*
	(0.0211)	(0.0190)	(0.0173)	(0.0214)
Bad Health	2.2518**	0.9355	1.2362	1.8031*
	(0.6308)	(0.2193)	(0.2899)	(0.4975)
Relationship				
Sibling SSS > Respondent SSS	1.0212	0.9574	1.0790	0.9220
	(0.3209)	(0.2329)	(0.2855)	(0.2553)
Sibling Older than Respondent	1.5039	0.8523	1.1352	1.1273
	(0.4735)	(0.2178)	(0.2984)	(0.3469)
Distance > 1 Hour	0.6077*	1.1638	0.8246	0.8428
	(0.1191)	(0.1927)	(0.1380)	(0.1546)
High Emotional Closeness	1.1686	3.5856***	1.5844	2.0563*
	(0.3114)	(0.9375)	(0.4167)	(0.6148)
Monthly Contact	2.0546*	1.3960	1.8756	1.9207
	(0.6898)	(0.4981)	(0.6268)	(0.7595)
Family				
Sibship Size	1.0380	0.9845	0.9628	1.0599
	(0.0631)	(0.0582)	(0.0448)	(0.0588)
Respondent has Children	2.1208***	1.3838	1.9066***	1.5869*
	(0.4763)	(0.2485)	(0.3549)	(0.3067)
Parental Support Received	2.7056***	3.3286***	2.6929***	3.4731***
	(0.7056)	(0.7303)	(0.6050)	(0.8317)
Immigration Background	2.1422**	1.4319*	2.1087***	1.3831
	(0.5036)	(0.2511)	(0.4061)	(0.3075)
Observations (Sibling Dyads)	2,302	3,254	2,760	2,796
AIC	1719.62	1532.53	1729.57	1566.98
BIC	1817.22	1636.02	1830.26	1667.89

Notes: This table reports odds ratios from logistic regressions predicting the receipt of financial support stratified by gender. Models 6a and 6b provide estimates for male and female respondents separately, and Models 7a and 7b provide estimates for brothers and sisters separately. Robust standard errors clustered at the respondent level are reported in parentheses. Significance levels are denoted as follows: $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

V. Perception vs. Reality: Wealth Disparities Between Urban and Rural Households in Germany

Co-authors:

Tristan Maier, Andreas Walter

Own share:

30%

This article has received a Revise and Resubmit from:

Regional Studies

Perception vs. Reality:
Wealth Disparities Between Urban and Rural Households in Germany

Martin G. Becker[a] Tristan Maier[b] Andreas Walter[c]

Abstract: This study shows key differences in both actual and perceived net wealth between urban and rural households in Germany. Contrary to the narrative of rural economic disadvantage, rural households have not less but rather more net wealth than urban households. However, rural households underestimate their relative wealth position by approximately one decile more than urban households, reflecting a profound "left behind" feeling. About two-thirds of the initial underestimation can be explained by ownership of low-salience real estate. Yet a significant rural underestimation persists, emphasizing that a feeling of being "left behind" can occur in direct contradiction to reality.

JEL Codes: D14, D31, D91, G51, R12

Keywords: household finance, "left behind" places, regional inequalities, wealth perception, salience bias

[a] martin.becker@wirtschaft.uni-giessen.de, Justus Liebig University Gießen, Department of Financial Services, Licher Straße 74, 35394 Gießen, Germany

[b] tristan.maier@uni-giessen.de, Justus Liebig University Gießen, Department of Financial Services, Licher Straße 74, 35394 Gießen, Germany

[c] andreas.walter@wirtschaft.uni-giessen.de, Justus Liebig University Gießen, Department of Financial Services, Licher Straße 74, 35394 Gießen, Germany

1. Introduction

The differences between urban and rural regions are manifold. Households in these regions differ in terms of education (Brown et al., 2010; Roscigno & Crowle, 2001), income (Boulant et al., 2016; Fuest & Immel, 2019), poverty rates (Partridge & Rickman, 2008; Weber et al., 2005), real estate ownership (Kohl, 2016; Lerbs & Oberst, 2014), health care outcomes (Sun & Monnat, 2022; Zahnd et al., 2018), mortality rates (Schwerter et al., 2023; Singh & Siahpush, 2014), life satisfaction, (Hoogerbrugge & Burger, 2024; Jurčišinová et al., 2025; Okulicz-Kozaryn, 2024) and, last but not least, their political attitudes (Gimpel et al., 2020; Johnston et al., 2016; Kenny & Luca, 2021).

Many of these differences systematically favor urban households and have become increasingly politicized. Over the past decade rural populism has gained momentum across numerous countries – driven by a widespread perception of economic disadvantage and a growing sense of being "left behind" (Deppisch, 2021; Rodríguez-Pose, 2018). In this context, both perceived and actual socioeconomic disparities have become central to understanding shifting political attitudes and the rise of rural populism (Deppisch et al., 2022; Velthuis et al., 2025). This resentment follows a clear geographic pattern and is most pronounced in rural areas facing economic decline and a feeling of marginalization (Deppisch, 2021; Rodríguez-Pose, 2018). These dynamics have become increasingly visible in recent elections.

In the UK, major cities largely supported remaining in the EU, while Brexit found strongest support in economically struggling rural areas, especially in the north and east (Essletzbichler et al., 2018; Rodríguez-Pose, 2018). This rural-urban divide was also evident in the 2016 and 2024 U.S. presidential elections, with Trump winning most rural counties[1] while Clinton and Harris drew support mainly from urbanized areas (Federal Election Commission, 2016, 2024).

[1] Trump won 85% (89%) of rural counties in 2016 (2024)

A similar link has also been observed in Germany, where support for the right-wing populist party Alternative für Deutschland (AfD) is higher in rural areas (Diermeier, 2020; Richter & Bösch, 2017), especially in rural districts of the new federal states that are facing population decline and economic difficulties (Bergmann et al., 2017). These areas are often portrayed as structurally weak and without prospects, reinforcing a stigmatizing "left behind" narrative (Rodríguez-Pose, 2018; Royer & Leibert, 2024), which the AfD actively employs (Deppisch, 2021; Goerres et al., 2018).

In academic debate, however, the notion of being "left behind" lacks a universally accepted definition and is generally understood as a multidimensional condition in which a group consistently falls behind others (Deppisch, 2021; Hertrich & Brenner, 2024; Pike et al., 2024). These dimensions can range from cultural, social, geographical, and economic to political aspects, yet the condition carries a dual meaning. It refers both to the objective reality of falling behind and the subjective perception of marginalization and comparative loss (Deppisch, 2021; MacKinnon et al., 2024; Pike et al., 2024).

Despite the concept's multidimensional nature, objective economic disparities have dominated academic and policy discourse. Among the various dimensions, economic indicators are not only more tangible and easier to quantify but are also widely seen as central to broader patterns of social and regional inequality (Hertrich & Brenner, 2024; Pike et al., 2024). As a result, many studies have operationalized "left-behindness" through measurable disparities in economic outcomes. This is most evident in research on income differences, where evidence from both international and German contexts consistently shows that rural households have significantly lower incomes than their urban counterparts (Ananian & Dellaferrera, 2024; Boulant et al., 2016; Fuest & Immel, 2019; Sicular et al., 2007).

While income provides a valuable snapshot of households' current financial flows, net wealth offers a broader, longer-term perspective of their economic position. It reflects the cumulative result of past income, consumption, cost of living, saving behavior, and investment decisions, including both assets and liabilities (Bricker et al., 2016; Tully & Sharma, 2022; Wolff & Zacharias, 2009). Wealth and income are therefore interrelated but differ in their structure and complexity, which is shown by their positive but far from perfect correlation (Balestra & Oehler, 2023; Gallusser & Krapf, 2019). Accordingly, income and net wealth are essential but distinct indicators of a household's financial condition (Gibson-Davis et al., 2022; Lamarche et al., 2020; Wolff & Zacharias, 2009).

Although research on rural-urban income differences is extensive, wealth disparities have, with a few exceptions in Norway (Galster & Wessel, 2024) and China (Hu & Gao, 2023; Wang et al., 2020; Xie & Jin, 2015), only been examined at the aggregate national level (Albers et al., 2020; Bartels & Schroeder, 2020; Grabka & Westermeier, 2014).

Beyond these objective indicators, subjective perceptions of economic standing represent the second part of the dual meaning of the "left behind" narrative, which contributes to rural populism (Rodríguez-Pose 2018; Deppisch et al. 2022). Such perceptions matter because they reveal not only how households evaluate their own position but also whether systematic misjudgments distort these evaluations. Research shows that individuals systematically misjudge their relative wealth, often underestimating their position in the distribution (Batista et al., 2023; Fessler & Rapp, 2023; Sussman & Shafir, 2012; Zagorsky, 2000). However, existing studies focus almost exclusively on national-level patterns and do not examine whether such misperceptions differ systematically between urban and rural households. If rural households consistently underestimate their position more than urban ones, this may signal a deeper sense of economic exclusion and a perceived feeling of being "left behind".

In terms of net wealth disparities, two important research gaps persist. First, it is unclear whether rural households in Germany are disadvantaged in terms of net wealth, despite their well-documented disadvantages in other domains, such as income. Second, little is known about whether urban and rural households differ in how they perceive their relative position in the wealth distribution and whether rural households in particular systematically underestimate their standing. This raises key empirical questions that remain largely unanswered: (i) Are rural households indeed “left behind” in terms of net wealth? (ii) Do rural households feel “left behind” in terms of net wealth?

We contribute to predominantly income-centered research on economic urban-rural disparities (e.g., Ananian & Dellaferrera, 2024; Boulant et al., 2016) by investigating whether rural households in Germany actually are less wealthy than urban households. In doing so, we enhance a sparse strand of literature that specifically addresses urban-rural wealth disparities. In addition, we examine how rural households perceive their relative net wealth position – particularly in comparison to their urban counterparts – and whether they themselves feel economically "left behind". Specifically, we compare net wealth perceptions and the actual net wealth distribution of urban and rural households in Germany, exploring whether rural households feel and indeed are “left behind” in terms of net wealth. In addition, we integrate insights from behavioral economics by exploring how asset salience – particularly real estate – affects potential gaps in actual and perceived wealth.

Our analysis shows, contrary to the narrative of rural economic disadvantage, that rural households in Germany hold not less but rather more net wealth than urban households. Despite this objective advantage, rural households underestimate their actual net wealth decile by about one decile more than urban households, exhibiting a profound “left behind” feeling. This stronger underestimation can largely be explained by the low salience of real estate ownership but persists even after accounting for it.

2. Data and Methodology

2.1 Sample Selection

The German Panel on Household Finances (PHF) (Schmidt et al., 2025), administered by the Research Centre of the Deutsche Bundesbank, provides nationally representative longitudinal data on private households in Germany. It gathers extensive information on financial circumstances, including assets, debt positions, income flows, and wealth components, together with a broad set of demographic variables. It also offers a detailed breakdown of household wealth, differentiating between real estate holdings and various types of financial investments, providing a comprehensive view of household financial portfolios. In addition, the dataset includes a wide range of self-reported measures, such as individuals' risk preferences and their perceived position in the national net wealth distribution.

The design of the PHF includes an oversampling of wealthy households to ensure robust representation of the top tail of the wealth distribution (Schmidt & Eisele, 2013). All analyses use the survey weights provided by the Bundesbank to correct for oversampling and ensure population representativeness. These weights also account for non-response and sampling variability, thereby enhancing the reliability and precision of parameter estimates.

Missing data points were imputed by the Deutsche Bundesbank, using the methodology proposed by Rubin (1987) to consider within-imputation variance and between-imputation variance. Furthermore, we use the replication weights included in the dataset to adjust standard errors for both imputation variance and the complex survey design. [2]

Although the PHF is designed as a household-level survey, the questionnaire is primarily answered by the Financially Knowledgeable Person (FKP). The FKP is the household member designated as having the greatest knowledge of the household's finances. Because financial

[2] For comprehensive information on the survey design, data collection procedures, and weighting methodology, see von Kalckreuth et al. (2012).

information in the PHF is collected at the household level, the FKP's demographic characteristics are linked to these household-level financial records for the analyses.

We use data from the most recent survey wave conducted in 2023, which initially contains 3,985 households. To maintain a robust rural–urban contrast, we first exclude households residing in municipalities with 20,000–100,000 inhabitants. The sample is further restricted to households with available information on municipality size.[3] We also exclude observations with missing information on any of the variables relevant to our study, resulting in a final sample of 2,777 households.

2.2 Variable Measurement

Our main variable of interest captures the geographical classification of households based on their degree of urbanization. We construct a binary $Rural\ Household\ Indicator$, which equals one if the household resides in a rural municipality with a population less than 20,000 inhabitants, and zero otherwise. The classification follows the criteria established by the German Federal Office for Building and Regional Planning (2025). Municipalities with fewer than 20,000 inhabitants are categorized as rural (including small towns and rural areas), whereas those with more than 100,000 inhabitants are defined as urban (medium-sized and major cities).[4]

The second set of variables comprises measures related to the discrepancy between actual and perceived net wealth. First, it includes the total household $Net\ Wealth$ in thousands of euros (k€), calculated as total assets minus total liabilities.[5]

[3] The Deutsche Bundesbank's anonymization rules led to the removal of 116 households in small eastern municipalities (PHF Survey Team, 2025).

[4] As of year-end 2023, 83 German cities had more than 100,000 inhabitants (Statista ,2024).

[5] Total assets include both real assets, such as real estate and valuables, and financial assets, such as deposits and mutual funds. Total liabilities comprise mortgage and non-mortgage debt, such as credit card debt and private loans. A detailed description of the components of net wealth is provided in Appendix V-A.

The set further contains three closely related measures of a household's position in the wealth distribution. The first is the actual $Net\ Wealth\ Decile$, an ordinal variable that assigns each household to one of ten deciles based on their total net wealth in the empirical wealth distribution across the entire dataset, prior to any exclusions. The second is the $Perceived\ Net\ Wealth\ Decile$, a self-assessed measure of relative wealth. Specifically, the household's FKP was asked to place their household within Germany's net wealth distribution on a scale from 1 (bottom 10%) to 10 (top 10%).

The third measure is $Net\ Wealth\ Decile\ Misperception$, defined as the difference between the perceived and actual net wealth decile. Positive values indicate an overestimation of relative wealth, while negative values reflect an underestimation. This variable quantifies both the magnitude and direction of misperceptions of net wealth across households. For instance, consider a household with a net wealth of €300,000 that places it in the 7th decile of the German net wealth distribution. If the household's FKP perceives itself to be in the 5th decile, the resulting net wealth misperception value is "–2", indicating an underestimation of its relative wealth.

We further include a set of demographic control variables. The first part of the demographic control variables refers solely to the FKP, while household size, income, real estate ownership, and the region indicator are measured at the household level.[6]

2.3 Descriptive Statistics

Table V-1 displays the descriptive statistics for the variables outlined in Section 2.2. The $Rural\ Household\ Indicator$ in panel A shows that 52.39% of households in the sample reside in rural areas. Specifically, 17.31% live in municipalities with fewer than 5,000 inhabitants, and 35.08% live in small towns with 5,000 to 20,000 inhabitants. The remaining 47.61% of households are located in urban areas, with 22.56% living in medium-sized (100,000 to 500,000

[6] Appendix V-A provides a detailed description of all control variables.

inhabitants) and 25.05% in major cities with populations exceeding 500,000.[7] Since municipalities with 20,000 to 100,000 inhabitants were excluded from the analysis due to their ambiguous classification as neither clearly rural nor urban, no households in our sample reside in such areas.

Table V-1: Descriptive Statistics

Variable	Mean	25%-Quantile	Median	75%-Quantile	N
Panel A: Municipality Size					
Rural Household Indicator	52.39%	0.00	1.00	1.00	2,777
Municipality Size < 5k	17.31%	0.00	0.00	0.00	2,777
Municipality Size 5k – 20k	35.08%	0.00	0.00	1.00	2,777
Municipality Size 20k – 100k	0.00%	0.00	0.00	0.00	2,777
Municipality Size 100k – 500k	22.56%	0.00	0.00	0.00	2,777
Municipality Size > 500k	25.05%	0.00	0.00	1.00	2,777
Panel B: Actual vs. Perceived Wealth					
Net Wealth (k€)	325.27	10.72	101.92	392.38	2,777
Net Wealth Decile	5.48	3.00	5.00	8.00	2,777
Perceived Net Wealth Decile	4.29	3.00	4.00	5.00	2,777
Net Wealth Decile Misperception	-1.19	-3.00	-1.00	0.00	2,777
Panel C: Demographics					
Age	53.31	38.00	53.00	67.00	2,777
Male	54.54%	0.00	1.00	1.00	2,777
University Degree	28.10%	0.00	0.00	1.00	2,777
Financial Literacy	2.48	2.00	3.00	3.00	2,777
Household Size	1.98	1.00	2.00	2.00	2,777
Income (k€)	71.85	29.65	49.04	82.65	2,777
Region East	22.13%	0.00	0.00	0.00	2,777
Real Estate Ownership	48.01%	0.00	0.00	1.00	2,777

Notes: This table presents descriptive statistics for our sample. Definitions of all variables are provided in Appendix V-A. Means and quantiles are computed using survey weights and combined across multiple imputations following Rubin's rules.

Panel B reports descriptives for the set of actual versus perceived net wealth variables. Average household net wealth amounts to approximately €325,000 while the median is about €102,000. This indicates that households' net wealth is highly right-skewed, aligning with existing evidence that a minority of households possess the majority of net wealth (Albers et al., 2020; Benhabib & Bisin, 2018). Although the average actual net wealth corresponds to the 5.48th

[7] This distribution aligns with a report of the German Federal Statistical Office (2021), which excluding municipalities with 20,000 to 100,000 inhabitants – shows that 19.2% of the population live in municipalities with fewer than 5,000 inhabitants, 36.7% in small towns (5,000 to 20,000), 20.9% in medium-sized cities (100,000 to 500,000), and 23.4% in major cities (500,000 or more).

decile[8], households perceive themselves, on average, to be at 4.29 on the decile scale. On average, households across the sample underestimate their actual net wealth by 1.19 deciles. This systematic underestimation aligns with earlier evidence from Zagorsky (2000) showing that for every dollar owned, the average individual believes they only hold 0.65 cents per dollar. It is also in line with the findings of Batista et al. (2023), demonstrating that individuals rate others as better off than themselves even under identical financial conditions.

Panel C summarizes the demographics of the FKP and household. The average FKP is 53.31 years old, and FKPs are slightly more often males (54.54%) than females (45.46%). The mean of 1.98 household members and the 75th percentile of 2.00 indicate that a large part of the German population either lives alone or with one other person. Moreover, 28.10% of FKPs in our sample have a university degree, and their average score on the Lusardi and Mitchell (2008) big three financial literacy test is 2.48.[9] The median indicates that more than 50% were able to correctly answer all three questions.

Household income in our sample averages around €72,000, compared to €63,000 reported by the German Federal Statistical Office (2024b). This difference likely reflects the PHF's highly detailed interview process, which records each income component separately and uses specific object lists for every account and portfolio.

22.13% of households in our sample reside in East Germany, slightly above the official share of 19.35% reported by the German Federal Statistical Office (2024a). Considering the East-West divide in Germany is essential when examining urban–rural wealth disparities. Historically, East German households have been economically worse off compared to those in

[8] Deciles are based on the full-sample net wealth distribution and then applied to the subset. Under a uniform 10% split, the expected mean decile is 5.5 (i.e., the average of deciles 1 to 10), assuming the subset does not differ systematically from the full sample.

[9] In the fifth PHF wave, the set of financial literacy questions was modified such that only half of the respondents received the original "Big Three" questions, while the other half were presented with a revised version of the diversification question. Restricting the sample to participants who completed the original version increases the mean number of correct responses to 2.71.

the West, facing lower wages (Smolny & Kirbach, 2011), higher unemployment rates (Falk et al., 2011), and lower net wealth (Kasinger et al., 2023). These persistent disparities continue to impact economic outcomes today and are often accompanied by a perception of being "left behind," which has also been linked to higher levels of political radicalization (Hilmar, 2022).

In our sample, 48.01% of households own property, which is almost identical to the official German homeownership-rate of 47.6% reported by Eurostat (2025). Since real estate represents the largest component of household portfolios (Arrondel et al., 2018; Flavin & Yamashita, 2002), ownership differences are particularly relevant for understanding wealth disparities and potential misperceptions of economic standing. Consistent with previous evidence (Kohl, 2016; Lerbs & Oberst, 2014), ownership rates differ markedly between urban and rural households in our sample: 62.98% of rural households own property compared to only 31.54% in urban areas (see Appendix V-B).

Table V-2 displays the correlation matrix for all control variables and the $Rural\ Household\ Indicator$. Net wealth and real estate ownership show the highest correlation (0.42), while all other correlations are substantially lower, indicating moderate associations but no evidence of severe multicollinearity.

Table V-2: Correlation Matrix

Variable	Rural Household Indicator	Age	Male	University Degree	Financial Literacy	Household Size	Household Income	Net Wealth	Region East	Real Estate Ownership
Rural Household Indicator	1.00***									
Age	0.08***	1.00***								
Male	0.05	-0.04	1.00***							
University Degree	-0.12***	-0.12***	0.06*	1.00***						
Financial Literacy	-0.04	-0.12***	0.07**	0.20***	1.00***					
Household Size	0.16***	-0.21***	0.02	0.02	0.06*	1.00***				
Income	0.05*	-0.06*	0.08**	0.17***	0.11***	0.25***	1.00***			
Net Wealth	0.07**	0.12***	0.08***	0.16***	0.09***	0.15***	0.34***	1.00***		
Region East	-0.02	0.04	-0.04	-0.04	-0.12***	0.01	-0.07**	-0.11***	1.00***	
Real Estate Ownership	0.31***	0.23***	0.11***	0.11***	0.12***	0.21***	0.23***	0.42***	-0.12***	1.00***

Notes: This table displays pairwise correlations between all control variables and the $Rural\ Household\ Indicator$. Correlations are derived using the survey weights and combined across multiple imputations according to Rubin's rules

2.4 Methodology

To address the research questions outlined above, we estimate a series of ordinary least squares (OLS) regression models. The general model specification is as follows:

$$Y_i = \beta_0 + \beta_1 Rural\ Household\ Indicator_i + \gamma c_i + \varepsilon_i \tag{V-1}$$

On the left-hand side, Y_i denotes the outcome of interest observed for the $Household_i$, specifically net wealth (in thousands) and the misperception between the actual and perceived decile. On the right-hand side, β_0 represents the constant term, while β_1 measures the average difference in the outcome between rural and urban households. Across all analyses, the $Rural\ Household$ indicator serves as a key explanatory variable of interest.

We present the regression models in up to three columns. Column (1) reports the bivariate relationship between the $Rural\ Household\ Indicator$ and the dependent variable. Column (2) adds demographic controls c_i, which include all FKP and household characteristics from Panel C of Table V-1, unless otherwise specified. Net wealth is included as an additional control in our second model, where it is not the dependent variable. The $Real\ Estate\ Ownership$ indicator is only added separately in Column (3) when analyzing net wealth decile misperception. ε_i denotes the error term that captures unobserved factors.

We estimate all models using the survey weights, and we pool coefficients with robust standard errors using replication weights according to Rubin's rules (Rubin, 1987). This approach accounts both for the complex survey design and the additional variability from multiple imputations.

3. Results

3.1 Objective Gap: Are Rural Households Indeed "Left Behind"?

We first examine whether rural households are indeed "left behind" in terms of net wealth compared to urban households, thereby putting the widespread narrative of rural economic disadvantage to an empirical test. We use household net wealth measured in thousands of euros (k€) as the dependent variable. Table V-3 reports the results, with all coefficients likewise denoted in k€.

Table V-3: Linear Regression on Net Wealth (k)

	Dependent Variable	
	(1) Net Wealth	(2) Net Wealth
Rural Household Indicator	98.5660**	61.8065
	(39.6428)	(44.3146)
Age		6.9744***
		(0.7504)
Male		89.9529**
		(34.9098)
University Degree		252.1451***
		(53.0698)
Financial Literacy		48.2705
		(41.6530)
Household Size		104.4698***
		(24.1672)
Region East		-174.2012***
		(44.8239)
Constant	273.6347***	-486.4794***
	(31.4042)	(95.0179)
Observations	2,777	2,777
R^2	0.0054	0.1030
F-Statistic	6.1820**	28.1920***

Notes: This table reports the results of our regressions on household $Net\ Wealth$, measured in thousands of euros. In column (1), $Net\ Wealth$ is regressed solely on the $Rural\ Household\ Indicator$. In column (2), $Net\ Wealth$ is regressed on the $Rural\ Household\ Indicator$ alongside a set of demographic control variables, which include Age, $Gender$, $University\ Degree$, $Financial\ Literacy$, $Household\ Size$, and $Region\ East$. Standard errors (in parentheses) are adjusted for imputation error following Rubin's Rules as well as for survey design using replicate weights. The data are weighted according to the sampling weights provided in the dataset. Significance levels are indicated as follows: *p < .1; **p < .05; ***p < .01.

