Essays in Behavioral Economics

Antal Ertl

Institute of Economics

Supervisor:
Hubert János Kiss, PhD
Institute of Economics, Corvinus University of Budapest

Corvinus University of Budapest

Doctoral School of Economics, Business and Informatics

Essays in Behavioral Economics

Ph.D Dissertation

Antal Ertl

Budapest, 2025

To my family

Contents

[1]This chapter is a joint work with Barna Bakó and Éva Holb.

[2]This chapter is a joint work with Hubert János Kiss.

List of Figures

List of Tables

Introduction

Behavioral and experimental economics have had a remarkable impact on the field of economics in recent decades. On one hand, these fields have renewed the effort to account for psychological factors in economic decision-making—an approach present since Adam Smith's The Theory of Moral Sentiments (A. Smith, 2010; V. L. Smith and Wilson, 2019). John Maynard Keynes also considered such factors, notably with his concept of ”animal spirits” influencing decisions (Keynes, 1937). On the other hand, behavioral and experimental economics have provided new tools—laboratory and field experiments—to test hypotheses about decision-making and the policy effects of interventions. These tools allow economists to test theories on data they generate themselves, rather than relying solely on naturally occurring events. Lab experiments enable economists to isolate external factors, while field experiments test the effects of interventions in natural settings.

In my dissertation, I analyze behavioral factors and use data generated by field and lab-in-the-field experiments. My aim is to showcase the broad range of issues that behavioral and experimental economics can address, from psychological factors in the labor market to how socio-economic status affects economic expectations and the relationship between economic preferences and outcomes.

In Chapter 1, I present a field experiment on the effects of loss aversion in academic performance. Along with Barna Bakó and Éva Holb, we tested whether university students perform better when they are endowed with the maximum achievable points upfront, and points are deducted for incorrect or incomplete tasks, compared to earning points from zero by completing tasks. The rationale is that, due to loss aversion, the negative feelings associated with losing points may be stronger than the positive feelings from earning them. If this effect is present, a simple and cost-effective tool—counting points backwards—could improve student performance.

In Chapter 2, I use survey methods to analyze perceptions of unfairness in workplace inequality. As a preliminary study, I examine whether people are willing to engage in harmful behavior—termed ”malicious envy”—when payoffs are unequal. I also investigate whether providing information about the causes of inequality, such as higher education, more experience, or greater effort, can reduce malicious envy. This study serves as a foundation for a subsequent laboratory experiment on perceived inequality in the labor market, which is still in progress.

In Chapter 3, I analyze the stability of economic preferences using data from a representative survey. While the stability of preferences is well-documented, I also investigate whether clusters of economic preferences remain stable over time. Specifically, I test whether similar clusters emerge in two waves of the survey. I use two methods: first, I apply the same clustering algorithm to both waves and compare the results. Additionally, I examine demographic and socio-economic differences, including gender. Second, – given that cluster analysis is highly sensitive to data–, I also project the second wave's data onto the results of the first wave to check for socio-economic and demographic differences while keeping the cluster composition constant.

In Chapter 4, together with Hubert János Kiss, I analyze how socio-economic status affects economic expectations, using representative monthly survey data from 2000 to 2010. We test whether higher income and education levels lead to more optimistic economic outlooks, and whether recessions have differential impacts across socio-economic strata. We also conduct several robustness checks.

These essays are linked not by their themes but by their methodological approaches, all made possible by advances in behavioral economics. Chapter 1 employs a field-experimental approach, while Chapter 3 uses measurements developed over the last two decades in behavioral economics. Chapter 2 uses survey methods to measure psychological effects in a hypothetical setting, while Chapter 4 applies a more traditional economic framework with a behavioral approach to a macroeconomic question.

The essays in this dissertation are relevant to their respective fields. Chapter 1 explores a cost-effective and simple method to improve student performance. Understanding whether this method is neutral—meaning it does not affect students differently based on gender or ability—is crucial, and existing literature is unclear on this. Furthermore, we are the first to test whether the impact of loss-framing diminishes over time, which has important implications for scalability.

Analyzing how people perceive and react to workplace inequalities is particularly relevant as income transparency gains support. While income transparency has been shown to have positive effects, such as increased productivity (Gutierrez et al., 2022) and a reduced gender wage gap (Bamieh and Ziegler, 2024; Bennedsen et al., 2023; Böheim and Gust, 2021), recent research suggests that its effects may be heterogeneous. For instance, income transparency might demotivate workers when they learn about peers' higher income, while motivating them in relation to supervisors' income (Cullen and Perez-Truglia, 2022). Understanding the mechanisms of income comparison is essential for assessing whether the overall effects are positive or negative.

The stability of economic preferences has been widely studied, especially regarding risk and time preferences. However, research on the stability of clusters of economic preferences is scarce. To my knowledge, no study has investigated whether these clusters remain stable over time or whether similar groupings reemerge. Chapter 3 addresses this novel question. Finally, Chapter 4 investigates whether socio-economic background influences economic expectations. If so, policymakers may need to account for this when designing targeted interventions. Furthermore, this study contributes to understanding a specific channel in the formation of economic expectations. The remainder of the thesis proceeds with the essays themselves, followed by a brief conclusion.