The baseline bivariate model in column (1) shows that rural households on average possess more net wealth than their urban counterparts. Initial results exhibit a negative urban-rural wealth gap that is both statistically significant and economically meaningful, as their net wealth on average is around 100k€ higher. This corresponds to a 36% difference relative to the urban mean of around 274k€, as shown by the constant term. However, this straightforward comparison masks key differences between the two populations – particularly the age and education of the FKP, as well as household size, which are documented in Appendix V-B.

Once these demographic characteristics[10] are explicitly accounted for in column (2), the initially observed rural coefficient decreases to around 62k€ and no longer indicates a statistically significant difference in net wealth. Overall, the results do not support the hypothesis that urban households are wealthier than rural households. With respect to the control variables, both age and household size show positive, statistically significant effects, while holding a university degree is also positively linked to net wealth. Specifically, relative to the urban households, rural FKP's older age and larger household size tend to raise the rural net wealth advantage, whereas their lower university attainment works in the opposite direction. Noteworthy, residing in East Germany is on average associated with a significantly lower net wealth of 174k, enhancing evidence regarding west-east disparities of economic outcomes in Germany (Falk et al., 2011; Smolny & Kirbach, 2011).

As prior research has barely examined urban-rural differences in household net wealth, we first position our findings in relation to broader socioeconomic indicators, which are widely regarded as central to social and regional inequality (Hertrich & Brenner, 2024; Pike et al., 2024). Accordingly, many studies have captured "left-behindness" through measurable disparities in income and gross domestic product (GDP). Our findings, however, diverge from this literature, which consistently documents lower household income (Ananian & Dellaferrera, 2024; Boulant et al., 2016; Fuest & Immel, 2019; Sicular et al., 2007) and regional GDP in rural areas (Lutz et al., 2013; Thünen Institute of Rural Studies, 2024).

This discrepancy may reflect the distinction between income and wealth, although closely linked, they capture different aspects of a household's financial condition (Gibson-Davis et al., 2022; Lamarche et al., 2020; Wolff & Zacharias, 2009). Whereas income reflects only the short-term inflows of a household, net wealth offers a longer-term measure of its economic standing.

[10] Income is deliberately excluded from the set of control variables, as it is closely correlated with net wealth. Controlling for income would risk absorbing part of the variation that is of direct interest when comparing wealth across households. Untabulated results, which include the income variable in the regression, confirm that the main findings remain qualitatively unchanged.

Studies that examine urban-rural wealth disparities are limited to evidence from Norway and China, both reporting a rural disadvantage in terms of net wealth. Wang et al. (2020) and Hu and Gao (2023) document that Chinese urban households have at least three times the wealth of rural ones, while Galster and Wessel (2024) show higher net wealth with an increasing Norwegian urbanization level. This difference in results might be due to the contrasting settlement patterns of these countries. While Germany is relatively polycentric, with numerous metropolitan areas and economically integrated rural regions, both China and Norway exhibit much stronger spatial concentration (Schmidt-Seiwert et al., 2019). In Norway, around 44% of the population resides in the capital region of Oslo (Birnbrich, 2023). In China, 19 metropolitan areas account for 83% of the population and only for one quarter of the total land area (Liu et al., 2024). Such pronounced concentration tends to create more structurally remote and economically weaker peripheries, which may help explain the urban–rural wealth gaps observed in these countries.

In summary, our findings challenge the conventional narrative that rural households are objectively “left behind” in terms of net wealth. While rural households may be disadvantaged in terms of income, and evidence from other countries suggests disadvantages in net wealth as well, this does not hold for Germany. When measured by net wealth, rural households in Germany are therefore not “left behind.”

3.2 Subjective Gap: Do Rural Households Feel “Left Behind”?

As mentioned in the introduction, the “left behind” narrative has a dual meaning. It refers not only to objective economic disadvantage but also to the subjective perception of falling behind. Hence, evidence of higher net wealth among rural households alone does not necessarily invalidate the narrative, as perceptions of relative decline may persist.

In this section we examine whether rural households perceive themselves as “left behind” in terms of net wealth. We operationalize this perception using the variable net-wealth-decile misperception, defined as the delta of household perceived net wealth decile and their actual position in the net wealth distribution as described in Section 2.2. Since households in our sample generally underestimate their position (see Table V-1), stronger underestimation among rural households compared to urban households can be interpreted as a perception of being "left behind." Table V-4 presents the results of these multivariate analyses with net-wealth-decile misperception as the dependent variable.

Table V-4: Linear Regression on Net Wealth Decile Misperception

	Dependent Variable		
	(1) Net Wealth Decile Misperception	(2) Net Wealth Decile Misperception	(3) Net Wealth Decile Misperception
Rural Household Indicator	-1.2575***	-0.9672***	-0.4062***
	(0.1535)	(0.1393)	(0.1289)
Real Estate Ownership			-2.2027***
			(0.1563)
Age		-0.0331***	-0.0189***
		(0.0037)	(0.0037)
Male		-0.1163	0.0097
		(0.1251)	(0.1153)
University Degree		0.1575	0.3674**
		(0.1486)	(0.1495)
Financial Literacy		-0.2852***	-0.1603*
		(0.1012)	(0.0839)
Household Size		-0.2272***	-0.0788
		(0.0646)	(0.0609)
Income		0.1451	0.2396**
		(0.1194)	(0.0972)
Net Wealth		-1.1931***	-0.6669***
		(0.2371)	(0.2006)
Region East		0.0589	-0.1230
		(0.1531)	(0.1306)
Constant	-0.5290***	1.9663***	1.4267***
	(0.1000)	(0.4233)	(0.3673)
Observations	2,777	2,777	2,777
R^2	0.0715	0.2640	0.4100
F-Statistic	67.0853***	33.4590***	80.7600***

Notes: This table reports the results of our regressions on $Net\ Wealth\ Decile\ Misperception$. In column (1), $Net\ Wealth\ Decile\ Misperception$ is regressed solely on the $Rural\ Household\ Indicator$. In column (2), $Net\ Wealth\ Decile\ Misperception$ is regressed on the $Rural\ Household\ Indicator$ alongside a set of demographic control variables, which include $Age, Gender, University\ Degree, Financial\ Literacy, Household\ Size, Income, Net\ Wealth, and\ Region\ East$. The $Real\ Estate\ Ownership$ indicator is introduced as an additional control in column (3). For interpretability, we z-standardized household income and household net wealth (i.e., subtracted the mean and divided by the standard deviation) prior to including them in the model. Standard errors (in parentheses) are adjusted for imputation error following Rubin's Rules as well as for survey design using replicate weights. The data are weighted according to the sampling weights provided in the dataset. Significance levels are indicated as follows: *p < .1; **p < .05; ***p < .01.

The bivariate baseline model in column (1) shows that rural households, on average, underestimate their relative wealth position by around 1.26 deciles more than urban households – a difference that is highly statistically significant. The estimates imply that urban households underestimate their net wealth by roughly 0.53 deciles, while rural households underestimate theirs by about 1.79 deciles.[11] Once demographic controls are introduced in column (2), the effect of the underestimation of rural households decreases to approximately 0.97 deciles but remains highly statistically significant. Even with demographic controls, rural respondents still underestimate their net wealth by about one decile more than comparable urban peers.

Reality and perception of net wealth of urban and rural households do therefore differ markedly: Even though rural households are not objectively "left behind" in terms of net wealth, they consistently perceive themselves as being worse off than comparable urban households.

Consistent with earlier studies (Gasiorowska, 2014; Maison et al., 2019; Tan et al., 2020), these results illustrate that subjective economic assessments are only weakly correlated with objective wealth indicators, reinforcing the notion that subjective and objective dimensions of economic well-being do not always move in tandem. In line with this, our findings provide evidence that rural households are not "left behind" in terms of net wealth, yet they nonetheless feel "left behind" (Deppisch, 2021; Rodríguez-Pose, 2018). This result highlights that an objective basis for being "left behind" is by no means a necessary condition for feeling "left behind" – and that such perceptions can even arise in direct contradiction to the objective reality.

To gain further insight into this divergence between objective wealth and perceived standing, we examine a key structural factor: real estate ownership. This asset class stands out both in its magnitude and in its distribution. Empirically, real estate represents by far the largest share of household wealth (Arrondel et al., 2018; Flavin & Yamashita, 2002) , and ownership rates are significantly higher among rural households (Kohl, 2016; Lerbs & Oberst, 2014). In our sample, 62.98% of rural households own property, compared to just 31.54% of urban households (see Appendix V-B). Incorporating a real estate indicator into our model allows us to directly test

[11] The constant of approximately −0.53 reflects the average underestimation among urban households, while the coefficient of −1.26 for the Rural Household Indicator captures the additional underestimation among rural households. Together, they indicate a total average underestimation of −1.79 deciles for rural households (see also Appendix V-B).

how much of the rural perception gap can be traced to differences in ownership rather than to location-specific factors.

We therefore include a real estate indicator in column (3) to analyze differences in wealth perception once ownership is held constant. This materially changes the estimates: Real estate ownership is associated with a highly significant underestimation of around 2.20 deciles, whereas the underestimation effect of the rural household indicator falls to around 0.41 deciles, nonetheless remaining highly significant. This suggests that real estate ownership may be a key driver of the perception gap.

One plausible explanation for the general underestimation and the additional perception gap between urban and rural households is the low salience of the financial value of real estate. The salience bias refers to the tendency to focus on information that is vivid, easily accessible, or cognitively prominent, while neglecting less visible or less frequently considered elements (Taylor & Thompson, 1982).

Limited salience of real estate arises because it is inherently illiquid (Gibson et al., 2022; He et al., 2018), seldom revalued (Edelstein & Quan, 2006; Koijen et al., 2025; Lin & Vandell, 2007), and rarely top-of-mind in day-to-day financial considerations. As a result, many households systematically discount – or even forget – the wealth locked up in real estate, possibly leading to substantial underestimation of their true economic position.

Although reliance on salient information often simplifies decision-making effectively, it can also lead to systematic biases and irrational judgments. (Bordalo et al., 2013; Tversky & Kahneman, 1973) In our context, the effect of systematically underestimating real estate is particularly pronounced among rural households, because they own real estate significantly more often. Given that real estate constitutes the largest single asset class in household portfolios overall, rural households are especially susceptible to undervaluing their net wealth compared to their urban counterparts. Our empirical findings clearly support our interpretation, with ownership accounting for over two deciles of households' underestimation of their net wealth position.

A related psychological mechanism is illustrated by Sussman and Shafir (2012), demonstrating that individuals perceive themselves as wealthier when holding less debt, even when net worth remains constant. Debt, especially in the form of loans, does appear more salient in the wealth perception of people. This is directly relevant in our context, as mortgage debt – although fully included in net-wealth calculations – may dominate the perception of housing-related finances, leading homeowners to feel less wealthy overall.[12] Nonetheless, extending this idea, we find that salience also matters on the asset side: when a large share of wealth is locked in less salient assets such as real estate, households further underestimate their overall wealth. Given that rural households tend to neglect their real estate assets when assessing their wealth, their subjective feeling of being "left behind" becomes plausible.

Finally, despite controlling for low salience real estate, a significant underestimation of rural households compared to their urban counterparts remains. This remaining difference indicates that apart from real estate, there is still a significant feeling of being "left behind" that is not attributable to the low salience of real estate or common socioeconomic characteristics. Thus, although our findings challenge the generalized view of rural economic disadvantage, they highlight a continued subjective perception of relative deprivation among rural households.

3.3 Robustness Checks

While our core findings indicate systematic urban–rural disparities in net wealth and its perception, methodological concerns such as endogeneity remain. Thus, we conduct a series of robustness checks to assess the consistency and sensitivity of our results.

[12] Untabulated regression results indicate that debt is also strongly associated with an underestimation of the household's wealth position. However, even after including debt in the regression, the rural and real estate indicators still remain highly significant.

Propensity Score Matching

While survey weights correct for non-response and improve representativeness at the population level, they do not eliminate systematic differences between rural and urban households. Rural samples typically comprise older, larger, and less-educated households than their urban counterparts – characteristics that also influence wealth levels and misperception. To ensure that differences in wealth and misperception are not merely driven by these compositional factors, we employ Propensity Score Matching (PSM) to balance observable characteristics between the two groups.

Specifically, we estimate a probit model predicting rural residence based on the demographic covariates listed in Panel C of Table 1. The fitted values from this model serve as each household's propensity score. Each rural household is then matched to the single urban household whose score is closest, using 1:1 nearest-neighbor matching with replacement.

To align the matching design with our regression framework of Section 2.4, we perform two versions of this matching procedure. In the first, standardized net wealth is omitted from the covariate set since it serves as the outcome in subsequent analyses.[13] In the second, standardized net wealth is included as an additional matching covariate. After matching, covariate balance improved considerably in both cases, with standardized mean differences for nearly all variables reduced below the conventional 0.1 threshold (Austin, 2009).[14]

With these balanced matched samples, we re-estimate the regression models for net wealth (Appendix V-C) and net wealth decile misperception (Appendix V-D). In the matched sample, the net wealth advantage of rural households is already statistically insignificant in our baseline specification. This result is expected, since matching ensures that rural and urban households no longer differ in demographics that originally drove the gap. Overall, our main results remain

[13] Income is again deliberately excluded from the covariate set, as it is closely correlated with net wealth.

[14] The only exception is the region east indicator in the first matching, which retained a standardized mean difference of 0.1870. This remaining imbalance reflects the higher concentration of rural households in eastern Germany, where comparable urban counterparts are relatively scarce when income and net wealth are not included as matching covariates.

robust in the matched sample, confirming that they are not driven by endogeneity from observable differences between rural and urban households.

Extended Sample (Including 20k – 100k Municipalities)

Until now, we excluded households residing in municipalities with 20,000 to 100,000 inhabitants. To evaluate the robustness of our results, we now add these households to the sample. The $Rural\ Household\ Indicator$ is accordingly modified: municipalities of up to 20,000 inhabitants are treated as rural, while larger ones are classified as urban.[15]

Appendices E–F replicate our analyses on the extended sample and show that our main results remain robust.

Spatial Extremes (<5k vs. >500k)

After confirming that our results hold in the extended sample, we strengthen the treatment contrast by focusing on the population extremes – municipalities with the clearest urban–rural divide where socioeconomic differences are likely to be most pronounced. Specifically, we exclude households residing in communities with 5,000 to 500,000 inhabitants. In this trimmed sample, the $Rural\ Household\ Indicator$ equals 1 for households in municipalities with fewer than 5,000 residents and 0 for those in major cities exceeding 500,000 residents.

The results for this spatial-extremes subsample, presented in Appendices G-H, differ slightly from the main analyses. In the net wealth regressions (Appendix V-G), the rural coefficient of the bivariate benchmark model remains around €89k but loses statistical significance, most likely due to the smaller sample size.

In column (3), where the net wealth decile misperception is the dependent variable, the rural coefficient dummy becomes statistically insignificant. Since the effect size with an

[15] We acknowledge that municipalities with 20,000 to 100,000 inhabitants do not fully conform to a standard definition of "urban." However, our empirical design requires a binary classification, which necessitates assigning them to the urban category.

underestimation of around 0.34 deciles is comparable with our main findings, the absence of significance is likewise probably due to the smaller sample size (Appendix V-H).

Single-Person Households

As an additional robustness check, we restrict the sample to single-person households. This restriction ensures that the demographic controls[16] describing the FKP align directly with household characteristics. Moreover, net-wealth allocation is straightforward in single households, allowing the FKP to assess actual and perceived wealth without the complexity of multiple household members.

In this new sample (Appendices I-J), our main findings remain unchanged, with one exception. In the bivariate benchmark model with net wealth as the dependent variable (Appendix V-I), we find no significant difference between urban and rural households. This finding is in line with our main results in Section 3.1, suggesting that the rural–urban wealth gap is driven rather by demographic factors, especially household size, than by location-specific factors alone.

Interquartile Wealth Sample

Our final robustness check addresses a methodological limitation inherent to the net wealth decile misperception measure. The scope for misperception is mechanically constrained at both ends of the distribution, which may influence the magnitude of observed deviations. Households in the top decile cannot overestimate, and those in the bottom decile cannot underestimate, their rank. To neutralize this constraint, we repeat the analysis with an interquartile sample, excluding the wealthiest and poorest 25% of households. Results in Appendix V-K confirm that our main findings are not artifacts of the decile-based measure.

[16] Household size was excluded as a control variable, as it does not vary within this subsample and would otherwise cause perfect collinearity.

4. Discussion and Conclusion

We find, contrary to the narrative of rural economic disadvantage, that rural households are on average not less wealthy than urban households but rather have more net wealth than urban households. However, once demographic differences are taken into account, the apparent rural advantage is no longer statistically distinguishable from zero. The observed wealth benefit of rural households is therefore not solely location-specific but largely reflects compositional differences in the populations under study.

Across the entire sample, households systematically underestimate their true wealth by about one decile. Building on this general pattern, our analysis shows that rural households underestimate their wealth position more strongly than otherwise comparable urban households – by roughly one additional decile. In other words, when comparing households that objectively occupy the same place in the wealth distribution, rural residents systematically perceive themselves as relatively worse off. This misperception mirrors the broader “left behind” narrative frequently associated with rural populations.

Furthermore, the limited salience of real estate appears to contribute to this misperception. Because real estate is illiquid, rarely revalued, and less salient in everyday financial considerations, households may undervalue the wealth locked in property. Since property ownership is much more common among rural households, low salience disproportionately affects them, perpetuating the perception of economic inferiority and reinforcing the “left behind” narrative.

Yet even after accounting for real estate ownership and other demographic characteristics, a sizable perception gap remains. Rural households continue to underestimate their relative wealth position compared to urban households, pointing to additional drivers of perceived disadvantage that cannot be explained by asset salience or observable demographics alone.

Our analysis, while robust, faces two main limitations that also open promising directions for future research. First, the limited granularity of the Deutsche Bundesbank’s survey data restricts our ability to distinguish between suburban households near metropolitan centers and genuinely remote rural communities. Future studies should use higher-resolution spatial data to better capture heterogeneity within the “rural” category and assess how proximity to metropolitan areas impacts both actual and perceived wealth.[17]

Secondly, the existing findings rely solely on German data from the Panel on Household Finances. Replicating the current analysis in cross-national studies is necessary to determine whether these patterns are country-specific or reflect broader trends across advanced economies. Such replication is particularly relevant for countries with a similarly polycentric peripheral structure, including the Netherlands, Switzerland, and Belgium (Schmidt-Seiwert et al., 2019).

Despite these limitations, our findings offer important implications for households and policymakers. Contrary to the conventional “left behind” narrative, rural households are not economically disadvantaged when viewed from a net wealth perspective. This underscores the need to broaden traditional measurement approaches: net wealth should complement income as a standard dimension in assessing spatial inequality. Media and official statistics should therefore present income and net wealth jointly to convey a more accurate and balanced picture of economic disparities while avoiding displaying rural regions as uniformly “left behind.”

Even when objective economic indicators reveal no evidence of rural deprivation, perceptions can still diverge sharply from these realities. Our results show that rural households feel disadvantaged in terms of net wealth even in the absence of any material deprivation. This divergence between perception and reality calls for policies that address not only material inequalities but also perceived relative disadvantages. Effective policy communication, symbolic inclusion, and public recognition of rural contributions can be decisive in this regard.

[17] The map of city and municipality types in Germany of the German Federal Office for Building and Regional Planning (2023) shows, however, that only a minority of rural regions are adjacent to metropolises. https://www.bbsr.bund.de/BBSR/DE/forschung/raumbeobachtung/Raumabgrenzungen/deutschland/gemeinden/StadtGemeindetyp/StadtGemeindetyp.html

As emphasized by Wolf (2021), visible and tangible gestures of appreciation – the “hanging baskets” of regional policy – may in some cases foster a stronger sense of belonging than complex financial programs that remain abstract or poorly communicated.

At the same time, it is important to counterbalance the prevailing “left behind” narrative of rural regions by emphasizing areas where rural life is comparatively advantageous. Apart from our results, rural households enjoy tangible benefits in many respects: lower living costs (Hawk, 2013; Ravallion & van de Walle, 1991), reduced income inequality (Frieden et al., 2023; Garbasevschi et al., 2023), higher homeownership rates (Kohl, 2016; Lerbs & Oberst, 2014), lower crime levels (Ceccato, 2015; Entorf & Spengler, 2000), less pollution (Kilpatrick et al., 2024; Strosnider et al., 2017), and higher life satisfaction across Northern and Western Europe and the U.S. (Hoogerbrugge & Burger, 2024; Jurčišinová et al., 2025; Okulicz-Kozaryn, 2024).

Yet these benefits often coexist with feelings of being “left behind”, making it crucial to address the remaining gaps between perception and reality, as they can strongly influence political behavior and collective attitudes. Both policy makers and households should always critically question whether a perceived disadvantage is supported by evidence or formed mainly by prevailing narratives and populism. This distinction matters: even without adverse economic outcomes, the feeling of being left behind can fuel political discontent and populist voting. Once politicized, such perceptions may be mobilized by political actors and translated into electoral outcomes, as seen with the AfD in Germany, the Brexit vote in the UK, and the election of Donald Trump in the United States.

Ultimately, subjective perceptions of socioeconomic standing – not just actual economic conditions – matter.

5. References

Albers, T. N. H., Bartels, C., & Schularick, M. (2020). The distribution of wealth in Germany, 1895–2018(001). https://www.econstor.eu/handle/10419/268554

Ananian, S., & Dellaferrera, G. (2024). *Employment and wage disparities between rural and urban areas* (ILO Working Paper No. 107). Geneva: International Labour Organization (ILO). https://www.econstor.eu/handle/10419/283544 https://doi.org/10.54394/LNZT9066

Arrondel, L., Rampazzi, C., Savignac, F., Bartiloro, L., Schürz, M., Lindner, P., Vermeulen, P., Fessler, P., Mathä, T. Y., & Schmidt, T. (2018). How do households allocate their assets? Stylized facts from the eurosystem Household Finance and Consumption Survey. *44th Issue (June 2016) of the International Journal of Central Banking*. https://www.ijcb.org/journal/ijcb16q2a4.htm

Austin, P. C. (2009). Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in Medicine*, *28*(25), 3083–3107. https://doi.org/10.1002/sim.3697

Balestra, C., & Oehler, F. (2023). OECD papers on well-being and inequalities: Measuring the joint distribution of household income, consumption and wealth at the micro level., *11*. https://doi.org/10.1787/f9d85db6-en

Bartels, C., & Schroeder, C. (2020). *Income, consumption and wealth inequality in Germany: Three concepts, three stories?* (Basic Papers 02/2020). Berlin: Forum for a New Economy. https://www.econstor.eu/handle/10419/281434

Batista, R. M., Sussman, A. B., & Trueblood, J. S. (2023). Self-other differences in perceptions of wealth. *Journal of Experimental Social Psychology*, *104*, 104420. https://doi.org/10.1016/j.jesp.2022.104420

Benhabib, J., & Bisin, A. (2018). Skewed wealth distributions: Theory and empirics. *Journal of Economic Literature*, *56*(4), 1261–1291. https://doi.org/10.1257/jel.20161390

Bergmann, K., Diermeier, M., & Niehues, J. (2017). Die AfD: Eine Partei der sich ausgeliefert fühlenden Durchschnittsverdiener? *Zeitschrift Für Parlamentsfragen*, *48*(1), 57–75. http://www.jstor.org/stable/26428857

Birnbrich, M. (2023). Country Profile of Norway. Hannover. = ARL Country Profiles;. *Academy for Territorial Development in the Leibniz Association*. https://www.arl-international.com/knowledge/country-profiles/norway/rev/3840

Bordalo, P., Gennaioli, N., & Shleifer, A. (2013). Salience and consumer choice. *Journal of Political Economy*. https://doi.org/10.1086/673885

Boulant, J., Brezzi, M., & Veneri, P. (2016). *Income levels and inequality in metropolitan areas: A comparative approach in OECD countries. OECD Working Papers*. https://www.oecd-ilibrary.org/urban-rural-and-regional-development/income-levels-and-inequality-in-metropolitan-areas_5jlwj02zz4mr-en

Bricker, J., Henriques, A., Krimmel, J., & Sabelhaus, J. (2016). Measuring income and wealth at the top using administrative and survey data. *Brookings Papers on Economic Activity*, *2016*(1), 261–331. https://doi.org/10.1353/eca.2016.0016

Brown, W. M., Newbold, K. B., & Beckstead, D. (2010). Growth and change in human capital across the Canadian urban hierarchy, 1996–2001. *Urban Studies*, *47*(7), 1571–1586. https://doi.org/10.1177/0042098009353625

Ceccato, V. (2015). *Rural crime and community safety* (First edition). *Routledge studies in crime and society: Vol. 18*. Routledge. https://doi.org/10.4324/9780203725689

Deppisch, L. (2021). “Where people in the countryside feel left behind populism has a clear path” – An analysis of the popular media discourse on how infrastructure decay, fear of social decline, and right-wing (extremist) values contribute to support for right-wing populism. Advance online publication. https://doi.org/10.3220/WP1613460042000

Deppisch, L., Osigus, T., & Klärner, A. (2022). How rural is rural populism? On the spatial understanding of rurality for analyses of right-wing populist election success in Germany. https://doi.org/10.1111/ruso.12397

Diermeier, M. (2020). Ist mehr besser? Politische Implikationen der disparaten Daseinsvorsorge in Deutschland [The more the better? Political implication of the diverging supply of public services in Germany]. *Zeitschrift für Politikwissenschaft*, *30*(4), 539–568. https://doi.org/10.1007/s41358-020-00239-y

Edelstein, R. H., & Quan, D. C. (2006). How does appraisal smoothing bias real estate returns measurement? *The Journal of Real Estate Finance and Economics*, *32*(1), 41–60. https://doi.org/10.1007/s11146-005-5177-9

Entorf, H., & Spengler, H. (2000). Socioeconomic and demographic factors of crime in Germany. *International Review of Law and Economics*, *20*(1), 75–106. https://doi.org/10.1016/s0144-8188(00)00022-3

Essletzbichler, J., Disslbacher, F., & Moser, M. (2018). The victims of neoliberal globalisation and the rise of the populist vote: A comparative analysis of three recent electoral decisions. *Cambridge Journal of Regions, Economy and Society*, *11*(1), 73–94. https://doi.org/10.1093/cjres/rsx025

Eurostat (2025). Distribution of population by tenure status, type of household and income group ; data code: ilc_lvho02. https://doi.org/10.2908/ILC_LVHO02

Falk, A., Kuhn, A., & Zweimüller, J. (2011). Unemployment and right-wing extremist crime *The Scandinavian Journal of Economics*, *113*(2), 260–285. https://doi.org/10.1111/j.1467-9442.2011.01648.x

Federal Election Commission. (2016). *Federal elections 2016: Election results for the U.S. president, the U.S. Senate and the U.S. House of Representatives*. https://www.fec.gov/resources/cms-content/documents/federalelections2016.pdf

Federal Election Commission (2024). Official 2024 presidential general election results. https://www.fec.gov/resources/cms-content/documents/2024presgeresults.pdf

Fessler, P., & Rapp, S. (2023). *The subjective wealth distribution: How it arises and why it matters to inform policy?* (Working Paper No. 249). Vienna: Oesterreichische Nationalbank (OeNB). https://www.econstor.eu/handle/10419/283417

Flavin, M., & Yamashita, T. (2002). Owner-occupied housing and the composition of the household portfolio. *American Economic Review*, *92*(1), 345–362. https://doi.org/10.1257/000282802760015775

Frieden, I., Peichl, A., & Schüle, P. (2023). Regional income inequality in Germany. *EconPol Forum*, *24*(2), 50–55. https://www.econstor.eu/handle/10419/272168

Fuest, C., & Immel, L. (2019). Ein zunehmend gespaltenes Land? – Regionale Einkommensunterschiede und die Entwicklung des Gefälles zwischen Stadt und Land sowie West- und Ostdeutschland. *Ifo Schnelldienst*, *72*(16), 19–28. https://www.econstor.eu/handle/10419/206905

Gallusser, D., & Krapf, M. (2019). Joint income-wealth inequality: An application using administrative tax data. *SSRN Electronic Journal.* Advance online publication. https://doi.org/10.2139/ssrn.3468011

Galster, G., & Wessel, T. (2024). Urban residence as a driver of wealth differentials: New evidence from Norway. *Population, Space and Place*, *30*(4), Article e2753, e2753. https://doi.org/10.1002/psp.2753

Garbasevschi, O. M., Taubenböck, H., Schüle, P., Baarck, J., Hufe, P., Wurm, M., & Peichl, A. (2023). Learning income levels and inequality from spatial and sociodemographic data in Germany. *Applied Geography*, *159*, 103058. https://doi.org/10.1016/j.apgeog.2023.103058

Gasiorowska, A. (2014). The relationship between objective and subjective wealth is moderated by financial control and mediated by money anxiety. *Journal of Economic Psychology*, *43*, 64–74. https://doi.org/10.1016/j.joep.2014.04.007

German Federal Office for Building and Regional Planning. (2023). *City and municipality types in Germany in 2021*. https://www.bbsr.bund.de/BBSR/DE/forschung/raumbeobachtung/Raumabgrenzungen/deutschland/gemeinden/StadtGemeindetyp/StadtGemeindetyp.html

German Federal Office for Building and Regional Planning. (2025). *Stadt- und Gemeindetypen in Deutschland; Bundesinstitut für Bau-, Stadt- und Raumforschung im Bundesamt für Bauwesen und Raumordnung (BBSR)*. https://www.bbsr.bund.de/BBSR/DE/forschung/raumbeobachtung/Raumabgrenzungen/deutschland/gemeinden/StadtGemeindetyp/StadtGemeindetyp.html

German Federal Statistical Office (2021). Land- und Forstwirtschaft, Fischerei. Bodenfläche nach Art der tatsächlichen Nutzung; Fachserie 3 Reihe 5.1. https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Landwirtschaft-Forstwirtschaft-Fischerei/Flaechennutzung/Publikationen/Downloads-Flaechennutzung/bodenflaechennutzung-2030510217004.pdf?__blob=publicationFile

German Federal Statistical Office (2024a). Population on December 31, 2023, by nationality and federal state. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Bevoelkerungsstand/Tabellen/bevoelkerung-nichtdeutsch-laender.html

German Federal Statistical Office (2024b). Sozialbericht 2024: Ein Datenreport für Deutschland. https://www.bpb.de/system/files/dokument_pdf/Sozialbericht_2024_bf.pdf

Gibson, R., Hoesli, M., & Shan, J. (2022). *The valuation of illiquid assets: A focus on private equity and real estate*. https://doi.org/10.3905/jai.2022.1.163

Gibson-Davis, C., Keister, L. A., Gennetian, L. A., & Lowell, W. (2022). Net worth poverty and child development. *Socius*, *8*. https://doi.org/10.1177/23780231221111672

Gimpel, J. G., Lovin, N., Moy, B., & Reeves, A. (2020). The urban–rural gulf in American political behavior. *Political Behavior*, *42*(4), 1343–1368. https://doi.org/10.1007/s11109-020-09601-w

Goerres, A., Spies, D. C., & Kumlin, S. (2018). The electoral supporter base of the Alternative for Germany. *Swiss Political Science Review*, *24*(3), 246–269. https://doi.org/10.1111/spsr.12306

Grabka, M. M., & Westermeier, C. (2014). Persistently high wealth inequality in Germany. *DIW Economic Bulletin*, *4*(6), 3–15. https://www.econstor.eu/handle/10419/98670

Hawk, W. (2013). Expenditures of urban and rural households in 2011. https://ecommons.cornell.edu/bitstream/1813/79405/1/bls_btn_expenditures_of_urban_and_rural_households.pdf

He, X., Lin, Z., & Liu, Y. (2018). Volatility and liquidity in the real estate market. *Journal of Real Estate Research*, *40*(4), 523–550. https://doi.org/10.1080/10835547.2018.12091511

Hertrich, T. J., & Brenner, T. (2024). *Looking behind the curtain: A model of left behind places and feelings* (Working Papers on Innovation and Space No. 01.24). Marburg: Philipps-University Marburg, Department of Geography. https://www.econstor.eu/handle/10419/282312

Hilmar, T. (2022). Fragile bonds of recognition: Exploring the social underpinnings of sentiments of exclusion in post-1989 East Germany. *European Journal of Sociology*, *63*(2), 247–278. https://doi.org/10.1017/s0003975622000236

Hoogerbrugge, M., & Burger, M. J. (2024). Housing and urban-rural differences in subjective wellbeing in the Netherlands. In C. Lenzi & V. Fedeli (Eds.), *Multidisciplinary movements in research. Spatial inequalities and wellbeing: A multidisciplinary approach* (pp. 97–118). Edward Elgar Publishing. https://doi.org/10.4337/9781802202632.00009

Hu, W., & Gao, T. (2023). Household wealth distribution and its impact in China: Evidence from the China Family Panel Studies. *Sustainability*, *15*(8), 6928. https://doi.org/10.3390/su15086928

Johnston, R., Manley, D., & Jones, K. (2016). Spatial polarization of presidential voting in the United States, 1992–2012: The "big sort" revisited. *Annals of the American Association of Geographers*, *106*(5), 1047–1062. https://doi.org/10.1080/24694452.2016.1191991

Jurčišinová, V., Želinský, T., & Hudec, O. (2025). The urban paradox and rural appeal: Subjective well-being across diverse European living environments. *Applied Research in Quality of Life*, 1–29. https://doi.org/10.1007/s11482-025-10437-x

Kalckreuth, U., Eisele M, Le Blanc J, Schmidt T, & Zhu J (2012). The PHF: A comprehensive panel survey on household finances and wealth in Germany. Bundesbank Discussion Paper. Deutsche Bundesbank. https://www.bundesbank.de/resource/blob/617148/e6ba33aa5cb2f0fd12db5a4205b5cd23/mL/2012-07-10-dkp-13-data.pdf

Kasinger, C., Braunheim, L., Beutel, M., & Brähler, E. (2023). Closing the "happiness gap" by closing the wealth gap: The role of wealth on life satisfaction between east and west-Germans. *Journal of Public Health*, *31*(9), 1433–1442. https://doi.org/10.1007/s10389-022-01716-0

Kenny, M., & Luca, D. (2021). The urban-rural polarisation of political disenchantment: An investigation of social and political attitudes in 30 European countries. *Cambridge Journal of Regions, Economy and Society*, *14*(3), 565–582. https://doi.org/10.1093/cjres/rsab012

Kilpatrick, D. J., Hung, P., Crouch, E., Self, S., Cothran, J., Porter, D. E., & Eberth, J. M. (2024). Geographic variations in urban-rural particulate matter (PM2.5) concentrations in the United States, 2010-2019. *GeoHealth*, *8*(9), e2023GH000920. https://doi.org/10.1029/2023GH000920

Kohl, S. (2016). Urban history matters: Explaining the German–American homeownership gap. *Housing Studies*, *31*(6), 694–713. https://doi.org/10.1080/02673037.2015.1121213

Koijen, R. S. J., Shah, N., & van Nieuwerburgh, S. (2025). The commercial real estate ecosystem. Advance online publication. https://doi.org/10.2139/ssrn.5120847

Lamarche, P., Oehler, F., & Rioboo, I. (2020). European household's income, consumption and wealth. *Statistical Journal of the IAOS*, *36*(4), 1175–1188. https://doi.org/10.3233/SJI-190528

Lerbs, O. W., & Oberst, C. A. (2014). Explaining the spatial variation in homeownership rates: Results for German regions. *Regional Studies*, *48*(5), 844–865. https://doi.org/10.1080/00343404.2012.685464

Lin, Z., & Vandell, K. D. (2007). Illiquidity and pricing biases in the real estate market. *Real Estate Economics*, *35*(3), 291–330. https://doi.org/10.1111/j.1540-6229.2007.00191.x

Liu, P., Zhong, F., & Han, N. (2024). Efficiency and equity: Effect of urban agglomerations' spatial structure on green development efficiency in China. *Sustainable Cities and Society*, *108*, 105504. https://doi.org/10.1016/j.scs.2024.105504

Lusardi, A., & Mitchell, O. S. (2008). Planning and financial literacy: How do women fare? *American Economic Review*, *98*(2), 413–417. https://doi.org/10.1257/aer.98.2.413

Lutz, R., Spies, M., Reusser, D. E., Kropp, J. P., & Rybski, D. (2013). Characterizing the development of sectoral gross domestic product composition. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, *88*(1), 12804. https://doi.org/10.1103/physreve.88.012804

MacKinnon, D., Béal, V., & Leibert, T. (2024). Rethinking 'left-behind' places in a context of rising spatial inequalities and political discontent. *Regional Studies*, *58*(6), 1161–1166. https://doi.org/10.1080/00343404.2023.2291581

Maison, D., Marchlewska, M., Sekścińska, K., Rudzinska-Wojciechowska, J., & Łozowski, F. (2019). You don't have to be rich to save money: On the relationship between objective versus subjective financial situation and having savings. *PloS One*, *14*(4), e0214396. https://doi.org/10.1371/journal.pone.0214396

Okulicz-Kozaryn, A. (2024). Urban regrets (Unhappy metros: Satisfaction with life scale (SWLS)). *Heliyon*, *10*(11), e30729. https://doi.org/10.1016/j.heliyon.2024.e30729

Partridge, M. D., & Rickman, D. S. (2008). Distance from urban agglomeration economies and rural poverty. *Journal of Regional Science*, *48*(2), 285–310. https://doi.org/10.1111/j.1467-9787.2008.00552.x

PHF Survey Team (2025). PHF User Guide Version 5.0. https://www.bundesbank.de/resource/blob/617424/ea2fb51664cbf2dcfa66114977790a8c/472B63F073F071307366337C94F8C870/phf-user-guide-data.pdf

Pike, A., Béal, V., Cauchi-Duval, N., Franklin, R., Kinossian, N., Lang, T., Leibert, T., MacKinnon, D., Rousseau, M., Royer, J., Servillo, L., Tomaney, J., & Velthuis, S. (2024). 'Left behind places': A geographical etymology. *Regional Studies*, *58*(6), 1167–1179. https://doi.org/10.1080/00343404.2023.2167972

Ravallion, M., & van de Walle, D. (1991). Urban-rural cost-of-living differentials in a developing economy. *Journal of Urban Economics*, *29*(1), 113–127. https://doi.org/10.1016/0094-1190(91)90030-b

Richter, C., & Bösch, L. (2017). Demokratieferne Räume? Wahlkreisanalyse zur Bundestagswahl 2017. https://www.fluechtlingsrat-brandenburg.de/wp-content/uploads/2017/12/studie_afd_wahlkreisanalyse_btw17.pdf

Rodríguez-Pose, A. (2018). The revenge of the places that don't matter (and what to do about it). *Cambridge Journal of Regions, Economy and Society 11.1: 189-209.* https://doi.org/10.1093/cjres/rsx024

Roscigno, V. J., & Crowle, M. L. (2001). Rurality, institutional disadvantage, and achievement/attainment *Rural Sociology*, *66*(2), 268–292. https://doi.org/10.1111/j.1549-0831.2001.tb00067.x

Royer, J., & Leibert, T. (2024). Suspended in time? Peripheralised and "left behind" places in Germany. *Geographica Helvetica*, *79*(3), 221–237. https://doi.org/10.5194/gh-79-221-2024

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. Wiley. https://doi.org/10.1002/9780470316696

Schmidt, T., & Eisele, M. (2013). Oversampling vermögender Haushalte im Rahmen der Studie „Private Haushalte und ihre Finanzen (PHF)". https://www.bundesbank.de/resource/blob/617146/92b765b0227577b65ccec0fdbd8a86bc/mL/phf-oversampling-data.pdf

Schmidt, T., Kocharkov, G., Paule-Paludkiewicz, H., Tzamourani, P., Wehrhöfer, N., Zhu, J., Bernard, R., Hebbat, M., Ibragimova, N., Kothmayr, L., Qiu, Z., Gerhardt, J., & Köhler, M. (2025). *Panel on Household Finances Wave 1- Wave 5.* https://doi.org/10.12757/BBK.PHF.W1-W5.01.01

Schmidt-Seiwert, V., Binot, R., Kiel, L., Duvernet, C., Golzio, P., & Pütz, T. (2019). *Die Unterschiede bestimmen die Vielfalt in Europa: Ein Atlas ausgewählter Aspekte der räumlichen Strukturen und Entwicklungen*. German Federal Office for Building and Regional Planning (BSSR). https://www.bbsr.bund.de/BBSR/DE/veroeffentlichungen/sonderveroeffentlichungen/2019/europa-atlas.html

Schwerter, J., Bleher, J., Doebler, P., & McElvany, N. (2023). *Metropolitan, urban, and rural regions – How regional differences affect elementary school students' academic achievement, well-being, and motivation.* https://doi.org/10.1177/23328584251331453

Sicular, T., Ximing, Y., Gustafsson, B., & Shi, L. (2007). The urban–rural income gap and inequality in China. *Review of Income and Wealth*, *53*(1), 93–126. https://doi.org/10.1111/j.1475-4991.2007.00219.x

Singh, G. K., & Siahpush, M. (2014). Widening rural-urban disparities in all-cause mortality and mortality from major causes of death in the USA, 1969-2009. *Journal of Urban Health*, *91*(2), 272–292. https://doi.org/10.1007/s11524-013-9847-2

Smolny, W., & Kirbach, M. (2011). Wage differentials between East and West Germany: Are they related to the location or to the people? *Applied Economics Letters*, *18*(9), 873–879. https://doi.org/10.1080/13504851.2010.511990

Statista. (2024). *Anzahl der Großstädte nach Bundesländern in Deutschland am 31. Dezember 2023*. https://de.statista.com/statistik/daten/studie/1244003/umfrage/grossstaedte-nach-bundeslaendern/

Strosnider, H., Kennedy, C., Monti, M., & Yip, F. (2017). Rural and urban differences in air quality, 2008-2012, and community drinking water quality, 2010-2015 - United States. *MMWR. Surveillance Summaries*, *66*(13), 1–10. https://doi.org/10.15585/mmwr.ss6613a1

Sun, Y., & Monnat, S. M. (2022). Rural-urban and within-rural differences in COVID-19 vaccination rates. *The Journal of Rural Health*, *38*(4), 916–922. https://doi.org/10.1111/jrh.12625

Sussman, A. B., & Shafir, E. (2012). On assets and debt in the psychology of perceived wealth. *Psychological Science*, *23*(1), 101–108. https://doi.org/10.1177/0956797611421484

Tan, J. J. X., Kraus, M. W., Carpenter, N. C., & Adler, N. E. (2020). The association between objective and subjective socioeconomic status and subjective well-being: A meta-analytic review. *Psychological Bulletin*, *146*(11), 970–1020. https://doi.org/10.1037/bul0000258

Taylor, S. E., & Thompson, S. C. (1982). Stalking the elusive "vividness" effect. *Psychological Review*, *89*(2), 155–181. https://doi.org/10.1037//0033-295x.89.2.155

Thünen Institute of Rural Studies. (2024). *Landatlas. Ausgabe 2024.* https://www.landatlas.de

Tully, S. M., & Sharma, E. (2022). Consumer wealth. *Consumer Psychology Review*, *5*(1), 125–143. https://doi.org/10.1002/arcp.1073

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(2), 207–232. https://doi.org/10.1016/0010-0285(73)90033-9

Velthuis, S., Royer, J., Le Petit-Guerin, M., Cauchi-Duval, N., Franklin, R., Leibert, T., MacKinnon, D., & Pike, A. (2025). Regional varieties of 'left-behindness' in the EU15. *Regional Studies*, *59*(1), Article 2417704. https://doi.org/10.1080/00343404.2024.2417704

Wang, Y., Li, Y., Huang, Y., Yi, C., & Ren, J. (2020). Housing wealth inequality in China: An urban-rural comparison. *Cities*, *96*, 102428. https://doi.org/10.1016/j.cities.2019.102428

Weber, B., Jensen, L., Miller, K., Mosley, J., & Fisher, M. (2005). A critical review of rural poverty literature: Is there truly a rural effect? *International Regional Science Review*, *28*(4), 381–414. https://doi.org/10.1177/0160017605278996

Wolf, R. (2021). Boris Johnson must sweat the small stuff to change lives post-Brexit. https://www.politico.eu/article/boris-johnson-must-sweat-the-small-stuff-to-change-lives-post-brexit/

Wolff, E. N., & Zacharias, A. (2009). Household wealth and the measurement of economic well-being in the United States. *The Journal of Economic Inequality*, *7*(2), 83–115. https://doi.org/10.1007/s10888-007-9068-6

Xie, Y., & Jin, Y. (2015). Household wealth in China. *Chinese Sociological Review*, *47*(3), 203–229. https://doi.org/10.1080/21620555.2015.1032158

Zagorsky, J. L. (2000). *Do individuals know how much they are worth? Journal of Financial Counseling and Planning 11.1: 13.*

Zahnd, W. E., James, A. S., Jenkins, W. D., Izadi, S. R., Fogleman, A. J., Steward, D. E., Colditz, G. A., & Brard, L. (2018). Rural-urban differences in cancer incidence and trends in the United States. *Cancer Epidemiology, Biomarkers & Prevention*, *27*(11), 1265–1274. https://doi.org/10.1158/1055-9965.EPI-17-0430

6. Appendix

Appendix V-A: Variable Dimensions and Descriptions

Variable	Dimension	Description
Panel A: Municipality Size		
Rural Household Indicator	1/0	Indicator variable that equals 1 if the household is located in a rural area (<20.000 inhabitants).
Municipality Size: < 5k	1/0	Indicator variable that equals 1 if the household lives in a municipality with less than 5.000 inhabitants.
Municipality Size: 5k – 20k	1/0	Indicator variable that equals 1 if the household lives in a municipality with 5.000–20.000 inhabitants.
Municipality Size: 20k –100k	1/0	Indicator variable that equals 1 if the household lives in a municipality with 20.000–100.000 inhabitants.
Municipality Size: 100k – 500k	1/0	Indicator variable that equals 1 if the household lives in a municipality with 100.000–500.000 inhabitants.
Municipality Size: > 500k	1/0	Indicator variable that equals 1 if the household lives in a municipality with over 500.000 inhabitants.
Panel B: Actual vs. Perceived Wealth		
Net Wealth	k€	Household's net wealth, expressed in thousands of euros, is defined as total assets minus total liabilities. Assets include real assets (e.g., main residence, other properties, vehicles, valuables, and businesses) and financial assets (e.g., deposits, funds, bonds, shares, insurance, and pension claims). Liabilities comprise mortgage debt on real estate and non-mortgage debt, such as consumer, student, credit card, and private loans.
Net Wealth Decile	[1;10]	Net wealth Decile of the household.
Perceived Net Wealth Decile	[1;10]	Perceived Net Wealth Decile of the household by FKP. FKPs were asked to assess the net wealth decile of their household on a scale from one to ten.
Net Wealth Decile Misperception	[-9;9]	Perceived Net Wealth Decile subtracted by Net wealth decile. A positive (negative) value indicates overestimation (underestimation) of the household's own net wealth.
Panel C: Demographics		
Age	[18;100]	Age of the financially knowledgeable person (FKP) of the household in years.
Male	1/0	Indicator variable that equals 1 if the FKP is male.
University Degree	1/0	Indicator variable that equals 1 if the financially knowledgeable person of the household has a university degree.
Financial Literacy	[0;3]	Indicates the score of the financially knowledgeable person in the big three financial literacy test introduced by Lusardi and Mitchell (2008).
Household Size	[1;8]	Number of people living in the household.
Income	k€	Household's gross annual income of the household in thousands of euros.
Region East	1/0	Indicator variable that equals 1 if the household lives in one of the eastern German states: Berlin, Brandenburg, Mecklenburg-West Pomerania, Saxony, Saxony-Anhalt, or Thuringia.
Real Estate Ownership	1/0	Indicator variable equal to 1 if the household owns real estate. Ownership includes, among others: single-family homes, owner-occupied apartments, multi-family houses, commercial buildings, land plots, garages, and mixed-use properties. Partial ownership is taken into account.

Notes: This table reports the dimensions and definitions of all variables. Dimensions indicate the values the variable can take as well as in which unit they are reported.

Appendix V-B: Mean Differences Between Urban and Rural Households

Variable	Mean Urban	Mean Rural	Mean Difference
Panel A: Actual vs. Perceived Wealth			
Net Wealth	273.63	372.20	98.57**
Net Wealth Decile	4.72	6.16	1.44***
Perceived Net Wealth Decile	4.19	4.37	0.18
Net Wealth Decile Misperception	-0.53	-1.79	-1.26***
Panel B: Demographics			
Age	51.71	54.77	3.06***
Male	51.84%	57.00%	5.15%*
Household Size	1.79	2.15	0.36***
University Degree	33.77%	22.96%	-10.81%***
Financial Literacy	2.51	2.45	-0.06
Household Income	65.80	77.36	11.56**
Region East	23.02%	21.32%	-1.70%
Real Estate Ownership	31.54%	62.98%	31.44%***

Notes: This table presents the mean differences between urban and rural households. Statistical significance is indicated by *, **, and ***, corresponding to the 10%, 5%, and 1% levels. p-values for continuous variables are obtained from two-sided t-tests, while binary variables are assessed using chi-square tests of independence. All means and test statistics incorporate survey weights and are pooled across multiple imputations following Rubin's Rules.

Appendix V-C: Linear Regression on Net Wealth with Matched Sample

	Dependent Variable	
	(1) Net Wealth	(2) Net Wealth
Rural Household Indicator	68.8962	61.6838
	(45.0451)	(46.8220)
Age		6.5965***
		(0.8322)
Male		62.7086
		(38.4831)
University Degree		243.6016***
		(42.7902)
Financial Literacy		59.8087**
		(23.7002)
Household Size		107.7302***
		(17.4154)
Region East		-155.4756***
		(45.1551)
Constant	296.5941***	-493.5404***
	(38.7373)	(82.0974)
Observations	2,144	2,144
R^2	0.0025	0.1000
F-Statistic	2.3394	24.7000***

Notes: This table reports the results of our regressions on household $Net\ Wealth$, measured in thousands of euros. The analysis is based on matched data obtained through a 1:1 nearest-neighbor matching algorithm with replacement using all control variables of Panel C except $Income$ and $Real\ Estate\ Ownership$. In column (1), $Net\ Wealth$ is regressed solely on the $Rural\ Household\ Indicator$. In column (2), $Net\ Wealth$ is regressed on the $Rural\ Household\ Indicator$ alongside a set of demographic control variables, which include $Age, Gender, University\ Degree, Financial\ Literacy, Household\ Size, and\ Region\ East$. Standard errors (in parentheses) are adjusted for imputation error following Rubin's Rules as well as for survey design using replicate weights. The data are weighted according to the sampling weights provided in the dataset. Significance levels are indicated as follows: *p < .1; **p < .05; ***p < .01.

Appendix V-D: Linear Regression on Net Wealth Decile Misperception with Matched Sample

	Dependent Variable		
	(1) Net Wealth Decile Misperception	(2) Net Wealth Decile Misperception	(3) Net Wealth Decile Misperception
Rural Household Indicator	-1.0726***	-0.9184***	-0.3816**
	(0.2155)	(0.1667)	(0.1564)
Real Estate Ownership			-2.2032***
			(0.1868)
Age		-0.0312***	-0.0172***
		(0.0048)	(0.0045)
Male		-0.1735	-0.0379
		(0.1713)	(0.1513)
University Degree		0.3980**	0.5823***
		(0.2025)	(0.1856)
Financial Literacy		-0.3139**	-0.2066*
		(0.1226)	(0.1055)
Household Size		-0.2453***	-0.0843
		(0.0799)	(0.0757)
Income		0.2104**	0.2700***
		(0.1054)	(0.0830)
Net Wealth		-1.5308***	-0.9158***
		(0.2589)	(0.2001)
Region East		0.1043	-0.0691
		(0.1839)	(0.1590)
Constant	-0.7145***	1.8155***	1.3440***
	(0.1844)	(0.5405)	(0.4619)
Observations	2,186	2,186	2,186
R^2	0.0616	0.2700	0.4190
F-Statistic	32.6231***	18.6000***	39.7500***

Notes: This table reports regression results on $Net\ Wealth\ Decile\ Misperception$. The analysis is based on matched data obtained through a 1:1 nearest-neighbor matching algorithm with replacement using all control variables of Panel C as well household's $Net\ Wealth$. In column (1), $Net\ Wealth\ Decile\ Misperception$ is regressed solely on the $Rural\ Household\ Indicator$. In column (2), $Net\ Wealth\ Decile\ Misperception$ is regressed on the $Rural\ Household\ Indicator$ alongside a set of demographic control variables, which include $Age, Gender, University\ Degree, Financial\ Literacy, Household\ Size, Income, Net\ Wealth, and\ Region\ East$. The $Real\ Estate\ Ownership$ indicator is introduced as an additional control in column (3). For interpretability, we z-standardized household income and household net wealth (i.e., subtracted the mean and divided by the standard deviation) prior to including them in the model. Standard errors (in parentheses) are adjusted for imputation error following Rubin's Rules as well as for survey design using replicate weights. The data are weighted according to the sampling weights provided in the dataset. Significance levels are indicated as follows: *p < .1; **p < .05; ***p < .01.

Appendix V-E: Linear Regression on Net Wealth – Extended Sample Incl. 20–100k Municipalities

	Dependent Variable	
	(1) Net Wealth	(2) Net Wealth
Rural Household Indicator	72.3744**	53.6431
	(33.9742)	(34.8062)
Age		7.4231***
		(0.7004)
Male		74.8826**
		(32.3227)
University Degree		229.0190***
		(46.3788)
Financial Literacy		67.4635**
		(28.3370)
Household Size		87.1562***
		(17.5397)
Region East		-185.8940***
		(38.1189)
Constant	299.8263***	-496.4075***
	(23.4264)	(79.0451)
Observations	3,822	3,822
R^2	0.0027	0.0901
F-Statistic	4.5381**	32.0480***

Notes: This table reports regression results on household $Net\ Wealth$, measured in thousands of euros, using the extended sample that includes municipalities with populations between 20,000 and 100,000. In column (1), $Net\ Wealth$ is regressed solely on the $Rural\ Household\ Indicator$. In column (2), $Net\ Wealth$ is regressed on the $Rural\ Household\ Indicator$ alongside a set of demographic control variables, which include $Age, Gender, University\ Degree, Household\ Size,$ and $Region\ East$. Standard errors (in parentheses) are adjusted for imputation error following Rubin's Rules as well as for survey design using replicate weights. The data are weighted according to the sampling weights provided in the dataset. Significance levels are indicated as follows: *p < .1; **p < .05; ***p < .01.

Appendix V-F: Linear Regression on Net Wealth Decile Misperception – Extended Sample Incl. 20–100k Municipalities

	Dependent Variable		
	(1) Net Wealth Decile Misperception	(2) Net Wealth Decile Misperception	(3) Net Wealth Decile Misperception
Rural Household Indicator	-0.9062***	-0.7001***	-0.2662**
	(0.1371)	(0.1239)	(0.1072)
Real Estate Ownership			-2.3007***
			(0.1250)
Age		-0.0358***	-0.0207***
		(0.0031)	(0.0030)
Male		-0.1467	-0.0155
		(0.1065)	(0.0944)
University Degree		0.1074	0.3158**
		(0.1323)	(0.1256)
Financial Literacy		-0.3041***	-0.1738***
		(0.0817)	(0.0655)
Household Size		-0.2322***	-0.0704
		(0.0678)	(0.0605)
Income		0.0882	0.2075**
		(0.0968)	(0.0842)
Net Wealth		-1.2041***	-0.6478***
		(0.1743)	(0.1424)
Region East		0.1770	-0.0546
		(0.1344)	(0.1120)
Constant	-0.8803***	1.9053***	1.4783***
	(0.0739)	(0.3234)	(0.2789)
Observations	3,822	3,822	3,822
R^2	0.0348	0.2490	0.4170
F-Statistic	43.6845***	42.9080***	99.5960***

Notes: This table reports the results of our regressions on $Net\ Wealth\ Decile\ Misperception$, using the extended sample that includes municipalities with populations between 20,000 and 100,000. In column (1), $Net\ Wealth\ Decile\ Misperception$ is regressed solely on the $Rural\ Household\ Indicator$. In column (2), $Net\ Wealth\ Decile\ Misperception$ is regressed on the $Rural\ Household\ Indicator$ alongside a set of demographic control variables, which include Age, $Gender$, $University\ Degree$, $Financial\ Literacy$, $Household\ Size$, $Income$, $Net\ Wealth$, and $Region\ East$. The $Real\ Estate\ Ownership$ indicator is introduced as an additional control in column (3). For interpretability, we z-standardized household income and household net wealth (i.e., subtracted the mean and divided by the standard deviation) prior to including them in the model. Standard errors (in parentheses) are adjusted for imputation error following Rubin's Rules as well as for survey design using replicate weights. The data are weighted according to the sampling weights provided in the dataset. Significance levels are indicated as follows: *p < .1; **p < .05; ***p < .01.

Appendix V-G: Linear Regression on Net Wealth – Spatial Extremes Sample (<5k vs. >500k Municipalities)

	Dependent Variable	
	(1) Net Wealth	(2) Net Wealth
Rural Household Indicator	89.2407	6.8344
	(58.4442)	(63.0812)
Age		7.7593***
		(1.4447)
Male		198.9230***
		(54.0781)
University Degree		261.5450***
		(57.8275)
Financial Literacy		28.6023
		(31.3971)
Household Size		98.9553***
		(35.2889)
Region East		-203.1966***
		(74.3860)
Constant	299.4315***	-483.8081***
	(45.1865)	(114.5222)
Observations	1,193	1,193
R^2	0.0031	0.0891
F-Statistic	2.3315	11.7827***

Notes: This table reports regression results on household $Net\ Wealth$, measured in thousands of euros, using the spatial extremes sample that includes only municipalities with fewer than 5,000 inhabitants or more than 500,000 inhabitants. In column (1), $Net\ Wealth$ is regressed solely on the $Rural\ Household\ Indicator$. In column (2), $Net\ Wealth$ is regressed on the $Rural\ Household\ Indicator$ alongside a set of demographic control variables, which include Age, $Gender$, $University\ Degree$, $Household\ Size$, and $Region\ East$. Standard errors (in parentheses) are adjusted for imputation error following Rubin's Rules as well as for survey design using replicate weights. The data are weighted according to the sampling weights provided in the dataset. Significance levels are indicated as follows: *p < .1; **p < .05; ***p < .01.

Appendix V-H: Linear Regression on Net Wealth Decile Misperception – Spatial Extremes Sample (<5k vs. >500k Municipalities)

	Dependent Variable		
	(1) Net Wealth Decile Misperception	(2) Net Wealth Decile Misperception	(3) Net Wealth Decile Misperception
Rural Household Indicator	-1.3486***	-1.0466***	-0.3388
	(0.2277)	(0.2037)	(0.2107)
Real Estate Ownership			-2.0924***
			(0.2317)
Age		-0.0302***	-0.0181***
		(0.0056)	(0.0052)
Male		-0.0007	0.2066
		(0.1839)	(0.1745)
University Degree		-0.1126	0.0955
		(0.1954)	(0.1894)
Financial Literacy		-0.2228	-0.1113
		(0.1689)	(0.1447)
Household Size		-0.1890**	-0.1110
		(0.0939)	(0.0864)
Income		0.1506	0.2390
		(0.2154)	(0.2124)
Net Wealth		-0.9458**	-0.5036
		(0.3886)	(0.3496)
Region East		0.3697	0.0499
		(0.2287)	(0.2084)
Constant	-0.5840***	1.5820**	1.1753**
	(0.1316)	(0.6272)	(0.5555)
Observations	1,193	1,193	1,193
R^2	0.0853	0.2260	0.3590
F-Statistic	35.0762***	13.1830***	29.7200***

Notes: This table reports regression results on *Net Wealth Decile Misperception*, using the spatial extremes sample that includes only municipalities with fewer than 5,000 inhabitants or more than 500,000 inhabitants. In column (1), *Net Wealth Decile Misperception* is regressed solely on the *Rural Household Indicator*. In column (2), *Net Wealth Decile Misperception* is regressed on the *Rural Household Indicator* alongside a set of demographic control variables, which include *Age*, *Gender*, *University Degree*, *Financial Literacy*, *Household Size*, *Income*, *Net Wealth*, and *Region East*. The *Real Estate Ownership* indicator is introduced as an additional control in column (3). For interpretability, we z-standardized household income and household net wealth (i.e., subtracted the mean and divided by the standard deviation) prior to including them in the model. Standard errors (in parentheses) are adjusted for imputation error following Rubin's Rules as well as for survey design using replicate weights. The data are weighted according to the sampling weights provided in the dataset. Significance levels are indicated as follows: *p < .1; **p < .05; ***p < .01.

Appendix V-I: Linear Regression on Net Wealth – Single Household Sample

	Dependent Variable	
	(1) Net Wealth	(2) Net Wealth
Rural Household Indicator	15.6306	27.5884
	(35.6275)	(33.5758)
Age		5.0194***
		(0.8276)
Male		63.2347*
		(32.6761)
University Degree		204.2819***
		(45.3696)
Financial Literacy		37.3436*
		(21.8366)
Region East		-136.9586***
		(28.6174)
Constant	190.6178***	-236.3064***
	(22.6023)	(84.2770)
Observations	807	807
R^2	0.0003	0.0871
F-Statistic	0.1925	11.9302***

Notes: This table reports the results of our regressions on household $Net\ Wealth$, measured in thousands of euros, using the single-person household sample. In column (1), $Net\ Wealth$ is regressed solely on the $Rural\ Household\ Indicator$. In column (2), $Net\ Wealth$ is regressed on the $Rural\ Household\ Indicator$ alongside a set of demographic control variables, which include Age, $Gender$, $University\ Degree$, and $Region\ East$. Standard errors (in parentheses) are adjusted for imputation error following Rubin's Rules as well as for survey design using replicate weights. The data are weighted according to the sampling weights provided in the dataset. Significance levels are indicated as follows: *p < .1; **p < .05; ***p < .01.

Appendix V-J: Linear Regression on Net Wealth Decile Misperception – Single Household Sample

	Dependent Variable		
	(1) Net Wealth Decile Misperception	(2) Net Wealth Decile Misperception	(3) Net Wealth Decile Misperception
Rural Household Indicator	-0.9517***	-0.8623***	-0.4237**
	(0.2275)	(0.2109)	(0.2056)
Real Estate Ownership			-2.1385***
			(0.2669)
Age		-0.0340***	-0.0223***
		(0.0055)	(0.0051)
Male		-0.0678	0.0151
		(0.1992)	(0.1760)
University Degree		-0.2484	-0.0033
		(0.2374)	(0.2232)
Financial Literacy		-0.1794	-0.1130
		(0.1471)	(0.1293)
Income		0.0312	0.1301
		(0.2651)	(0.2124)
Net Wealth		-0.9754**	-0.4329
		(0.3876)	(0.2826)
Region East		0.0758	-0.0368
		(0.2221)	(0.2007)
Constant	-0.2635**	1.9421***	1.6569***
	(0.1300)	(0.6130)	(0.5394)
Observations	807	807	807
R^2	0.0433	0.2720	0.4080
F-Statistic	17.4975***	14.4010***	35.7840***

Notes: This table reports the results of our regressions on $Net\ Wealth\ Decile\ Misperception$, using the single-person household sample. In column (1), $Net\ Wealth\ Decile\ Misperception$ is regressed solely on the $Rural\ Household\ Indicator$. In column (2), $Net\ Wealth\ Decile\ Misperception$ is regressed on the $Rural\ Household\ Indicator$ alongside a set of demographic control variables, which include Age, $Gender$, $University\ Degree$, $Financial\ Literacy$, $Income$, $Net\ Wealth$, and $Region\ East$. The $Real\ Estate\ Ownership$ indicator is introduced as an additional control in column (3). For interpretability, we z-standardized household income and household net wealth (i.e., subtracted the mean and divided by the standard deviation) prior to including them in the model. Standard errors (in parentheses) are adjusted for imputation error following Rubin's Rules as well as for survey design using replicate weights. The data are weighted according to the sampling weights provided in the dataset. Significance levels are indicated as follows: $^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$.

Appendix V-K: Linear Regression on Net Wealth Decile Misperception – Interquartile Sample

	Dependent Variable		
	(1) Net Wealth Decile Misperception	(2) Net Wealth Decile Misperception	(3) Net Wealth Decile Misperception
Rural Household Indicator	-1.0833***	-0.4545***	-0.4114***
	(0.1706)	(0.1366)	(0.1367)
Real Estate Ownership			-0.2762*
			(0.1621)
Age		-0.0087**	-0.0077*
		(0.0039)	(0.0040)
Male		0.1981	0.1914
		(0.1240)	(0.1236)
University Degree		0.9564***	0.9430***
		(0.1567)	(0.1601)
Financial Literacy		-0.0436	-0.0450
		(0.1022)	(0.1017)
Household Size		-0.1184	-0.0987
		(0.0807)	(0.0818)
Income		0.4332***	0.4264***
		(0.0943)	(0.0937)
Net Wealth		-1.1359***	-1.0795***
		(0.0770)	(0.0863)
Region East		-0.5339***	-0.5259***
		(0.1453)	(0.1482)
Constant	-1.4535***	-1.5413***	-1.4428***
	(0.1451)	(0.4250)	(0.4313)
Observations	1,343	1,343	1,343
R^2	0.0713	0.4300	0.4320
F-Statistic	40.3265***	38.9900***	36.4830***

Notes: This table reports regression results on $Net\ Wealth\ Decile\ Misperception$ using the interquartile sample, which excludes households in the top and bottom quartile of the wealth distribution. In column (1), $Net\ Wealth\ Decile\ Misperception$ is regressed solely on the $Rural\ Household\ Indicator$. In column (2), $Net\ Wealth\ Decile\ Misperception$ is regressed on the $Rural\ Household\ Indicator$ alongside a set of demographic control variables, which include Age, $Gender$, $University\ Degree$, $Financial\ Literacy$, $Household\ Size$, $Income$, $Net\ Wealth$, and $Region\ East$. The $Real\ Estate\ Ownership$ indicator is introduced as an additional control in column (3). For interpretability, we z-standardized household income and household net wealth (i.e., subtracted the mean and divided by the standard deviation) prior to including them in the model. Standard errors (in parentheses) are adjusted for imputation error following Rubin's Rules as well as for survey design using replicate weights. The data are weighted according to the sampling weights provided in the dataset. Significance levels are indicated as follows: $^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$.

VI. The Power of ESG Transparency: The Effect of the New SFDR Sustainability Labels on Mutual Funds and Individual Investors

Co-authors:

Fabio Martin, Andreas Walter

Own share:

45%

This article has been published as:

Becker, M. G., Martin, F., & Walter, A. (2022). The power of ESG transparency: The effect of the new SFDR sustainability labels on mutual funds and individual investors. *Finance Research Letters*, *47*, 102708. https://doi.org/10.1016/j.frl.2022.102708

Previous versions of this paper have been presented at the following conferences and workshops:

- 2023 FDIR Research Initiative (Ecole polytechnique and Toulouse School of Economics)

The Power of ESG Transparency: The Effect of the New SFDR Sustainability Labels on Mutual Funds and Individual Investors

Martin G. Becker[a] Fabio Martin[b] Andreas Walter[c]

Abstract: This paper analyzes the effect of the Sustainable Finance Disclosure Regulation (SFDR) on mutual funds and individual investors in the EU. First, we study whether affected funds increase their sustainability compared to a control group. Second, we examine if the regulation makes individual investors allocate more capital into more sustainable funds. In a difference-indifferences setting, we analyze the influence of the regulation on ESG fund scores and fund net inflows. Our results show that affected funds increase their sustainability rating after the policy intervention. Additionally, we find that a better ESG label leads to larger fund net inflows.

JEL Codes: G11, G18, G23

Keywords: Mutual fund, ESG, Sustainability ratings, Fund flows, Policy intervention

[a] martin.becker@wirtschaft.uni-giessen.de, Justus Liebig University Gießen, Department of Financial Services, Licher Straße 74, 35394 Gießen, Germany

[b] fabio.martin@wirtschaft.uni-giessen.de, Justus Liebig University Gießen, Department of Financial Services, Licher Straße 74, 35394 Gießen, Germany

[c] andreas.walter@wirtschaft.uni-giessen.de, Justus Liebig University Gießen, Department of Financial Services, Licher Straße 74, 35394 Gießen, Germany

1. Introduction

On the 27th of November 2019, the European Parliament and Council published the Regulation (EU) 2019/2088 on sustainability related disclosure in the financial services sector (SFDR) being effective as of March 10, 2021. The preamble of the regulation states that to fight climate change "urgent action is needed to mobilize capital not only through public policies but also by the financial services sector". Introducing this new kind of regulation, the EU tries to change behavior patterns in the financial sector, discouraging greenwashing, and promoting responsible and sustainable investments.

The new policy applies to all European financial market participants (FMPs). These include investment firms, pension providers, and insurance-based investors, as well as qualifying venture capital and social entrepreneurship activities. Besides the increasing reporting duty, one of its main requirements for the FMPs is the classification of ESG-related products and non-ESG products as either article 6, 8 or 9 funds depending on the degree of ESG integration.[1]

Here, article 8 comprise those funds that do consider ESG aspects in their investment process but are focused on financial materiality, whereas article 9 products aim to create an environmental and social impact alongside generating a financial return. This can usually be done by aligning the portfolio to the UN Sustainable Development Goals (SDGs) or to the Paris Agreement. More specifically, article 8 applies where "a financial product promotes, among other characteristics, environmental or social characteristics, or a combination of those characteristics, provided that the companies in which the investments are made follow good governance practices" (Regulation (EU) 2019/2088). In contrast, article 9 refers to funds which have generating a real impact as their primary goal alongside a financial return. Finally, article 6 products are those funds which do not fulfill the requirements to be labeled as article 8 or 9 and thus represent all funds that do not or to a very low degree integrate sustainability in their

[1] The name of the fund categories refer to the respective articles in the sustainable finance disclosure regulation.

investment process. Moreover, while article 7 of the SFDR deals with pre-contractual disclosures, it does not constitute its own classification.

The goal of this paper is to analyze the effect of the policy intervention on achieving the stated goal of fighting climate change by incentivizing FMPs to become more green and by mobilizing capital in the financial services sector. To analyze the extent to which this goal has been accomplished we study both, the demand and supply side. Thus, we examine how mutual fund managers and private customers react to the policy intervention, respectively.

If investors are better informed about the sustainability of funds, this creates an incentive for funds to invest in more sustainable ways (Hartzmark & Sussman, 2019). Firstly, we expect an increase in the sustainability scores of funds that are affected by the intervention compared to the unaffected funds after the public disclosure of the upcoming regulation in December 2019.

Following Mugerman et al. (2022) we use a difference-in-differences methodology as our main identification strategy and divide our sample into two groups. The first one, being the treatment group, contains all European funds that are affected by the regulation. For our control group, we use all U.S. based mutual funds since they are not exposed to the intervention and represent the largest part of mutual funds worldwide. In order to control for potential differences between the treatment and the control group, we use a 1:1 nearest neighbor matching (Ammann et al., 2019; Bilbao-Terol et al., 2017). We then estimate if European funds increased their ESG scores relatively to the control group as a result of the SFDR.

The demand side then implies that an increase in transparency and sustainability leads to more inflows towards sustainable funds (Alda, 2020; Ghoul & Karoui, 2021). The disclosure of being an article 6, 8 or 9 fund could directly influence the investment decisions of private customers. Huang et al. (2020) show that funds which are being given some sort of performance label should experience an increase in their inflows due to jumps in reputation. However, this depends on how investment firms promote their labels. Again, we examine this in a difference-in-differences setting to analyze whether the intervention had a significant impact on the fund net inflows within the first four months after March 10, 2021, the day on which the funds label were first publicly disclosed.

Overall, our results are consistent with the literature: For the supply side, we find that the increasing transparency of sustainability enforced by the new regulation incentivize mutual funds to increase their ESG efforts. EU funds, which are affected by the new SFDR rule increased their ESG scores more than funds in the non-EU control group.

For the demand side, our results indicate that the intervention had a statistically significant impact on the fund flows within the first four months after the intervention. Article 8 and 9 funds did see positive net inflows compared with less sustainable EU funds. This is in accordance with Ghoul and Karoui (2021), Aasheim et al. (2022), Ammann et al. (2019) and Huang et al. (2020) who show that funds which are associated with a higher ESG alignment attract higher inflows from investors.

Our paper makes two contributions to the mutual fund literature. First, we add to the scarce literature of policy interventions and their effect on capital markets. Zhang et al. (2021) examine the impact of the implementation of “Guidelines for Establishing a Green Financial System” in China and show that afterwards the risk-adjusted return for the highest ESG portfolio nearly doubles.

Second, we contribute to the literature on the relationship between fund flows and sustainability. Ammann et al. (2019) examine the effect of the introduction of Morningstar’s Sustainability Rating on mutual fund flows. They find strong evidence that retail investors shift money away from low-rated and into high-rated funds. Ghoul and Karoui (2021) show that funds which have changed their names to a sustainability-related appellation exhibit larger inflows. Alda (2020) show that a higher ESG screening intensity triggers larger inflows. Ceccarelli et al. (2019) find that active funds which missed a “low carbon designation” label by Morningstar at its release, shifted their holdings towards less carbon-intensive firms. Lastly, Rzeźnik et al. (2021) show that some investors buy assets after a misconceived ESG score upgrade. This is evidence for the fact that it is not the true sustainability that seems to matter but only the ESG label. Therefore, we extend the literature by examining the unique setting of a policy intervention and its effects on mutual funds.

2. Data

Our study is based on 9,722 EU mutual funds and 15,896 U.S. funds for the period between September 2019 and June 2021.[2] While we are mostly interested in the effect of the regulation on European mutual funds, the U.S. data is used as a control group lacking the policy intervention. We gather data from the Morningstar database on the portfolio (monthly) ESG scores as well as the (monthly) fund size and the inception date. Following Ammann et al. (2019), we analyze the sustainability of a mutual fund using the Morningstar Sustainability Rating. The Morningstar Sustainability Rating is being calculated based on the individual securities in each fund. In doing so, Morningstar evaluates how well an issuer manages environmental, social and governance risks and opportunities. The rating of the fund is then calculated based on a peer group comparison. It ranks mutual funds on a scale from one (worst) to five (best) within their global category. Further, Morningstar also provides data on each fund's SFDR classification. Here, the funds are either labeled as an article 8 fund, article 9 fund or not classified (i.e., article 6). Similarly to other studies, we retrieve data on our control variables – fund age, fund size, total net asset values and returns – and drop all observations with missing data (e.g. Alda (2020), Barber et al. (2005) and Morey (2002)).

[2] In September 2019 Morningstar changed how the sustainability rating is calculated (Ferriani & Natoli, 2021). Rzeźnik et al. (2021) show that investors misconceived the changes in the Morningstar methodology. To ensure our results are not driven by the change in methodology, we exclude data before September 2019.

Table VI-1: Summary Statistics

Panel A: EU vs. US						
EU funds (All)						
	No. of Obs.	Mean	Median	SD	Min	Max
Sustainability rating	92,310	3.29	3.00	1.07	1.00	5.00
Fund age (in Months)	92,310	160	145	108	2	439
Fund size (in Million Euro)	92,310	489	138	1,191	1	30,593
Total returns (in %)	92,310	1.57	1.89	4.94	−17.33	13.26
Fund flows (in % of TNA)	92,310	0.22	−0.35	9.52	−30.59	62.01
US funds (All)						
Sustainability rating	94,173	2.96	3.00	0.98	1.00	5.00
Fund age (in Months)	94,173	147	126	99	2	439
Fund size (in Million Euro)	94,173	621	295	1,195	1	40,449
Total returns (in %)	94,173	1.75	2.40	5.78	−17.33	13.26
Fund flows (in % of TNA)	94,173	−0.35	−0.86	8.97	−30.59	62.01
Panel B: EU						
EU funds (Article 6)						
	No. of Obs.	Mean	Median	SD	Min	Max
Sustainability rating	32,835	3.12	3.00	1.05	1.00	5.00
Fund age (in Months)	32,835	155	140	106	2	421
Fund size (in Million Euro)	32,835	500	159	865	1	4,594
Total Returns (in %)	32,835	1.38	1.95	5.58	−22.19	10.18
Fund flows (in % of TNA)	32,835	0.00	−0.45	9.14	−28.48	53.44
EU funds (Article 8)						
Sustainability rating	29,960	3.58	4.00	1.01	1.00	5.00
Fund age (in Months)	29,960	160	140	111	2	421
Fund size (in Million Euro)	29,960	591	213	931	1	4,594
Total returns (in %)	29,960	1.51	2.05	5.55	−22.19	10.18
Fund flows (in % of TNA)	29,960	0.46	−0.22	9.32	−28.48	53.44
EU funds (Article 9)						
Sustainability rating	2,799	4.01	4.00	0.89	1.00	5.00
Fund age (in Months)	2,799	138	119	103	4	421
Fund size (in Million Euro)	2,799	571	265	827	1	4,594
Total returns (in %)	2,799	1.59	2.21	5.67	−22.19	10.18
Fund flows (in % of TNA)	2,799	1.96	0.83	11.01	−28.48	53.44

Notes: This table reports summary statistics of the monthly portfolio sustainability scores as well as on the different fund characteristic measures. All control variables are winsorized at the 1% and 99% level.

Following Sirri and Tufano (1998) net mutual inflows are calculated as the growth in total assets reduced by the monthly returns as a percentage of total net assets at the beginning of the previous month:

$$Flow_{i,t} = \frac{TNA_{i,t} - TNA_{i,t-1}\left(1 + R_{i,t}\right)}{TNA_{i,t-1}} \tag{VI-1}$$

Whereby $TNA_{i,t}$ indicates the total net assets of a given fund i at the end of month t and $R_{i,t}$ is the return of fund during the month t. Since we are comparing EU to U.S funds as well as fund flows among the different SFDR categories, we create two different samples where the first one contains all U.S and EU and the second one all EU funds that have a SFDR classification in Morningstar. All control variables in both samples and the fund flows are winsorized at the 1% and 99% level.

Finally, we apply the 1:1 nearest-neighbor matching method from Rubin (1973). This matches funds from our control groups to the treated funds using fund age, fund size, fund returns and fund net inflows. Non-matched funds from the control group are removed from the sample. This ensures that we avoid any bias from inadequate comparison and improves parametric statistical models (Ammann et al., 2019; Bilbao-Terol et al., 2017; Joliet & Titova, 2018). Table VI-1 offers an overview of the summary statistics. The mean sustainability rating for EU funds in the sample is 0.43 higher than for U.S. funds. Also, as expected, EU funds that are classified as article 9 have the highest sustainability rating while article 6 funds have the lowest one within the EU sample.

3. The Influence of the EU Directive on Sustainability Scores

First, we examine the introduction of the new SFDR policy and its effect on the sustainability rating of the affected funds by estimating the following model:

$$\begin{aligned} ESG_{i,t} = \; & \beta_0 * Treated_i + \beta_1 * Post_t + \beta_2 * Treated_i \; x \; Post_t + \beta_3 * Size_{i,t-1} \\ & + \beta_4 * Age_{i,t-1} + \beta_5 * Ret_{i,t-1} + \beta_6 * Flow_{i,t-1} + \mu_i \end{aligned} \tag{VI-2}$$

where $ESG_{i,t}$ describes the sustainability rating of fund i at month t. The dummy $Treated_i$ equals one if fund i is a EU-based mutual fund and thus affected by the SFDR. The dummy $Post_t$ equals one for all months after November 2019. November 2019 marks the date when the European Commission passed the new regulation. Thus, FMPs had time since 2019 to adjust their portfolios and make them more ESG aligned, whereas customers had no information about the fund labels before March 2021. $Size_{i,t-1}$ are the total net assets of fund i at month $t-1$; $Age_{i,t-1}$ describes the total months between $t-1$ and the inception date of fund i; $Flows_{i,t-1}$ are the flows of fund i at month $t-1$; and $Ret_{i,t-1}$ is the return of fund i and month $t-1$. Our difference-in-differences estimator $Treated_i \times Post_t$ indicates observations for EU funds in the period after the introduction of the SFDR policy. In addition, we use fund fixed effects μ_i to control for any time-invariant effects and estimate our model using fund-clustered standard errors.

Table VI-2 reports the results for Eq. (2). The interaction term in column (2) shows that the intervention achieved its desired effect. The ESG rating for European mutual funds significantly rose after the announcement of the SFDR regulation relatively to the U.S peers. The average difference in fund ratings between EU and U.S. funds rose by nearly 0.03 rating grades. This is in accordance with our initial expectations since mutual funds might anticipate higher fund inflows if they are publicly being labeled as a green investment. Hence, EU funds increase their sustainability level more than U.S. funds. Further, column (1) shows that the average EU funds score is 0.313 higher than for US funds while the average base level in the sample is about 2.99.

Table VI-2: The Influence of the EU Directive on Sustainability Scores

Dependent var.	(1) ESG fund rating	(2) ESG fund rating
Intercept	2.992***	
	(102.00)	
Treated	0.313***	
	(24.52)	
Post	−0.017*	−0.003
	(−1.74)	(−0.48)
Treated x Post	0.019	0.027**
	(1.37)	(2.62)
Controls	Yes	Yes
Fixed effects (Fund)	No	Yes
R^2 adj.	0.027	0.001
Observations	186,483	186,483

Notes: The dummy $Treated$ takes the value one for all EU funds and zero otherwise. The dummy $Post$ indicate the time period after November 2019. T-statistics (in parentheses) are based on fund-clustered standard errors. ***, **, * denote statistical significance at the 1%, 5% and 10% level, respectively. Significance levels are calculated using fund-clustered standard errors.

4. Impact on Investors

After analyzing the supply side, we now shift our focus onto the demand side. Huang et al. (2020) and Ammann et al. (2019) show that mutual funds with a better rating or label attract more inflows compared to less sustainable funds. Consequently, funds that received the article 8 or article 9 label should experience higher net inflows compared to article 6 funds. Again, we use the 1:1 nearest-neighbor matching to create three samples: for column (1), we exclude all article 9 funds and match each article 8 fund with the respective article 6 fund; for column (2) we exclude all article 8 funds and match each article 9 fund with the respective article 6 and for column (3) we match all article 8 and 9 funds with article 6 funds. In all specifications, unmatched article 6 funds are removed from the sample. We examine this hypothesis by estimating the following model:

$$
\begin{aligned}
Flow_{i,t} = {} & \beta_0 * Treated_i + \beta_1 * Post_t + \beta_2 * Treated_i \; x \; Post_t \\
& + \beta_3 * Size_{i,t-1} + \beta_4 * Age_{i,t-1} + \beta_5 * Ret_{i,t-1} + \mu_i
\end{aligned} \qquad \text{(VI-3)}
$$

with $Flows_{i,t}$ being the net flows of fund i in month t. $Post_t$ now marks the effective date of the intervention, i.e. all observations beginning with March 2021. We use this date since the sustainability labels were not disclosed before March 2021 and thus customers could not take it into account when allocating their money. In column (1), (2) and (3) the dummy $Treated_i$ takes the value one for all funds classified as article 8, article 9 and article 8 or 9, respectively and zero otherwise. Control variables remain unchanged. We again use fund fixed effects μ_i and fund-clustered standard errors.

Table VI-3 displays the results of the difference-in-differences estimation. The interaction term in column (3) shows that funds which were labeled as either article 8 or 9 were able to significantly increase their net fund flows after the intervention. In particular, more sustainable funds are able to generate 0.5 percentage points per month more inflows than less sustainable funds.

Table VI-3: The Influence of the EU Directive on Fund Flows

	(1) Art. 8	(2) Art. 9	(3) Art. 8 & 9
Dependent var.	Fund flows	Fund flows	Fund flows
Post	0.017***	0.019***	0.017***
	(10.56)	(3.57)	(11.09)
Treated x Post	0.005***	0.006	0.005***
	(2.69)	(0.89)	(2.93)
Controls	Yes	Yes	Yes
Fixed effects (Fund)	Yes	Yes	Yes
R^2 adj.	0.020	0.012	0.018
Observations	60,013	5,581	65,594

Notes: In columns (1), (2) and (3) the dummy $Treated$ takes the value one for all funds classified as article 8, article 9 and article 8 or 9, respectively and zero otherwise. The dummy $Post$ indicates the time period beginning with March 2021. T-statistics (in parentheses) are based on standard errors. ***, **, * denote statistical significance at the 1%, 5% and 10% level, respectively. Significance levels are calculated using fund-clustered standard errors.

The significance of the interactions terms in columns (1) and (2) indicate that this result might be largely driven by article 8 funds given the smaller sample size of article 9 funds. This is in line with the findings of Aasheim et al. (2022) or Ammann et al. (2019) who find abnormal flows of 1.83% during the first 6 months after the publication of Morningstar sustainability fund ratings. This supports our hypothesis that an increase in transparency and sustainability will lead to more sustainable investments. However, it is not entirely clear how much of the effect can be attributed to increase in transparency or sustainability. Rzeźnik et al. (2021) show that investors care more about the label itself than the actual degree of ESG integration.

5. Conclusion

In this paper, we study the impact of the SFDR – a legislation regarding sustainability disclosure for mutual funds – on the sustainability and fund flows of mutual funds. Using difference-in-differences regressions and 1:1 nearest neighbor matching, we compare funds affected by the legislation (EU-based funds) with unaffected funds (U.S.-based funds). Our results show a significantly higher increase in sustainability ratings for the EU-based funds after the announcement of the SFDR. This shows that, on the fund level, the intervention so far achieved its purpose of moving capital into more sustainable investments. To observe whether the legislation also has an impact on investors we investigated the changes in fund flows of different sustainability classifications introduced by the SFDR. Here, we find that investors appreciate a higher degree of ESG alignment and allocate their capital accordingly. Funds with classifications indicating a more advanced level sustainability integration experience significantly higher net fund flows after the public disclosure of fund labels. Our findings have direct implications for investors and practitioners. First, due to investors investing more in article 8 and article 9 funds, asset manager should increase their sustainability efforts according to article 8 and article 9 of the SFDR. Second, it is likely that the newly introduced labels will increase the threat of a possible ESG overvaluation (Bofinger et al., 2022). An increasing amount in indications for sustainable investments can potentially lead to even higher investments towards overvalued firms. In summary, our study shows the effectiveness of the newly introduced regulation on sustainability-related disclosures in the financial services sector and points towards the SFDR mobilizing capital towards sustainable investments.

6. References

Aasheim, L. K., Miguel, A. F., & Ramos, S. B. (2022). Star rating, fund flows and performance predictability: Evidence from Norway. *Financial Markets and Portfolio Management*, *36*(1), 29–56. https://doi.org/10.1007/s11408-021-00390-8

Alda, M. (2020). ESG fund scores in UK SRI and conventional pension funds: Are the ESG concerns of the SRI niche affecting the conventional mainstream? *Finance Research Letters*, *36*, 101313. https://doi.org/10.1016/j.frl.2019.101313

Ammann, M., Bauer, C., Fischer, S., & Müller, P. (2019). The impact of the Morningstar Sustainability Rating on mutual fund flows. *European Financial Management*, *25*(3), 520–553. https://doi.org/10.1111/eufm.12181

Barber, B. M., Odean, T., & Zheng, L. (2005). Out of sight, out of mind: The effects of expenses on mutual fund flows. *The Journal of Business*, *78*(6), 2095–2120. https://doi.org/10.1086/497042

Bilbao-Terol, A., Álvarez-Otero, S., Bilbao-Terol, C., & Cañal-Fernández, V. (2017). Hedonic evaluation of the SRI label of mutual funds using matching methodology. *International Review of Financial Analysis*, *52*, 213–227. https://doi.org/10.1016/j.irfa.2017.06.003

Bofinger, Y., Heyden, K. J., & Rock, B. (2022). Corporate social responsibility and market efficiency: Evidence from ESG and misvaluation measures. *Journal of Banking & Finance*, *134*, 106322. https://doi.org/10.1016/j.jbankfin.2021.106322

Ceccarelli, M., Ramelli, S., & Wagner, A. F. (2019). When investors call for climate responsibility, how do mutual funds respond? *SSRN Electronic Journal.* Advance online publication. https://doi.org/10.2139/ssrn.3353239

Ferriani, F., & Natoli, F. (2021). ESG risks in times of COVID-19. *Applied Economics Letters*, *28*(18), 1537–1541. https://doi.org/10.1080/13504851.2020.1830932

Ghoul, S. E., & Karoui, A. (2021). What's in a (green) name? The consequences of greening fund names on fund flows, turnover, and performance. *Finance Research Letters*, *39*, 101620. https://doi.org/10.1016/j.frl.2020.101620

Hartzmark, S. M., & Sussman, A. B. (2019). Do investors value sustainability? A natural experiment examining ranking and fund flows. *The Journal of Finance*, *74*(6), 2789–2837. https://doi.org/10.1111/jofi.12841

Huang, C., Li, F. E., & Weng, X. I. (2020). Star ratings and the incentives of mutual funds. *The Journal of Finance*, *75*(3), 1715–1765. https://doi.org/10.1111/jofi.12888

Joliet, R., & Titova, Y. (2018). Equity SRI funds vacillate between ethics and money: An analysis of the funds' stock holding decisions. *Journal of Banking & Finance*, *97*, 70–86. https://doi.org/10.1016/j.jbankfin.2018.09.011

Morey, M. R. (2002). Mutual fund age and Morningstar ratings. *Financial Analysts Journal*, *58*(2), 56–63. https://doi.org/10.2469/faj.v58.n2.2523

Mugerman, Y., Steinberg, N., & Wiener, Z. (2022). The exclamation mark of Cain: Risk salience and mutual fund flows. *Journal of Banking & Finance*, *134*, 106332. https://doi.org/10.1016/j.jbankfin.2021.106332

Rubin, D. B. (1973). Matching to remove bias in observational studies. *Biometrics*, *29*(1), 159. https://doi.org/10.2307/2529684

Rzeźnik, A., Hanley Kathleen W., & Pelizzon, L. (2021). The salience of ESG ratings for stock pricing: Evidence from (potentially) confused investors.

Sirri, E. R., & Tufano, P. (1998). Costly search and mutual fund flows. *The Journal of Finance*, *53*(5), 1589–1622. https://doi.org/10.1111/0022-1082.00066

Zhang, X., Zhao, X., & Qu, L. (2021). Do green policies catalyze green investment? Evidence from ESG investing developments in China. *Economics Letters*, *207*, 110028. https://doi.org/10.1016/j.econlet.2021.110028

VII. When a B Becomes an A: Causal Evidence on the Effects of a Journal Ranking Update on Academics' Publication Behavior

Co-authors:

Mario Fernandes, Kerstin Pull, Andreas Walter

Own share:

30%

This article has been published as:

Fernandes, M., Becker, M. G., Pull, K., & Walter, A. (2025). When a B becomes an A: causal evidence on the effects of a journal ranking update on academics' publication behavior. *Studies in Higher Education*, 1-22. https://doi.org/10.1080/03075079.2024.2447788

When a B Becomes an A:
Causal Evidence on the Effects of a Journal Ranking Update on Academics' Publication Behavior

Mario Fernandes[a] Martin G. Becker[b] Kerstin Pull[c] Andreas Walter[d]

Abstract: Exploiting a quasi-natural experiment and using a difference-in-differences approach, we provide causal evidence on academics' response to the update of a survey-based journal ranking. Focusing on business researchers in Germany, Austria and Switzerland, we find that academics who are affected by the journal ranking shift their publication strategies, publishing relatively more in journals that have been upgraded than in journals that have retained their rank. More specifically, affected academics shift to journals that have been upgraded from B to A and less to those that have been upgraded from C to B. Younger and more actively publishing academics are more likely to react to journal upgrades. By identifying academics' causal response to a journal ranking update, our study highlights the power of incentives in academia and allows us to derive important implications on the governance of higher education systems.

JEL Codes: A14, I23, J45

Keywords: difference-in-differences, journal rankings, incentives in academia, publication strategies, academic publishing

[a] mario.fernandes-2@wirtschaft.uni-giessen.de, Justus Liebig University Gießen, Department of Financial Services, Licher Straße 74, 35394 Gießen, Germany

[b] martin.becker@wirtschaft.uni-giessen.de, Justus Liebig University Gießen, Department of Financial Services, Licher Straße 74, 35394 Gießen, Germany

[c] kerstin.pull@uni-tuebingen.de, Eberhard Karls University of Tübingen, School of Business and Economics, Nauklerstraße 47, 72074 Tübingen, Germany

[d] andreas.walter@wirtschaft.uni-giessen.de, Justus Liebig University Gießen, Department of Financial Services, Licher Straße 74, 35394 Gießen, Germany

1. Introduction

'An A is an A' – this is how Aguinis et al. (2020) describe that publishing in highly ranked journals has become 'the new bottom line for valuing academic research'. But what happens when 'a B becomes an A'? This is what we study in our paper. More specifically, we analyze whether and to what extent academics change their publication behavior in response to the update of a journal ranking.

Journal rankings are used to assess the value of research, to inform hiring or tenure committees or to allocate resources (see, e.g., Aguinis et al. (2020); Anderson et al. (2022); Heckman and Moktan (2020); Hudson (2024); Rafols et al. (2012); Walker, Fenton, et al. (2019)). As a result, publications in highly ranked journals have become 'the currency' for the evaluation of academics, departments, and universities (Osterloh & Frey, 2020). Yet, the adequacy of journal rankings to assess the quality of the research published therein and the resulting pressure to publish in highly ranked journals is subject to considerable debate (see, e.g., Aguinis et al. (2020); Anninos (2014); Heckman and Moktan (2020); Mingers and Willmott (2013); Osterloh and Frey (2020); Salandra et al. (2022)) with critics faulting the 'fetishism' (Willmott, 2011) and 'seductive power' (Nkomo, 2009) of journal rankings.

In our paper, we shed more light on the effects of journal rankings by analyzing potential shifts in academics' publication behavior following the update of a journal ranking that is only relevant for some academics, but not for others. Theoretically, when for some academics a B becomes an A, these academics will publish relatively more in the respective journal, because – for them and not for others – publishing in this journal is linked to tangible and intangible benefits: e.g., direct monetary payments (Osterloh & Frey, 2020), enhanced career perspectives (Bajo et al., 2020; Heckman & Moktan, 2020) and/or prestige (Roach & Sauermann, 2010; Sauermann & Stephan, 2013; Stern, 2004). Following a journal ranking update, we thus expect affected academics (in comparison to non-affected ones) to publish relatively more in journals that have been upgraded than in journals that retained their rank – at least if the costs of publishing in the upgraded journals did not increase to the same degree at the same time. While academics' reactions to journal ranking updates have been analyzed before (e.g., Śpiewanowski and Talavera (2021) and Hudson (2024)), we are the first to causally identify academics' reactions to a journal ranking update by exploiting a quasi-natural experiment and employing a difference-in-differences approach. This causal identification is crucial because it allows for an unbiased assessment of affected academics' behavioral response – which then delivers the basis on which to derive well-grounded implications for policy makers and university leaders.

In our empirical analyses, we build on a setting which proves ideal for the analysis of our research question. More specifically, we focus on business researchers in Austria, Germany, and Switzerland and on the update of the journal ranking issued by the German Academic Association for Business Research (VHB) in early 2015. We choose this setting as it allows for a causal analysis of the shifts in academics' publication behavior following a journal ranking update. More specifically, we exploit a quasi-natural experiment and use a difference-in-differences approach to analyze how the update of the survey-based VHB-journal-ranking influences affected academics' publication behavior such that affected academics – in comparison to unaffected ones – publish relatively more in journals whose rank improved in the update than in journals whose rank remained constant.

Our empirical findings confirm our prediction: Following the update of the VHB-journal-ranking, affected academics in German-speaking countries as compared to academics from other countries publish relatively more in journals that were upgraded than in journals whose rank did not change. Other than the preceding literature, we causally identify this effect. Exploring potential heterogeneities at the level of journal ranks, we find our results to be driven by journals that were upgraded from B to A. Additionally exploring potential heterogeneities at the level of individual researchers by using a second, hand-collected data set, we find younger and more actively publishing researchers to react more strongly to the journal update.

To conclude, our study offers three major contributions. First, we causally identify the effects of a change in the incentives to publish in one journal versus another on the publication strategies of affected as opposed to unaffected academics, thus detecting strategic shifts in academics' publication behavior. Our results are strengthened by placebo tests where we do not find any such strategic shifts for researchers that are affiliated to U.S. (or French) institutions and who are arguably unaffected by the ranking update under consideration. Second, we find this effect to be driven by academics publishing relatively more in journals that were upgraded from B to A and not so much by academics publishing relatively more in journals that were upgraded from C to B. Third, we find younger and more actively publishing researchers to react more strongly to the journal update than older and less active ones.

In sum, our study highlights the power of journal rankings in guiding academics' publication behavior and has important practical implications: If changes in a journal ranking do have sizeable effects on affected academics' publication behavior, then rankings should be very carefully designed and not be taken as the sole indicator of the value of academic research. Rather, hiring and tenure committees should refrain from an 'excessive attention to journal lists' (Aguinis et al., 2020), in particular survey-based ones, but instead assess the works of the academics who apply for a position or who are evaluated for tenure without inferring their merits from the rank of the journal that their work was published in.

Our paper proceeds as follows. In Section 2, we describe our setting. In Section 3, we discuss the effects of a journal upgrade and introduce our hypothesis. In Section 4, we describe our data and methodology. Section 5 investigates the causal effect of a journal ranking update on affected academics' publication behavior. In Section 6, we conduct a set of robustness tests and provide further analyses. Section 7 offers the discussion of our findings and concludes.

2. Setting and Identification

Our setting is based on the journal ranking issued by the German Academic Association for Business Research (VHB): the VHB-journal-ranking. With its almost 3,000 members, most business researchers in Austria, Germany, and Switzerland are members of the association. While there are no comprehensive studies examining the utilization of journal rankings in evaluating business researchers in German-speaking countries (as, e.g. Walker, Fenton, et al. (2019) or Heckman and Moktan (2020) provide them for the U.K. or the U.S., respectively), several authors have underscored the significance of the VHB-journal-ranking for business researchers in German-speaking countries (e.g., Eisend (2011) or Schrader and Hennig-Thurau (2009)). Specifically, tenure and appointment committees in Austria, Germany, and Switzerland regularly use the VHB ranking to comparatively assess academics' publication portfolios, and business academics applying for a position in the respective countries often explicitly include a journal's VHB-journal-rank in their publication lists (M. Fernandes & Walter, 2022).[1] In addition, a prominent German business magazine (i.e., Wirtschaftswoche) uses the VHB-journal-ranking to construct a ranking of individual business researchers in the German-speaking countries. This ranking is also freely accessible on the web.[2] Hence, it can safely be assumed that the VHB-journal-ranking directly impacts the career progression of business researchers in German speaking countries and that business researchers in the respective countries are aware of its relevance and have been socialized to recognize it as a career-relevant authority. To study our research questions, we refer to the update in the VHB-journal-ranking that was released in early 2015, when the former VHB-journal-ranking (Jourqual 2) was replaced by its third version VHB-journal-ranking (Jourqual 3). The last larger overhaul prior to the update in 2015 took place in 2008. It is only very recently (April 2024) that the results of a new survey on perceived journal quality have been published by the

[1] At the time of our data, the San Francisco Declaration on Research Assessment (DORA), which denounces the practice of using journal impact factors as a substitute for measuring the quality of individual research in hiring, promotion, or funding decisions, had not yet gained influence. Rather, DORA had only been published in May 2013 and had at that time been signed by only 75 scientific organizations. The German Research Foundation, for example, ratified DORA in 2021, which was after the period covered in our data. Additionally, a recent study by Morgan-Thomas et al. (2024) finds for the U.K. that journal rankings remain strongly predictive of expert evaluations of research, despite institutional commitments to the principles of DORA.

[2] https://www.forschungsmonitoring.org/ranking/bwl/authors

association. However, the association changed the construction of the ranking, e.g., one journal might now have different ratings for different fields of specialization within business studies. Therefore, these ratings cannot be used to derive a new comprehensive VHB-journal-ranking that could form the basis of our analysis. The VHB-journal-ranking and its updates are survey-based, i.e., ranks are allocated to journals based on a survey among the members of the association.[3]

With more than 1,100 members of the association having participated in the survey for Jourqual 3, an individual academic was not in the position to affect a journal's rank with his or her own voting behavior. Hence, an upgrade of a specific journal can safely be regarded as exogenous from an individual researcher's perspective, thus allowing us to causally identify shifts in affected academics' publication strategies following a ranking update. More specifically, we study the (relative) representation of 'affected' business researchers located in Austria, Germany, and Switzerland who have been socialized to acknowledge the VHB-journal-ranking and its updates to be career-relevant in 'treated' journals (i.e., journals that were upgraded) vs. 'untreated' journals (i.e., journals that retained their rank) pre- and post the ranking update with the help of a comprehensive self-collected journal-data set.

3. The Effects of a Journal Upgrade

Academics compete for limited publication space in academic journals, particularly in highly ranked ones. Highly ranked journals are, on average, more visible, thus enhancing academics' chances that their work will be well received and impact the work of others. Further, publishing in highly ranked journals is associated with more reputation and prestige (Roach & Sauermann, 2010; Sauermann & Stephan, 2013; Stern, 2004), and academics often even define themselves by the status of the journals in which they publish (Salandra et al., 2022). At times, academics might even receive direct monetary compensation for having published in a highly ranked

[3] For details on the construction of the ranking see https://vhbonline.org/vhb4you/vhb-jourqual/vhb-jourqual-3 [archived at http://web.archive.org/web/20240711025206/http://vhbonline.org/service/vhb-jourqual/vhb-jourqual-3].

journal (Osterloh & Frey, 2020), but, more importantly, having published in a highly ranked journal will positively affect an academic's career, e.g. in terms of receiving tenure (Bajo et al., 2020; Heckman & Moktan, 2020).

Because publishing in highly ranked journals will typically also be associated with higher costs in terms of time and effort, academics will weigh the benefits and costs of publishing in a journal when deciding on where to submit their work. If the benefits to publish in a journal increase, e.g., because of an upgrade of the respective journal without the costs increasing at the same time, academics are ceteris paribus more likely to choose the corresponding journal as their publication outlet. This leads us to our hypothesis:

H: Following the update of a journal ranking, affected academics (in comparison to unaffected ones) will publish relatively more often in upgraded journals than in journals with a constant ranking.

4. Data and Methodology

4.1 Data and Variables

We begin our data collection by downloading publicly available files from the webpage of the VHB[4], which display the ranks of journals that were covered in the VHB-journal-ranking preceding and following the ranking update in 2015. While the VHB journal ranking is focused on business journals, it also includes other journals which are deemed relevant for the profession, i.e. journals from neighboring disciplines (such as 'economics') as well as cross-disciplinary journals such as 'Science'. For our analysis, we included all journals listed in the VHB journal ranking – not only those that would classify as a business journal in a narrower sense. The VHB-journal-ranking classifies journals into six categories: A+, A, B, C, D, and not ranked. We restrict our data to journals that are ranked between A+ (outstanding, world-leading

[4] Please visit https://vhbonline.org/vhb4you/vhb-jourqual/vhb-jourqual-3/tabellen-zum-download [archived at http://web.archive.org/web/20240814145429/http://www.vhbonline.org/service/vhb-jourqual/vhb-jourqual-3/tabellen-zum-download] to retrieve the respective files.

scientific business administration journals) and C (recognized scientific business administration journals) according to both versions of the ranking. We excluded journals that were ranked D preceding and/or following the update or that were not ranked in either of the rankings. Journals that were ranked D are not strictly scientific journals but journals that address practitioners, for example "*Der Betriebswirt – Management in Wissenschaft und Praxis*". This results in 424 journals. Because only one single journal has been downgraded, we cannot analyze the effects of journal downgrades and remove this journal from our data set.[5] Further, we remove two journals with 'inconclusive' rankings according to the updated ranking, resulting in 421 journals.

We then scrape the online literature database Microsoft Academics[6] using an API client[7] to retrieve all papers published in these 421 journals between 2010 and 2021 where information regarding the affiliations of all authors is available. We did not start earlier with our data collection because the last larger overhaul prior to the update in 2015 took place in 2008 and we did not want to confound our results. Allowing for publication lags, this leaves us with six observation years before the treatment and six observation years thereafter. We are able to retrieve data for 400 journals, which constitute our final sample. For these journals, we retrieve data for 437,270 publications for which we have information regarding the journal in which the paper is published, the year of publication and the affiliations of all authors that contributed to the paper. Our final journal sample includes five journals that were upgraded from A to A+, 21 journals that were upgraded from B to A, and 63 journals that were upgraded from C to B. 311

[5] The fact that we observe 89 upgrades and only one downgrade in the range of ranks under consideration (A+ to C) may raise concerns about a potential grade inflation with the comparative value of, e.g., a publication in an A journal being subsequently reduced. While this may in fact be the case (in terms of academics subsequently needing more A-publications than before to get tenure), it still holds true that it becomes more attractive for an individual academic to publish in a journal after it has been upgraded as compared to publishing in that very same journal before the upgrade.

[6] While Microsoft Academics was retired at the end of 2021, the data retrieved from Microsoft Academics are available from the authors upon request.

[7] More specifically, we use the r package 'microdemic' as introduced by Chamberlain and Baker (2021).

journals retained their original ranking. Specifically, 17 (45) journals kept their A+ (A) status and 130 (119) journals were consistently rated as B (C) journals.[8]

Next, we assign all affiliations in our data to countries to identify affected authors. The authors in our data are affiliated with 12,376 unique affiliations. To assign these affiliations with the respective country, we scrape the Wikipedia webpages of the affiliations to assign the country of the affiliation. Random checks by a human rater hint at the reliability of this method. In those cases where this procedure does not render a result (2,447 of 12,276 affiliation), we assign the country by a human rater. The accuracy of the assignment was randomly checked by a second human rater. Via the countries of their affiliations, we identify 'affected' authors, i.e. academics who are subject to our treatment. Given our setting, we define affected authors as authors who are affiliated in Austria, Germany, or Switzerland.[9] We then aggregate these data points for each of the papers into journal-years where each journal-year represents all of the papers in our sample for a given journal in a given year. Based on these journal-years we create a panel data set where the level of observation are journal-years. Our panel consists of 4,584 journal-year observations[10].

For each journal-year, we calculate two variables that we use in our regression models as dependent variables. The first variable ($\%\ Papers\ with\ at\ least\ one\ Affected\ Author$) is

[8] While journals such as *Research Policy* or the *Strategic Management Journal* did retain their original (A) ranks, journals such as the *Journal of Management* or the *Journal of Management Studies* were upgraded from B to A. Detailed lists including all journals that received an upgrade or that retained their original ranking can be found following this link: https://vhbonline.org/vhb4you/vhb-jourqual/vhb-jourqual-3/tabellen-zum-download [archived at http://web.archive.org/web/20240814145429/http://www.vhbonline.org/service/vhb-jourqual/vhb-jourqual-3/tabellen-zum-download]

[9] In so doing, we also classify authors as being 'affected' who are not affected by the VHB ranking update – be it because they stem from a different discipline but publish in a journal that is listed in the VHB journal ranking or be it because they already know that they will not stay in the German-speaking university system, even though they currently do have an affiliation at a university in a German-speaking country. Similarly, authors that currently do not have an affiliation at a university in a German-speaking country, but plan to switch to a university in a German-speaking country in the future, are classified as 'unaffected' even though in fact they are affected by the VHB ranking update. As a result of the resulting measurement error, it will be harder for us to identify the hypothesized effect. More specifically, our point estimator will render the lower bound of the hypothesized effect.

[10] Our data covers 400 journals over a period of 12 years, which would result in 4,800 journal-year observations. 216 journal-years are missing in the data that we scrape from Microsoft Academics – one reason being that the *VHB-journal-ranking* also covers German language journals, such as, e.g., *Kredit und Kapital (Credit and Capital Markets)*, which are not that well covered by Microsoft Academics.

the fraction of papers with at least one affected author in a given journal and year. To construct this variable, we count the number of papers in any journal-year where at least one of the authors is identified as 'affected' and divide this number by the total number of papers in the same journal-year. E.g., if a journal publishes 200 papers in a specific year and if on 40 of these we detect at least one author from Austria, Germany, or Switzerland, the variable $\%\ Papers\ with\ at\ least\ one\ Affected\ Author$ would take the value of 20% for this journal-year. As can be inferred from Table VII-1, the average over all journal-years of $\%\ Papers\ with\ at\ least\ one\ Affected\ Author$ in our data equals 11.70%.

Table VII-1: Descriptive Statistics

Journal Ranking (before and after the ranking update)	# Journals	# Journal-years	Mean % Papers with at least one Affected Author	Mean % Affected Authors
A+ → A+	17	203	0.0713	0.0440
A → A+	5	60	0.0444	0.0274
A → A	45	535	0.1571	0.1122
B → A	21	250	0.1088	0.0779
B → B	130	1514	0.1357	0.1068
C → B	63	742	0.1224	0.0925
C → C	119	1280	0.0872	0.0715
Total	400	4,584	0.1170	0.0898

Notes: This table reports the descriptive statistics for our sample. The first column $Journal\ Ranking$ indicates the ranking a journal had before and after the ranking update. $\#\ Journal$ indicates how many journals each category contains. $\#\ Journal - years$ indicates how many total journal-years in the category are included in the sample. $Mean\ \%\ Papers\ with\ at\ least\ one\ Affected\ Author$ indicates the average fraction of papers with at least one affected author for all journal-years in this category. $Mean\ \%\ Affected\ Authors$ indicates the average fraction of affected authors for all journal-years in this category. Please note that downgrades were largely absent and therefore do not appear in the table.

Because $\%\ Papers\ with\ at\ least\ one\ Affected\ Author$ could potentially be affected by coauthor team sizes increasing over time (M. Fernandes & Walter, 2023; Jones, 2021; Kwiek, 2021), we employ a second variable: $\%\ Affected\ Authors$. This variable is the fraction of affected authors in any journal-year, thus accounting for the size of coauthor teams increasing over time. E.g., if a journal publishes 200 papers in a specific year which are written by a total of 500 distinct authors and if 50 of these authors are located in Austria, Germany, or Switzerland, $\%\ Affected\ Authors$ takes the value of 10% for this journal-year. The average over all journal-years of $\%\ Affected\ Authors$ in our data is 8.98%.

4.2 Methodology

To test whether the representation of affected authors in upgraded ('treated') journals increases more strongly following the ranking update compared to their representation in journals with a constant ranking ('untreated'), we estimate a difference-in-differences model of the following form:

$$y_{i,t} = \beta_1 * Post\ Update_t + \beta_2 * Upgraded\ Journal_i + \beta_3 * Post\ Update_t\ x\ Upgraded\ Journal_i + \theta_i \quad \text{(VII-1)}$$

In this model, $y_{i,t}$ is one of the two dependent variables, i.e., $\%\ of\ Papers\ with\ at\ least\ one\ Affected\ Author$ or $\%\ Affected\ Authors$ in journal i in year t. The dummy $Post\ Update_t$ equals one for the year 2016 and later years. While the ranking update was released at the beginning of 2015, we anticipate that its effects would not become apparent in the very same year because it takes some until a paper is published. In total, this yields a publication lag of one to two years to account for the time-consuming nature of publication processes, which is in line with literature (see, e.g., Bäker (2015)). $Upgraded\ Journal_i$ is a dummy variable that equals one if a journal receives an upgrade in the update of the VHB-journal-ranking. We are particularly interested in the coefficient of the difference-in-differences estimator $Post\ Update_t\ x\ Upgraded\ Journal_i$, which indicates whether affected authors did increase their representation in upgraded journals more strongly compared to journals with a constant ranking following the ranking update. In addition, we use journal fixed effects θ_i to control for any time-invariant journal characteristics, which absorb the coefficient of the dummy variable $Upgraded\ Journal_i$. We estimate our models using standard errors clustered at the journal-level.

4.3 Common Trends

A crucial assumption for any difference-in-differences model is that the variables of interest follow parallel trends for the treatment and the control group prior to the treatment. In order to provide evidence that the parallel trends assumption holds in our setting, we follow recent literature such as Braghieri et al. (2022), M. Fernandes et al. (2023), He and Wang (2017), or Miller et al. (2021) and conduct an event study. More precisely, we estimate the following fixed effects model:

$$y_{it} = \sum_{k \geq -5, k \neq -1}^{k=6} D_{it}^{k}\, \delta_k + \theta_i + \varepsilon_{it} \qquad \text{(VII-2)}$$

where y_{it} represents the dependent variables, i.e., $\%\,Papers\ with\ at\ least\ one\ Affected\ Author$ or $\%\,Affected\ Authors$. D_{it}^{k} is a set of binary variables that take the value one if the journal-year i was k years away from the update of the VHB-journal-ranking. We omit the dummy for $k = -1$ so that the post-treatment effects are relative to the period immediately prior to the year in which the VHB-journal-ranking was updated. θ_i indicates the journal-fixed effects that we include in the regression model.

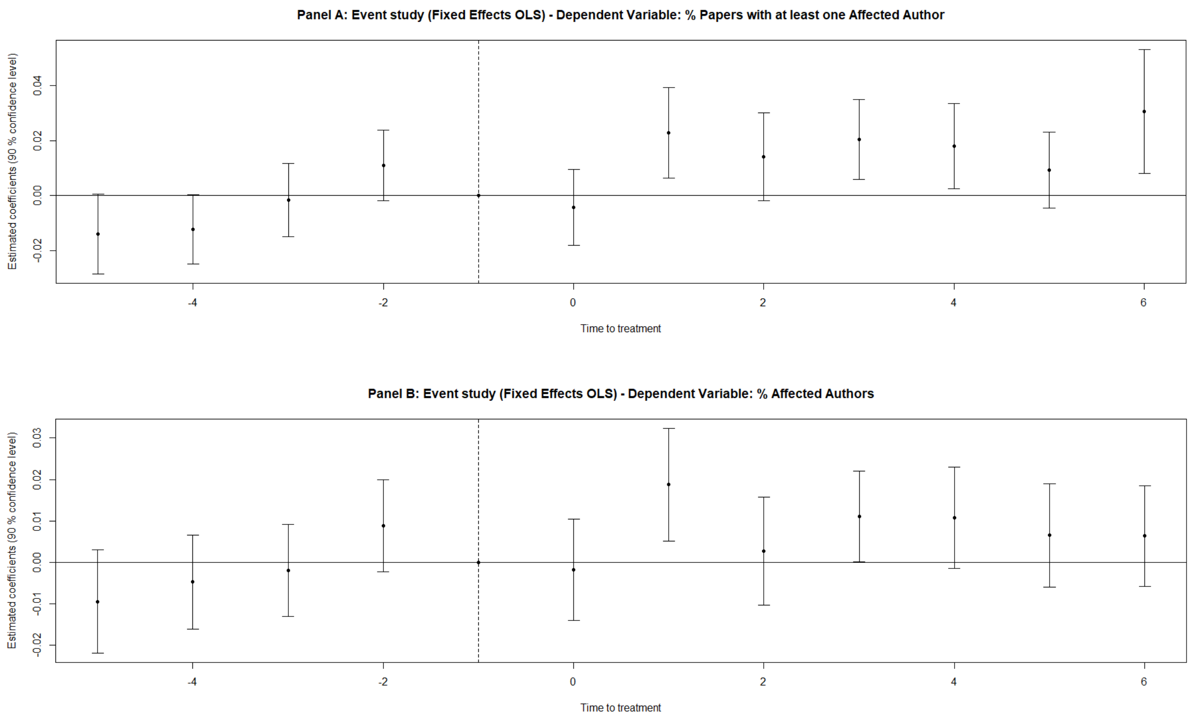

Figure VII-1: Estimated Effect of the Update of the VHB-Journal-Ranking

Notes: This figure displays the event study plots for both dependent variables in study 1 based on a fixed effects model, where we include journal fixed effects. Panel A provides the results for our first dependent variable, i.e. $\%\,Papers\ with\ at\ least\ one\ Affected\ Author$ as defined in Section 4.1.1. Panel B provides the results regarding our second dependent variable, i.e. $\%\,Affected\ Authors$ as defined in Section 4.1.1. The time variable is based on the distance (in years) between the update of the VHB-journal-ranking and the respective journal-year. Bars represent 90 % intervals.

Panel A of Figure 1 shows that the estimates are in line with the parallel-trends assumption regarding our first dependent variable $\%\, Papers\ with\ at\ least\ one\ Affected\ Author$. Before the update of the VHB-journal-ranking, all coefficients are close to zero, showing no noticeable trends before treatment. Hence, we conclude that journals that did not receive an upgrade serve as a suitable control group for journals that did receive an upgrade. Panel B of Figure 1 displays similar results regarding our second dependent variable $\%\, Affected\ Authors$. Again, we observe that the estimates are in line with the parallel-trends assumption as they show no discernible differences before treatment. This finding confirms that journals that did not receive an upgrade are a suitable control group for journals that did receive an upgrade.[11]

5. Results

Table VII-2 displays regression results regarding affected academics' publication behavior caused by the update of the VHB-journal-ranking. Column (1) reports our results regarding the first dependent variable, $\%\, of\ Papers\ with\ at\ least\ one\ Affected\ Author$. The coefficient of the interaction term is statistically significant and positive. This indicates that the fraction of papers with at least one affected author increases more strongly in journals that receive an upgrade compared to journals whose rank remains constant. The magnitude of the coefficient equals 1.29%, which is economically meaningful, given that the average fraction of papers with at least one affected author equals 11.70%.

[11] Please find the exact coefficient estimates from our fixed-effects regressions in Appendix VII-A.

Table VII-2: Panel Regressions – Main Effect

All (upgraded) journals	(1)	(2)
	% Papers with at least one Affected Author	% Affected Authors
Mean LHS	0.1170	0.0898
Post Update	0.0097^{**}	0.0016
	(0.0048)	(0.0042)
Post Update x Upgraded Journal	0.0129^{*}	0.0092^{*}
	(0.0069)	(0.0055)
Journal Fixed Effects	Yes	Yes
Observations (journal-years)	4,584	4,584
R^2	0.0056	0.0013
F Statistic	11.8610^{***}	2.6966^{*}

Notes: This table reports the results of our main difference-in-differences estimation. $Post\ Update_t$ is a dummy variable that equals one, for the years 2016 and later. The difference-in-differences estimator is $Post\ Update_t x\ Upgraded\ Journal_i$, which indicates whether affected authors did increase their representation in upgraded journals in consequence of the ranking update as compared to those journals with a constant ranking. In addition, we use journal fixed effects to control for any time-invariant journal characteristics which absorb the coefficient of the dummy variable $Upgraded\ Journal_i$. All models are estimated using standard errors clustered at the journal-level. Significance levels are denoted as follows: *p<0.1; **p<0.05; ***p<0.01.

Column (2) shows our results regarding the second dependent variable, % $Affected\ Authors$. Again, we find that the coefficient of the interaction term is statistically significant and positive. This shows that the fraction of affected authors increases more strongly in upgraded journals following the ranking update as compared to journals with a constant rank. The size of the coefficient (0.92%) is again economically relevant, as the average fraction of affected authors in our sample equals 8.98%.

Taken together, our results from Table VII-2 support our hypothesis that affected academics react to the update of a journal ranking by publishing relatively more frequently in upgraded journals (compared to journals with constant rankings) after the new version of the ranking has been released.

The fact that we also find a significantly positive coefficient for the $Post\ Update$ dummy when analyzing % $of\ Papers\ with\ at\ least\ one\ Affected\ Author$ indicates that the fraction of papers with at least one affected author also increases in journals with a constant ranking. This finding is in line with previous literature highlighting that business researchers in the German-speaking countries in general increasingly focus on journal publications in the respective time frame (as opposed to the publication of monographs or chapters in edited books, see, e.g., Ayaita et al. (2019)) – thus catching up with business researchers from other countries who had focused on journal publications before.

Next, we explore potential heterogeneities with respect to different journal ranks, i.e., we analyze if it makes a difference whether a journal is upgraded from B to A or from C to B.[12] Panel A of Table VII-3 displays our results when comparing journals that receive an upgrade from B to A with journals that retain their original A and/or B rank. We find a significant increase in both our dependent variables for journals that are upgraded from B to A when we compare these journals to journals that retain their original A or B rank (columns (1) and (4)). E.g., a coefficient of 2.58% for the interaction term displayed in column (1) of Panel A is economically meaningful, given that the average of the dependent variable equals 13.78%. We find similar results when comparing journals that were upgraded from B to A to journals that retained their original B rank (columns (3) and (6)). When we compare journals that were upgraded from B to A to journals that were already A journals before the update and that retained their A rank (columns (2) and (5)), we find no significant increase following the ranking update. This means that academics do not strategically shift their publications from journals that were already ranked A before to the newly added A journals, but they rather shift away from B journals who retained their ranking.

[12] Because only five journals were upgraded from A to A+ and because only very few business researchers in German speaking-countries published in an A+ journal at the time of our data, we refrain from interpreting our results for this subsample as it is too small to derive meaningful results. For the sake of completeness, the results can, however, be found in Appendix VII-D.

Table VII-3: Panel Regressions – Ranking Heterogeneity

Panel A: VHB-journal-rank (B) to VHB-journal-rank (A)

	% Papers with at least one Affected Author			% Affected Authors		
Mean LHS	0.1378	0.1417	0.1319	0.1049	0.1012	0.1027
Comparison group: Journals that retained their original … rank	(1) (A)/(B)	(2) (A)	(3) (B)	(4) (A)/(B)	(5) (A)	(6) (B)
Post Update	0.0177***	0.0408***	0.0095	0.0072	0.0230***	0.0016
	(0.0067)	(0.0098)	(0.0082)	(0.0057)	(0.0075)	(0.0072)
Post Update x Upgraded Journal	0.0258**	0.0027	0.0340***	0.0165*	0.0006	0.0221**
	(0.0116)	(0.0136)	(0.0126)	(0.0094)	(0.0106)	(0.0104)
Journal Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations (journal-years)	2,299	785	1,764	2,299	785	1,764
R^2	0.0193	0.0726	0.0093	0.0043	0.0380	0.0029
F Statistic	14.8179***	28.0592***	7.5408***	4.5362**	14.1774***	2.3806*

Panel B: VHB-journal-rank (C) to VHB-journal-rank (B)

	% Papers with at least one Affected Author			% Affected Authors		
Mean LHS	0.1154	0.1314	0.1001	0.0910	0.1021	0.0792
Comparison group: Journals that retained their original … rank	(1) (B)/(C)	(2) (B)	(3) (C)	(4) (B)/(C)	(5) (B)	(6) (C)
Post Update	0.0028	0.0095	-0.0054	-0.0030	0.0016	-0.0086
	(0.0038)	(0.0082)	(0.0075)	(0.0051)	(0.0072)	(0.0071)
Post Update x Upgraded Journal	0.0125	0.0058	0.0208**	0.0092	0.0046	0.0148*
	(0.0083)	(0.0100)	(0.0094)	(0.0065)	(0.0083)	(0.0082)
Journal Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations (journal-years)	3,536	2,256	2,022	3,536	2,256	2,022
R^2	0.0015	0.0041	0.0031	0.0006	0.0006	0.0026
F Statistic	2.4226*	4.2637**	2.8342*	0.9316	0.6076	2.4080*

Notes: This table reports the results of our difference-in-differences estimations regarding different levels of journal rankings. Panel A compares journals that receive an upgrade from B to A with journals that retain their original B and/or A ranking. Panel B compares journals that receive an upgrade from C to B with journals that retain their original C and/or B ranking. $Post\ Update_t$ is a dummy variable that equals one, for the years 2016 and later. The difference-in-differences estimator is $Post\ Update_t x\ Upgraded\ Journal_i$, which indicates whether affected authors did increase their representation in upgraded journals in consequence of the ranking update as compared to the respective control group. In addition, we use journal fixed effects to control for any time-invariant journal characteristics which absorb the coefficient of the dummy variable $Upgraded\ Journal_i$. All models are estimated using standard errors clustered at the journal-level. Significance levels are denoted as follows: *p<0.1; **p<0.05; ***p<0.01.

Panel B of Table VII-3 presents the results of our analysis when comparing journals that receive an upgrade from C to B with journals that retain their original B and/or their original C ranking. We find weaker effects for journals upgraded to the B rank. In particular, we find no statistically significant interaction terms in models (1), (2), (4), and (5). This indicates that both the share of papers with at least one affected author and the share of affected authors do not increase more strongly in journals upgraded from C to B, compared to journals that retained their B or C (columns (1) and (4)), and compared to journals that retained their B (columns (2) and (5)). Only if we compare journals who were upgraded from C to B to journals that retained their original C rank, we do find weak statistically significant interaction terms in both models (columns (3) and (6)).

In summary, our results in Table VII-3 suggest that affected academics react strongly to journals that receive an upgrade from B to A and not to those that receive an upgrade from C to B – which makes sense because tenure and appointment decisions at the time of our data in the German-speaking countries were typically based on having published in A journals.[13]

6. Robustness Tests and Additional Analyses

6.1 Robustness Tests

To test the robustness of our findings, we first conduct a placebo test where we focus on business researchers affiliated to U.S. institutions instead of business researchers affiliated to institutions in German-speaking countries. Arguably, business researchers from the U.S. should not be affected by the update of the VHB-journal-ranking (with the exception of those few academics that might consider switching to an institution in a German-speaking country in the future).

[13] While, internationally, even at the time of our data, it was important for business researchers to also have published in A+ journals (and not 'only' in A journals), tenure and appointment decision in German-speaking countries at the time of our data typically depended on business researchers having published in A journals. In fact, in 2018, 75% of professors in business studies who were affiliated at German universities did not have one single publication in an A+ journal (M. Fernandes & Walter, 2022).

Table VII-4: Panel Regressions – Main Effect – Robustness Check (U.S. Based Authors)

All (upgraded) journals

	(1) % Papers with at least one U.S. Based Author	(2) % U.S. Based Authors
Mean LHS	0.4101	0.3262
Post Update	-0.0281***	-0.0435***
	(0.0060)	(0.0052)
Post Update x Upgraded Journal	-0.0007	-0.0063
	(0.0109)	(0.0088)
Journal Fixed Effects	Yes	Yes
Observations (journal-years)	4,584	4,584
R^2	0.0130	0.0407
F Statistic	27.5447***	88.6616***

Notes: This table reports the results of two difference-in-differences estimations where our dependent variables are based on U.S based authors instead of authors affected by the journal ranking update. $Post\ Update_t$ is a dummy variable that equals one, for the years 2016 and later. The difference-in-differences estimator is $Post\ Update_t x\ Upgraded\ Journal_i$, which indicates whether U.S. based authors did increase their representation in upgraded journals in consequence of ranking update. In addition, we use journal fixed effects to control for any time-invariant journal characteristics which absorb the coefficient of the dummy variable $Upgraded\ Journal_i$. All models are estimated using standard errors clustered at the journal-level. Significance levels are denoted as follows: *p<0.1; **p<0.05; ***p<0.01.

Tables 4 and 5 present the respective results. As expected, we find no reaction of U.S. affiliated authors to the ranking update in question – neither across all ranking levels nor for specific ranking levels. In particular, the interaction terms of our difference-in-differences models are statistically insignificant in all our estimations. This highlights that, as expected, U.S. affiliated authors are not affected and hence do not react to the update of the VHB-journal-ranking, thus supporting our claim that the observed changes in the publication behavior of academics from German-speaking countries can be causally attributed to the update of the VHB-journal-ranking.

Table VII-5: Panel Regressions – Ranking Heterogeneity – Robustness Checks (U.S. Based Authors)

Panel A: VHB-journal-rank (B) to VHB-journal-rank (A)

	% Papers with at least one U.S. Based Author			% U.S. Based Authors		
Mean LHS	0.4441	0.5409	0.4218	0.3493	0.4183	0.3353
Comparison group: Journals that retained their original … rank	(1) (A)/(B)	(2) (A)	(3) (B)	(4) (A)/(B)	(5) (A)	(6) (B)
Post Update	-0.0226***	-0.0146	-0.0255***	-0.0400***	-0.0369***	-0.0411***
	(0.0071)	(0.0117)	(0.0086)	(0.0063)	(0.0102)	(0.0077)
Post Update x Upgraded Journal	-0.0214	-0.0295	-0.0186	-0.0239	-0.0271	-0.0228
	(0.0174)	(0.0197)	(0.0180)	(0.0155)	(0.0175)	(0.0162)
Journal Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations (journal-years)	2,299	785	1,764	2,299	785	1,764
R^2	0.0126	0.0160	0.0145	0.0421	0.0549	0.0424
F Statistic	13.4253***	5.8237***	11.8415***	46.2064***	20.8398***	35.6256***

Panel B: VHB-journal-rank (C) to VHB-journal-rank (B)

	% Papers with at least one U.S. Based Author			% U.S. Based Authors		
Mean LHS	0.3523	0.3770	0.3212	0.2816	0.2985	0.2577
Comparison group: Journals that retained their original … rank	(1) (B)/(C)	(2) (B)	(3) (C)	(4) (B)/(C)	(5) (B)	(6) (C)
Post Update	-0.0303***	-0.0255***	-0.0362***	-0.0439***	-0.0411***	-0.0472***
	(0.0070)	(0.0086)	(0.0115)	(0.0061)	(0.0077)	(0.0097)
Post Update x Upgraded Journal	0.0059	0.0011	0.0118	0.0015	-0.0012	0.0049
	(0.0135)	(0.0144)	(0.0163)	(0.0105)	(0.0115)	(0.0129)
Journal Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations (journal-years)	3,536	2,256	2,022	3,536	2,256	2,022
R^2	0.0130	0.0123	0.0144	0.0363	0.0416	0.0365
F Statistic	21.2821***	12.8773***	13.4055***	60.7154***	44.7503***	34.7937***

Notes: This table reports the results of our difference-in-differences estimations regarding different levels of journal rankings where our dependent variables are based on U.S based authors instead of authors affected by the journal ranking update. Panel A compares journals that receive an upgrade from B to A with journals that retain their original B and/or A ranking. Panel B compares journals that receive an upgrade from C to B with journals that retain their original C and/or B ranking. $Post\ Update_t$ is a dummy variable that equals one, for the years 2016 and later. The difference-in-differences estimator is $Post\ Update_t x\ Upgraded\ Journal_i$, which indicates whether U.S: based authors did increase their representation in upgraded journals in consequence of the ranking update as compared to the respective control group. In addition, we use journal fixed effects to control for any time-invariant journal characteristics which absorb the coefficient of the dummy variable $Upgraded\ Journal_i$. All models are estimated using standard errors clustered at the journal-level. Significance levels are denoted as follows: *p<0.1; **p<0.05; ***p<0.01.

To further corroborate the robustness of our results, we perform a second placebo test, where we replace U.S authors with authors affiliated to institutions in France (Appendix VII-B and C). Again, and as expected, we observe no statistically significant interaction effect in our estimations. Arguably, for authors affiliated to institutions in France, the VHB-journal-ranking is not relevant and hence these authors do not react to the update – again supporting our claim that the observed changes in the publication behavior of academics affiliated to institutions in German-speaking countries can be causally attributed to the update of the VHB-journal-ranking.

To make sure that our estimated effects are not confounded by the upgrade of 'The Academic Journal Guide' conducted by the Chartered Association of Business School, CABS, that was also updated in 2015[14], we perform additional analyses where we create a dummy variable ($AJG\ Upgraded\ Journal$) that equals one for journals that received an upgrade in the Academic Journal Guide in 2015 and replicate our initial analysis from Table VII-2.

Table VII-6: Panel Regressions – Main Effect – Robustness Check (Journals Upgraded in the Academic Journal Guide)

All (upgraded) journals

	(1) % Papers with at least one Affected Author	(2) % Affected Authors
Post Update	0.0143***	0.0055
	(0.0050)	(0.0044)
Post Update x AJG Upgraded Journal	-0.0047	-0.0049
	(0.0078)	(0.0066)
Journal Fixed Effects	Yes	Yes
Observations (journal-years)	4,584	4,584
R^2	0.0049	0.0009
F Statistic	10.3094***	1.8018

Notes: This table reports the results of two difference-in-differences estimation. $Post\ Update_t$ is a dummy variable that equals one, for the years 2016 and later. The difference-in-differences estimator is $Post\ Update_t\ x\ AJG\ \ Upgraded\ Journal_i$, which indicates whether affected authors did increase their representation in journals that received an upgrade in the AJG in consequence of the ranking update as compared to those journals with a constant ranking. In addition, we use journal fixed effects to control for any time-invariant journal characteristics which absorb the coefficient of the dummy variable $AJG\ Upgraded\ Journal_i$. All models are estimated using standard errors clustered at the journal-level. Significance levels are denoted as follows: *p<0.1; **p<0.05; ***p<0.01.

Table VII-6 presents the respective results. Since the interaction term between $Post\ Update$ and $AJG\ Upgraded\ Journal$ is statistically insignificant, we conclude that researchers from German-speaking countries do not react to changes in the Academic Journal Guide. Thus, our

[14] Other than the VHB-journal-ranking, this ranking does not only rely on expert opinions when assigning values to journals, but also employs citations. For details on its construction see https://charteredabs.org/academic-journal-guide-2021-available-now/. Journals are classified in the categories 1 (bottom) to 4* (top). For a descriptive analysis regarding the overlap of ranking changes between the VHB-journal-ranking and the AJG please refer to Appendix VII-E.

results are unlikely to be biased by the upgrades that took place in the Academic Journal Guide as at the same time as the updates in the VHB-journal-ranking.[15]

6.2 Were the Upgrades Warranted?

Even though our predictions do not depend on whether the upgrades in the VHB-journal-ranking were warranted or not (either may be true and will lead to the very same predictions), our data offers the unique opportunity to also shed light on whether the upgrades in the update of the VHB-journal-ranking were warranted or not by tracing the development of the journals' citation-based SCImago Journal Ranking (SJR) (González-Pereira et al., 2010) as a more 'objective' measure of journal quality (Śpiewanowski and Talavera (2021); Hudson (2024)).

A graphical inspection of the SJR values of journals that were upgraded compared to journals that retained their original ranking in Figure 2, before and after the ranking update, shows that the journal upgrades were, on average, warranted and hence not driven by strategic voting behavior of academics. Rather, the journal upgrades seem to have corrected for a former misalignment between their original VHB-journal-ranking and the costs of publishing in the respective journal, where the latter may be proxied by a journal's SJR value as an established outside criterion of journal quality.

[15] In a different set of robustness tests, we build different sub-samples where we replicate our analyses with samples that either exclude journals that receive an upgrade in the VHB-journal-ranking or the AJG. These robustness tests further strengthen our findings that our results are not biased by upgrades in the AJG. The respective results are available upon request.

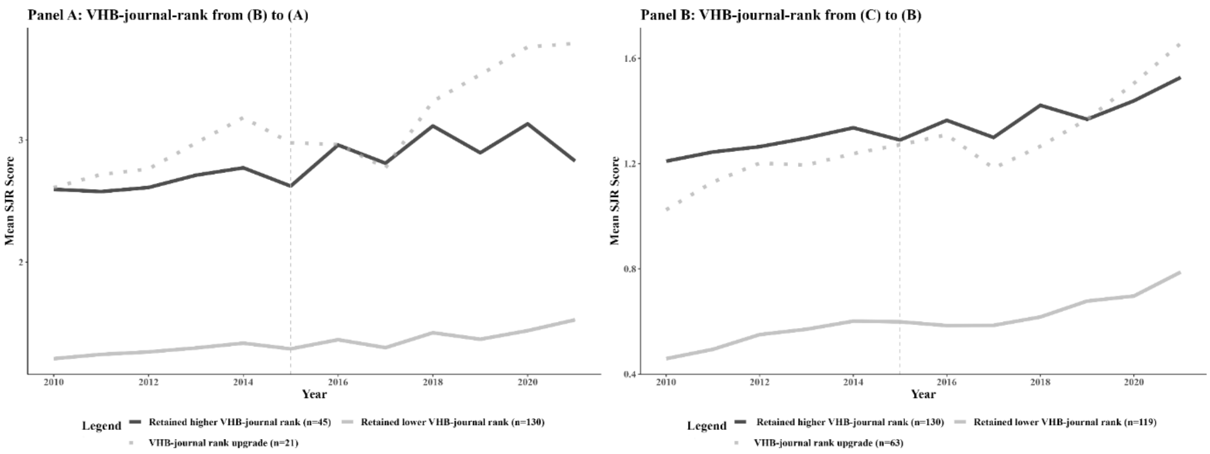

Figure VII-2: SJR Score of all upgrades

Notes: This figure displays the average yearly SJR Scores for journals that either received an upgrade or retained their original ranking. Panel A displays the average scores for journals that received an upgrade from B to A (dotted grey line) as well as journals that retained their original B ranking (continuous grey line) or their original A ranking (continuous black line). Panel B displays the average scores for journals that received an upgrade from C to B (dotted grey line) as well as journals that retained their original C ranking (continuous grey line) or their original B ranking (continuous black line).

6.3 Further Analysis – Researcher Heterogeneity

To explore whether certain affected researchers – e.g. based on their age and publication record – react differently to the update of the VHB-journal-ranking, we use a second hand-collected dataset with individual researchers (and not journal-years) as the level of analysis. We start by identifying all business researchers who hold a tenured professorship at a university in a German-speaking country by the end of 2018 (1,536 individual researchers). Next, we collect data (year and institution of each career step, year of birth, gender etc.) from publicly available CVs of these researchers. Omitting researchers with missing information, we gain a final sample of 1,124 researchers. To these, we merge publication data from the online research-monitoring portal Forschungsmonitoring (cf. Ayaita et al. (2019), Bäker et al. (2021)). The publication data cover all journal publications of the researchers in our data set and contain information regarding year, journal and number of coauthors of each publication. To capture the researchers' reaction to the journal upgrades we calculate two variables: $\# \: Pubs \: Upgraded$ is the number of publications in upgraded journals published by a researcher in our data set between 2016 and 2018. Researchers in our data publish an average of 0.67 publications in upgraded journals between 2016 and 2018. $Higher \: \# \: Pubs \: Upgraded$ is a dummy variable that equals one if a researcher has more publications in upgraded journals between 2016 and

2018 (i.e. following the update) compared to the period of 2013 and 2015 (i.e. preceding the update). Roughly 23% of the researchers in our sample publish more in upgraded journals in the first three years after the ranking update as compared to the three years before.

Based on our merged data set, we define a series of independent variables that we use to explore potential heterogeneities in terms of researchers' characteristics. First, we create a set of dummy variables to capture the age of the researchers at the time of the ranking update in 2015. The first dummy equals one if a researcher is younger than 40 (22% of the researchers), the second equals one, if a researcher is between 40 and 49 years (41% of the researchers), and the third equals one, if a researcher is older than 50 years (37% of the researchers). Next, we include one variable $\#\,Publications\,pre\,Update$ measured as the number of publications a researcher in our sample has published until 2015 to proxy for researchers' publication output prior to the ranking update.

We control for researchers' field of specialization within business research, and we include a dummy variable that equals one if a researcher is a woman ($Female$) to account for potential gender differences (e.g., Jokinen and Pehkonen (2017)). We also include a dummy variable to control for whether a researcher obtained his or her PhD in a German-speaking country ($PhD\ German - speaking\ country$), as the update of the-VHB-journal-ranking is arguably more relevant for researchers who have been educated and socialized within the German-speaking university system. Finally, we include a dummy variable ($Elite\ University$), which – based on Clermont (2016) – controls for particularly renowned universities/business schools at which the researchers are affiliated. Table VII-7 presents our results.

Table VII-7: OLS Regressions – Researcher Heterogeneity

	(1) OLS	(2) LPM
	# Pubs Upgraded	Higher # Pubs Upgraded
Mean LHS	0.6681	0.2340
Age < 40	0.5234***	0.1985***
	(0.1096)	(0.0353)
40 ≥ Age < 50	0.1364*	0.0746***
	(0.0808)	(0.0257)
# Publications pre Update	0.0442***	0.0066***
	(0.0062)	(0.0010)
Female	-0.0596	0.0192
	(0.0848)	(0.0327)
Elite University	0.1344	0.0345
	(0.1215)	(0.0372)
PhD German-speaking country	-0.3130	-0.0163
	(0.1982)	(0.0450)
Constant	0.3284	-0.0051
	(0.2308)	(0.0516)
Field of Specialization	Yes	Yes
Observations	1,124	1,124
R^2	0.2630	0.1074
F Statistic	33.0422***	11.1434***

Notes: This table reports the results of two different OLS/ LPM regressions. $\#\ Pubs\ Upgraded$ is the number of publications in upgraded journals published by a researcher between 2016 and 2018. $Higher\ \#\ Pubs\ Upgraded$ is a dummy variable that equals one if a researcher has more publications in upgraded journals between 2016 and 2018 (i.e. following the update) compared to the period of 2013 and 2015 (i.e. preceding the update). $\#\ Publications\ pre\ Update$ measures the number of publications a researcher in our sample has published until 2015. $Female$ is a dummy variable which equals 1 if the researcher is female. $Elite\ University$ is a dummy variable which equals 1 if the researcher is currently affiliated with a particularly renowned university/business school based on Clermont (2016). $PhD\ German-speaking\ country$ is a dummy variable which equals one if the researcher obtained their PhD in a German-speaking country. $Field\ of\ Specialization$ refers to a set of field-specific dummy variables indicating the field of specialization of the researcher. Both models are estimated using heteroscedasticity-robust standard errors. Significance levels are denoted as follows: *p<0.1; **p<0.05; ***p<0.01.

Overall, our analyses show that researchers do not react homogeneously to upgrades in a journal ranking. While there are no differences between genders, younger and more active researchers react more strongly to the ranking update compared to their peers. For the younger researchers whom we find to more strongly react to the update, the benefits of publishing in highly ranked journals are arguably larger. Hence, we would expect them to react more strongly to the ranking update and be more eager to change their publication strategy.[16] For more actively publishing researchers the (individual) costs to switch between production outlets will most likely be

[16] While it may be questioned whether younger researchers are able to affect the decision on where to submit a paper that might be co-authored with more senior colleagues (who might not be affiliated to a German-speaking university), our results hint at that being actually the case.

lower. As a result, these researchers, too, will be more likely to change their publication strategy in response to a ranking update.

7. Discussion and Conclusion

7.1 Summary of Results

Based on our unique data sets, our paper provides novel evidence regarding how academics respond to changing incentives to publish in certain journals as opposed to others. More specifically, we analyze how the update of a journal ranking causally affects the publication behavior of academics for whom the ranking is relevant. As predicted, we find affected academics to increasingly publish in journals that are upgraded – both, compared to journals that retain their original rank and compared to unaffected academics for whom the ranking and its update are not relevant. Seemingly, affected researchers have been successfully socialized to accept the authority of the journal ranking such that they respond to its update in the predicted manner. This strategic shift in publication behavior is driven by journals that were upgraded from B to A, and not so much by journals that were upgraded from C to B. Younger and more actively publishing researchers were found to more strongly react to the update.

7.2 Discussion of Findings in Relation to Literature

Our findings complement the studies of Śpiewanowski and Talavera (2021) and Hudson (2024) on the effects of a journal rating change on U.K. based researchers' choice of publication outlets. However, both studies do not causally identify the effect. Śpiewanowski and Talavera (2021) employ a different empirical approach, examining papers uploaded to the IDEAS/RePEc online repository from 2010 to 2014, and tracking their outcomes. Śpiewanowski and Talavera (2021) found that the share of papers published in journals upgraded to AJG 4 increased significantly. However, they did not provide the corresponding unconditional probability, preventing a direct comparison of the effect sizes of an upgrade to AJG 4 with the effects documented in our study. Interestingly, Śpiewanowski and Talavera (2021) did not observe a significant effect for upgrades to the most prestigious category, AJG 4*. This result aligns with our findings regarding upgrades to A+ status, as we also report an insignificant effect for these upgrades (see Appendix VII-D). Going beyond the analyses conducted by Śpiewanowski and Talavera (2021) and explicitly addressing their call for corresponding analyses, our study further explores the heterogeneities in researchers' responses to journal upgrades. We find that younger and more active researchers react more strongly to the changes in journal ranks. Hudson (2024) employs a similar approach to ours, utilizing data of publications from journals listed in the AJG to analyze whether U.K. business researchers responded to the 2015 update of the AJG ratings in the 2011–2021 period. However, Hudson (2024) does not causally identify the effect by employing a difference-in-difference estimation.

In addition, our study relates to research that analyzes the effects of journal upgrades on other outcomes, e.g. researchers' choice of topic or the likelihood of their work being cited. While Buehling (2021) shows that researchers do not react to changes in journal ranks by their choice of topic, Drivas and Kremmydas (2020) find that papers published in journals that receive an upgrade experience a significant increase in citations. The increase in citations is attributed primarily to an intentional 'signaling effect', where researchers cite these highly ranked journals to associate their work with prestigious outlets. Being more likely to be cited – even if only by the community of affected researchers – will create yet another incentive for researchers based in a German-speaking country to shift their publication strategy towards upgraded journals.

Regarding researchers' perceptions of journal rankings, Walker, Salter, et al. (2019) show that those researchers who published in journals upgraded in the AJG were less hostile and more positive toward the ranking compared to those who did not benefit from the re-grading. However, upward shifts in journal rankings were only significant for upgrades from AJG 3- to AJG 4-star, with researchers becoming less hostile and more positive toward the ranking. In contrast, upgrades from AJG 2- to AJG 3-star did not influence researchers' attitudes, likely due to the greater emphasis placed on AJG 4-star journals by university research managers. Our analysis aligns with the findings of Walker, Salter, et al. (2019), as we observe that affected researchers in our study do not react strongly to upgrades in the lower tiers of the journal ranking.

7.3 Limitations and Avenues for Future Research

Our paper is, of course, not without limitations. First, our empirical analysis refers to a specific setting: the update of a survey-based journal ranking which is relevant for a specific group of researchers (business researchers who are affiliated to universities in German-speaking countries) and not for others. While this setting is ideal to causally identify the hypothesized effect, it might limit the generalizability of our results and implications. Future research might address this but analyzing the hypothesized effect in other settings. Other settings might also allow for an assessment of lead authorship and its effects – something which is hard to study in the field of business research where the ordering of coauthors is alphabetical in most instances (J. M. Fernandes & Cortez, 2020).

Further, because we obtain our data from a specific platform (Microsoft Academic), we rely on the coverage of this very platform. While other platforms, such as, e.g., Web of Science, might provide a higher coverage, recent literature (e.g., Harzing and Alakangas (2017), Martín-Martín et al. (2021) shows that Microsoft Academic compares reasonably well to other sources of publication data. Harzing (2019, p. 341) even concludes that '…Google Scholar and Microsoft Academic maintain their position as the most comprehensive free sources for publication and citation data'. Thus, our choice of data source should not pose a problem.

When analyzing shifts in academics' publication strategies, we treat the update of the journal ranking as exogenous from the perspective of individual academics. While we are confident that our results can be trusted in picking up a causal effect, in an ideal setting, the ranking update would, of course, have to be completely exogenous. To be able to measure the effects of a change in the incentives to publish in treated versus untreated journals, this exogenous update would however have to take place in a way that the benefits and costs of publishing in treated versus untreated journals are not affected at the same time. As this may be difficult to achieve in a real-world setting, future research might seek to address this point experimentally.

Further, future research might analyze whether groups of researchers with aligned interests successfully coordinated to strategically upvote journals in which they regularly publish or expect to publish in. While we found the journal upgrades on average to be warranted, this does not mean that there was no strategic upvoting involved. Detecting explicit (or implicit) coordination amongst groups of researchers to strategically upvote certain journals was, however, beyond the scope of this paper and is left to further research.

By exploring how academics respond to changing incentives, our paper adds to recent research that analyzes the effects of (changing) incentives in academia (Blomfield & Vakili, 2023; Hudson, 2024; Pietilä, 2019; Śpiewanowski & Talavera, 2021). With our difference-in-differences approach, we are the first to estimate the causal effect of changing incentives, thus substantially extending previous research. Unfortunately, our setting does not allow us to analyze the effect of ranking downgrades (as they were investigated by Śpiewanowski and Talavera (2021) and Hudson (2024)), because there are almost no downgrades in the journal ranking update we study. Hence, the investigation of causal effects regarding the downgrades of a journal rank is an open issue for future research. Similarly, future research might also include the lowest journal rankings to derive a more holistic picture of the effects of a ranking update.

7.4 Implications for Practice

Our study highlights the power of journal rankings in terms of affecting academics' publication behavior. Even though the academics that we focus on in our study will often not receive a direct monetary payment for having published in a journal of a particular rank, the fact that there are other tangible and intangible benefits attached to publishing in a higher-ranked journal does obviously guide academics' publication behavior.

Thus, journal rankings are by no means neutral or innocent, and using them to assess the publication performance of academics is a sensitive issue. If a journal ranking is biased in the sense that it overvalues weak journals, the publication strategy of academics for whom the ranking is relevant will likely be distorted. The risk of a journal ranking being biased and not reflecting the 'true' quality of journals is greatest for journal rankings based on surveys – where survey participants may strategically overstate their assessment of the journals in which they publish or expect to publish. Even though, on average, this does not seem to have happened in the journal ranking update under consideration, department and university leaders are well advised to check the validity of journal rankings before basing, e.g., tenure or appointment decisions on these.

8. References

Aguinis, H., Cummings, C., Ramani, R. S., & Cummings, T. G. (2020). "An a is an a": The new bottom line for valuing academic research. Academy of Management Perspectives, 34(1), 135–154. https://doi.org/10.5465/amp.2017.0193

Anderson, C. G., McQuaid, R. W., & Wood, A. M. (2022). The effect of journal metrics on academic resume assessment. *Studies in Higher Education*, *47*(11), 2310–2322. https://doi.org/10.1080/03075079.2022.2061446

Anninos, L. N. (2014). Research performance evaluation: Some critical thoughts on standard bibliometric indicators. *Studies in Higher Education*, *39*(9), 1542–1561. https://doi.org/10.1080/03075079.2013.801429

Ayaita, A., Pull, K., & Backes-Gellner, U. (2019). You get what you 'pay' for: Academic attention, career incentives and changes in publication portfolios of business and economics researchers. *Journal of Business Economics*, *89*(3), 273–290. https://doi.org/10.1007/s11573-017-0880-6

Bajo, E., Barbi, M., & Hillier, D. (2020). Where should I publish to get promoted? A finance journal ranking based on business school promotions. *Journal of Banking & Finance*, *114*, 105780. https://doi.org/10.1016/j.jbankfin.2020.105780

Bäker, A. (2015). Non-tenured post-doctoral researchers' job mobility and research output: An analysis of the role of research discipline, department size, and coauthors. *Research Policy*, *44*(3), 634–650. https://doi.org/10.1016/j.respol.2014.12.012

Bäker, A., Breuninger, S., & Pull, K. (2021). Pushing performance by building bridges: Human and social capital as mechanisms behind the mobility-performance link. *Journal of Vocational Behavior*, *129*, 103613. https://doi.org/10.1016/j.jvb.2021.103613

Blomfield, M., & Vakili, K. (2023). Incentivizing effort allocation through resource allocation: Evidence from scientists' response to changes in funding policy. *Organization Science*, *34*(1), 100–128. https://doi.org/10.1287/orsc.2021.1565

Braghieri, L., Levy, R., & Makarin, A. (2022). Social media and mental health. *American Economic Review*, *112*(11), 3660–3693. https://doi.org/10.1257/aer.20211218

Buehling, K. (2021). Changing research topic trends as an effect of publication rankings: The case of German economists and the Handelsblatt ranking. *Journal of Informetrics*, *15*(3), 101199. https://doi.org/10.1016/j.joi.2021.101199

Chamberlain, S., & Baker, C. (2021). *microdemic: 'Microsoft Academic' API client*. https://github.com/ropensci/microdemic

Clermont, M. (2016). Effectiveness and efficiency of research in Germany over time: An analysis of German business schools between 2001 and 2009. *Scientometrics*, *108*(3), 1347–1381. https://doi.org/10.1007/s11192-016-2013-3

Drivas, K., & Kremmydas, D. (2020). The Matthew effect of a journal's ranking. *Research Policy*, *49*(4), 103951. https://doi.org/10.1016/j.respol.2020.103951

Eisend, M. (2011). Is VHB-JOURQUAL2 a good measure of scientific quality? Assessing the validity of the major business journal ranking in German-speaking countries. *Business Research*, *4*(2), 241–274. https://doi.org/10.1007/BF03342756

Fernandes, J. M., & Cortez, P. (2020). Alphabetic order of authors in scholarly publications: A bibliometric study for 27 scientific fields. *Scientometrics*, *125*(3), 2773–2792. https://doi.org/10.1007/s11192-020-03686-0

Fernandes, M., Hilber, S., Sturm, J.-E., & Walter, A. (2023). Closing the gender gap in academia? Evidence from an affirmative action program. *Research Policy*, *52*(9), 104865. https://doi.org/10.1016/j.respol.2023.104865

Fernandes, M., & Walter, A. (2022). Publication behavior in different fields of business administration: From anecdotal to empirical evidence. *Schmalenbach Journal of Business Research*, *74*(3), 265–306. https://doi.org/10.1007/s41471-022-00137-9

Fernandes, M., & Walter, A. (2023). The times they are a-changin': Profiling newly tenured business economics professors in Germany over the past thirty years. *Journal of Business Economics*, *93*(5), 929–971. https://doi.org/10.1007/s11573-022-01132-6

González-Pereira, B., Guerrero-Bote, V. P., & Moya-Anegón, F. (2010). A new approach to the metric of journals' scientific prestige: The SJR indicator. *Journal of Informetrics*, *4*(3), 379–391. https://doi.org/10.1016/j.joi.2010.03.002

Harzing, A.-W. (2019). Two new kids on the block: How do Crossref and Dimensions compare with Google Scholar, Microsoft Academic, Scopus and the Web of Science? *Scientometrics*, *120*(1), 341–349. https://doi.org/10.1007/s11192-019-03114-y

Harzing, A.-W., & Alakangas, S. (2017). Microsoft Academic: Is the phoenix getting wings? *Scientometrics*, *110*(1), 371–383. https://doi.org/10.1007/s11192-016-2185-x

He, G., & Wang, S. (2017). Do college graduates serving as village officials help rural China? *American Economic Journal: Applied Economics*, *9*(4), 186–215. https://doi.org/10.1257/app.20160079

Heckman, J. J., & Moktan, S. (2020). Publishing and promotion in economics: The tyranny of the top five. *Journal of Economic Literature*, *58*(2), 419–470. https://doi.org/10.1257/jel.20191574

Hudson, R. (2024). Responding to incentives or gaming the system? How UK business academics respond to the Academic Journal Guide. *Research Policy*, *53*(9), 105082. https://doi.org/10.1016/j.respol.2024.105082

Jokinen, J., & Pehkonen, J. (2017). Promotions and earnings – gender or merit? Evidence from longitudinal personnel data. *Journal of Labor Research*, *38*(3), 306–334. https://doi.org/10.1007/s12122-017-9254-7

Jones, B. F. (2021). The rise of research teams: Benefits and costs in economics. *Journal of Economic Perspectives*, *35*(2), 191–216. https://doi.org/10.1257/jep.35.2.191

Kwiek, M. (2021). What large-scale publication and citation data tell us about international research collaboration in Europe: Changing national patterns in global contexts. *Studies in Higher Education*, *46*(12), 2629–2649. https://doi.org/10.1080/03075079.2020.1749254

Martín-Martín, A., Thelwall, M., Orduna-Malea, E., & Delgado López-Cózar, E. (2021). Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: A multidisciplinary comparison of coverage via citations. *Scientometrics*, *126*(1), 871–906. https://doi.org/10.1007/s11192-020-03690-4

Miller, S., Johnson, N., & Wherry, L. R. (2021). Medicaid and mortality: New evidence from linked survey and administrative data. *The Quarterly Journal of Economics*, *136*(3), 1783–1829. https://doi.org/10.1093/qje/qjab004

Mingers, J., & Willmott, H. (2013). Taylorizing business school research: On the 'one best way' performative effects of journal ranking lists. *Human Relations*, *66*(8), 1051–1073. https://doi.org/10.1177/0018726712467048

Morgan-Thomas, A., Tsoukas, S., Dudau, A., & Gąska, P. (2024). Beyond declarations: Metrics, rankings and responsible assessment. *Research Policy*, *53*(10), 105093. https://doi.org/10.1016/j.respol.2024.105093

Nkomo, S. M. (2009). The seductive power of academic journal rankings: Challenges of searching for the otherwise. *Academy of Management Learning & Education*, *8*(1), 106–112. https://doi.org/10.5465/amle.2009.37012184

Osterloh, M., & Frey, B. S. (2020). How to avoid borrowed plumes in academia. *Research Policy*, *49*(1), 103831. https://doi.org/10.1016/j.respol.2019.103831

Pietilä, M. (2019). Incentivising academics: Experiences and expectations of the tenure track in Finland. *Studies in Higher Education*, *44*(6), 932–945. https://doi.org/10.1080/03075079.2017.1405250

Rafols, I., Leydesdorff, L., O'Hare, A., Nightingale, P., & Stirling, A. (2012). How journal rankings can suppress interdisciplinary research: A comparison between innovation studies and business & management. *Research Policy*, *41*(7), 1262–1282. https://doi.org/10.1016/j.respol.2012.03.015

Roach, M., & Sauermann, H. (2010). A taste for science? PhD scientists' academic orientation and self-selection into research careers in industry. *Research Policy*, *39*(3), 422–434. https://doi.org/10.1016/j.respol.2010.01.004

Salandra, R., Salter, A., & Walker, J. T. (2022). Are academics willing to forgo citations to publish in high-status journals? Examining preferences for 4* and 4-rated journal publication among UK business and management academics. *British Journal of Management*, *33*(3), 1254–1270. https://doi.org/10.1111/1467-8551.12510

Sauermann, H., & Stephan, P. (2013). Conflicting logics? A multidimensional view of industrial and academic science. *Organization Science*, *24*(3), 889–909. https://doi.org/10.1287/orsc.1120.0769

Schrader, U., & Hennig-Thurau, T. (2009). VHB-JOURQUAL2: Method, results, and implications of the German Academic Association for Business Research's journal ranking. *Business Research*, *2*(2), 180–204. https://doi.org/10.1007/BF03342710

Śpiewanowski, P., & Talavera, O. (2021). Journal rankings and publication strategy. *Scientometrics*, *126*(4), 3227–3242. https://doi.org/10.1007/s11192-021-03891-5

Stern, S. (2004). Do scientists pay to be scientists? *Management Science*, *50*(6), 835–853. https://doi.org/10.1287/mnsc.1040.0241

Walker, J. T., Fenton, E., Salter, A., & Salandra, R. (2019). What influences business academics' use of the Association of Business Schools (ABS) list? Evidence from a survey of UK academics. *British Journal of Management*, *30*(3), 730–747. https://doi.org/10.1111/1467-8551.12294

Walker, J. T., Salter, A., Fontinha, R., & Salandra, R. (2019). The impact of journal re-grading on perception of ranking systems: Exploring the case of the Academic Journal Guide and business and management scholars in the UK. *Research Evaluation*, *28*(3), 218–231. https://doi.org/10.1093/reseval/rvz010

Willmott, H. (2011). Journal list fetishism and the perversion of scholarship: Reactivity and the ABS list. *Organization*, *18*(4), 429–442. https://doi.org/10.1177/1350508411403532

9. Appendix

Appendix VII-A: Event Study – Test for Parallel Trends

All (upgraded) journals

	(1) % Papers with at least one Affected Author	(2) % Affected Authors
5 Years Before	-0.0139	-0.0094
	(0.0088)	(0.0076)
4 Years Before	-0.0123	-0.0047
	(0.0077)	(0.0069)
3 Years Before	-0.0017	-0.0019
	(0.0081)	(0.0067)
2 Years Before	0.0109	0.0088
	(0.0078)	(0.0068)
Year of Update	-0.0043	-0.0018
	(0.0083)	(0.0074)
1 Year Later	0.0227**	0.0187**
	(0.0100)	(0.0083)
2 Years Later	0.0140	0.0027
	(0.0097)	(0.0079)
3 Years Later	0.0203**	0.0111*
	(0.0088)	(0.0066)
4 Years Later	0.0180*	0.0107
	(0.0094)	(0.0074)
5 Years Later	0.0092	0.0065
	(0.0084)	(0.0076)
6 Years Later	0.0305**	0.0064
	(0.0137)	(0.0074)
Journal Fixed Effects	Yes	Yes
Observations (journal-years)	4,584	4,584
R^2	0.5971	0.6354

Notes: This table reports the results of our event-study models as described in Section 4.2. The dummy indicating one-year prior treatment status is omitted from the regression. Standard errors are clustered at the journal level and are reported in parentheses. Significance levels are denoted as follows: *p<0.1; **p<0.05; ***p<0.01.

Appendix VII-B: Panel Regressions – Main Effect – Robustness Check (France Based Authors)

All (upgraded) journals

	(1)	(2)
	% Papers with at least one France Based Author	% France Based Authors
Mean LHS	0.0407	0.0257
Post Update	0.0037*	-0.0007
	(0.0020)	(0.0017)
Post Update x Upgraded Journal	0.0048	0.0003
	(0.0032)	(0.0023)
Journal Fixed Effects	Yes	Yes
Observations (journal-years)	4,584	4,584
R^2	0.0027	0.0407
F Statistic	5.7499***	0.1616

Notes: This table reports the results of two difference-in-differences estimations where our dependent variables are based on France based authors instead of authors affected by the journal ranking update. $Post\ Update_t$ is a dummy variable that equals one, for the years 2016 and later. The difference-in-differences estimator is $Post\ Update_t x\ Upgraded\ Journal_i$, which indicates whether France based authors did increase their representation in upgraded journals in consequence of ranking update. In addition, we use journal fixed effects to control for any time-invariant journal characteristics which absorb the coefficient of the dummy variable $Upgraded\ Journal_i$. All models are estimated using standard errors clustered at the journal-level. Significance levels are denoted as follows: *p<0.1; **p<0.05; ***p<0.01.

Appendix VII-C: Panel Regressions – Ranking Heterogeneity – Robustness Checks (France Based Authors)

Panel A: VHB-journal-rank (B) to VHB-journal-rank (A)

	% Papers with at least one France Based Author			% France Based Authors		
Mean LHS	0.0450	0.0473	0.0436	0.0278	0.0268	0.0275
Comparison group: Journals that retained their original … rank	(1) (A)/(B)	(2) (A)	(3) (B)	(4) (A)/(B)	(5) (A)	(6) (B)
Post Update	0.0039*	0.0061**	0.0031	-0.0009	0.0000	-0.0013
	(0.0023)	(0.0031)	(0.0029)	(0.0022)	(0.0024)	(0.0028)
Post Update x Upgraded Journal	0.0051	0.0029	0.0059	0.0028	0.0019	0.0032
	(0.0061)	(0.0064)	(0.0063)	(0.0041)	(0.0042)	(0.0045)
Journal Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations (journal-years)	2,299	785	1,764	2,299	785	1,764
R^2	0.0028	0.0101	0.0022	0.0002	0.0006	0.0003
F Statistic	2.9063*	3.6399**	1.7708	0.2587	0.2231	0.2749

Panel B: VHB-journal-rank (C) to VHB-journal-rank (B)

	% Papers with at least one France Based Author			% France Based Authors		
Mean LHS	0.0391	0.0430	0.0357	0.0257	0.0272	0.0237
Comparison group: Journals that retained their original … rank	(1) (B)/(C)	(2) (B)	(3) (C)	(4) (B)/(C)	(5) (B)	(6) (C)
Post Update	0.0036	0.0031	0.0043	-0.0008	-0.0013	-0.0001
	(0.0025)	(0.0029)	(0.0041)	(0.0016)	(0.0028)	(0.0030)
Post Update x Upgraded Journal	0.0043	0.0049	0.0036	-0.0010	-0.0005	-0.0016
	(0.0036)	(0.0040)	(0.0049)	(0.0035)	(0.0034)	(0.0035)
Journal Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations (journal-years)	3,536	2,256	2,022	3,536	2,256	2,022
R^2	0.0021	0.0032	0.0027	0.0002	0.0004	0.0002
F Statistic	3.3119**	3.3225**	2.4702*	0.2621	0.4260	0.1487

Notes: This table reports the results of our difference-in-differences estimations regarding different levels of journal rankings where our dependent variables are based on France based authors instead of authors affected by the journal ranking update. Panel A compares journals that receive an upgrade from B to A with journals that retain their original B and/or A ranking. Panel B compares journals that receive an upgrade from C to B with journals that retain their original C and/or B ranking. $Post\ Update_t$ is a dummy variable that equals one, for the years 2016 and later. The difference-in-differences estimator is $Post\ Update_t x\ Upgraded\ Journal_i$, which indicates whether France based authors did increase their representation in upgraded journals in consequence of the ranking update as compared to the respective control group. In addition, we use journal fixed effects to control for any time-invariant journal characteristics which absorb the coefficient of the dummy variable $Upgraded\ Journal_i$. All models are estimated using standard errors clustered at the journal-level. Significance levels are denoted as follows: *p<0.1; **p<0.05; ***p<0.01.

Appendix VII-D: Panel Regressions – Ranking Heterogeneity – Upgrades to A+

	% Papers with at least one Affected Author			% Affected Authors		
Mean LHS	0.1268	0.0651	0.1457	0.0884	0.0402	0.1036
Comparison group: Journals that retained their original … rank	(1) (A+)/(A)	(2) (A+)	(3) (A)	(4) (A+)/(A)	(5) (A+)	(6) (A)
Post Update	0.0355***	0.0215*	0.0408***	0.0191***	0.0086	0.0230***
	(0.0078)	(0.0099)	(0.0098)	(0.0057)	(0.0052)	(0.0075)
Post Update x Upgraded Journal	-0.0096	0.0043	-0.0149	-0.0027	0.0078	-0.0067
	(0.0190)	(0.0207)	(0.0199)	(0.0135)	(0.0133)	(0.0144)
Journal Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations (journal-years)	798	263	595	798	263	595
R^2	0.0050	0.0276	0.0635	0.0303	0.0296	0.0353
F Statistic	19.4471***	3.3911**	18.4079***	11.3961***	3.6465**	9.9440***

Notes: This table reports additional results of our difference-in-differences estimations regarding different levels of journal rankings. The table compares journals that receive an upgrade from A to A+ with journals that retain their original A and/or A+ ranking. $Post\ Update_t$ is a dummy variable that equals one, for the years 2016 and later. The difference-in-differences estimator is $Post\ Update_t x\ Upgraded\ Journal_i$, which indicates whether affected authors did increase their representation in upgraded journals in consequence of the ranking update as compared to the respective control group. In addition, we use journal fixed effects to control for any time-invariant journal characteristics. All models are estimated using standard errors clustered at the journal-level. Significance levels are denoted as follows: *p<0.1; **p<0.05; ***p<0.01.

Appendix VII-E: Overlap of Ranking Changes between the VHB-journal-ranking and the AJG

	AJG Constant	AJG Downgrade	AJG Upgrade
Constant VHB-journal-rank	186 journals (47 %)	11 journals (3 %)	114 journals (29 %)
Upgraded VHB-journal-rank	57 journals (14 %)	3 journals (1 %)	29 journals (7 %)

Notes: This cross table reports changes in journal rankings according to the AJG and the VHB-journal-ranking for all journals included in our sample. Columns represent different status changes in the AJG (constant, downgrade, upgrade) and rows represent changes in the VHB-journal-ranking (constant, upgrade). Each cell displays the number of journals that fall into each category, with the percentage of journals in each category provided in parentheses.

Affidavit

Ich erkläre hiermit, dass ich die vorgelegten und nachfolgend aufgelisteten Aufsätze selbstständig und nur mit den Hilfen angefertigt habe, die im jeweiligen Aufsatz angegeben oder zusätzlich in der nachfolgenden Liste aufgeführt sind. In der Zusammenarbeit mit den angeführten Koautoren war ich mindestens anteilig beteiligt. Bei den von mir durchgeführten und in den Aufsätzen erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis niedergelegt sind, eingehalten.

Martin G. Becker
Gießen 13.02.2026

Submitted Papers:

I. Becker, M. G., Stolper, O. & Walter, A. (2026). *Investing by example: Leveraging peer information in digital banking* (Working paper).

II. Becker, M. G., & Walter, A. (2026). *From branches to browsers: A comparative analysis of direct and traditional bank clients in Germany* (Working paper).

III. Becker, M. G., & Walter, A. (2024). Anlageverhalten und Kundenprofile im Vergleich: Unterschiede zwischen Sparkassen, Genossenschaftsbanken und Großbanken. *Zeitschrift für Bankrecht und Bankwirtschaft, 36*(6), 382–392. https://doi.org/10.15375/zbb-2024-0606

IV. Becker, M. G., Maier, T., & Walter, A. (2026). *Perception vs. reality: Wealth disparities between urban and rural households in Germany* (Working paper).

V. Becker, C. C., & Becker, M. G. (2026). *Financial support among siblings: The relevance of personal and family characteristics* (Working paper).

VI. Becker, M. G., Martin, F., & Walter, A. (2022). The power of ESG transparency: The effect of the new SFDR sustainability labels on mutual funds and individual investors. *Finance Research Letters, 47*, 102708. https://doi.org/10.1016/j.frl.2022.102708

VII. Fernandes, M., Becker, M. G., Pull, K., & Walter, A. (2025). When a B becomes an A: Causal evidence on the effects of a journal ranking update on academics' publication behavior. *Studies in Higher Education, 50(12),* 2942–2963. https://doi.org/10.1080/03075079.2024.2447788