Acknowledgments

This dissertation is a result of years of reading, discussions, learning and writing. I would like to thank all of those who helped me along the way. First of all, I thank Hubert János Kiss for helping me, both professionally and on more personal matters as well: I thoroughly enjoyed our discussions, and I am very thankful for the support I received from him. I know that without him, finishing the PhD would not have been such a good experience.

I thank colleagues from Corvinus University of Budapest for their comments and conversations. I am especially grateful to Dániel Horn, Barna Bakó, and Áron Horváth for giving me the opportunity to work with them. I have learned a great deal from them—be it scientific writing, asking the right questions, or tackling complex issues in a methodologically sound way. I also thank Éva Holb, Péter Horváth, Ágoston Reguly and Gergely Hajdu for discussions and comments. I thank Ágnes Pintér and Philipp Külpmann for their valuable comments, suggestions, and critiques on the draft version.

I thank my parents for their unwavering support throughout my studies, for their encouragement, and for passing on a share of their strong work ethic. I thank my brother for helping our family, even with the smallest of things, and for being the best (and yes, the funniest) little brother I could have asked for.

Finally, I would like to thank my wife, Csenge, for everything. She encouraged me through my doubts, supported me during the hardest moments, and celebrated my achieve-

ments. Without her, this thesis may never have been completed.

Chapter 1

Learning to Win by Fearing to Lose

Exploring the Positive Effects of Loss Aversion on Academic Achievement and Motivation in Education[1]

1.1 Introduction

According to Kahneman and Tversky, 1979, individuals perceive losses as more impactful than equivalent gains, meaning the emotional impact of losing something is greater than the satisfaction of gaining the same thing. This phenomenon, known as loss aversion, suggests that people are more motivated to avoid losses than to seek gains. Tversky and Kahneman, 1991 further supported these findings, while Novemsky and Kahneman, 2005 demonstrated that loss aversion is evident across various domains, including consumer behavior, finance, and health. Furthermore, Mrkva et al., 2020 argue that although certain factors can mitigate its effects, loss aversion remains a significant and measurable phenomenon.

In the context of education, loss aversion offers a compelling framework for enhancing academic performance. The basic idea is that if students are more motivated by the fear of losing points than by the prospect of gaining them, reframing the grading system to emphasize losses could lead to improved performance. By granting students a full score and deducting points for incorrect answers – rather than allowing them to accumulate points for correct responses – educators can potentially engage students' natural aversion to losses, thereby motivating greater effort and focus. The psychological pressure of trying

[1]This chapter is a joint work with Barna Bakó and Éva Holb.

to avoid losses could result in students being more attentive to their studies, thus enhancing both engagement and performance.

Several studies have examined whether students perform better when they start with a full score and lose points for incorrect answers during a test. For example, Bies-Hernandez, 2012 found that framing results as losses negatively impacted learning, and the use of a point deduction system significantly impaired students' perceptions of the course. However, as Smith et al., 2019 points out, the study only analyzed raw results without accounting for other socio-demographic and academic factors.

Apostolova-Mihaylova et al., 2015 conducted an experiment in an economics course where students were divided into two groups: a control group, which earned points for correct answers, and a treatment group, which started with maximum points and lost them for incorrect responses. They found no significant difference in overall performance between the two groups but did observe a notable gender disparity: females in the loss-framed group performed worse than males. In contrast, McEvoy, 2016, using a similar experimental design, found that students subjected to loss-framed grading performed better as the semester progressed, without any gender differences. Smith et al., 2019 further supported the idea that framing scores as losses can improve academic achievement, particularly when controlling for students' pre-existing knowledge and proficiency.

Another study by Faulk et al., 2019 tested loss aversion using a different approach. In their experiment, students could earn points for attending seminars, but the treatment group was told that they had all the points at the start and would lose them for non-attendance. This design confirmed that loss framing can indeed motivate positive behaviors beyond test performance, as students in the loss-framed group were more likely to attend seminars regularly.

Meanwhile, Levitt et al., 2016 explored both financial and non-financial incentives for improving test scores in primary and secondary school students. Although they found that incentives generally improved scores, there was no evidence that loss-framed incentives led to greater effort. Moreover, they identified heterogeneous effects, particularly with boys responding more strongly to non-financial rewards than girls, hinting at a nuanced interplay between gender, incentives, and loss aversion.

Roy and Lewis, 2023 extended the application of loss aversion into the domain of goal-setting. They examined whether framing academic goals as gains or losses affected performance and found that both approaches improved outcomes, though there were no significant differences between them. This suggests that goal-setting itself can enhance academic performance, but the framing of those goals may not always make a critical difference.

As shown by these results, the findings are mixed: some studies find significant differences between treatments, while others report heterogeneous treatment effects based on gender. However, none of the studies have examined the distinction between medium-term and short-term effects. Specifically, it remains unclear whether the observed effect is due to the act of changing the status quo – shifting from 'gaining points' to 'losing points' – a concept that may be unfamiliar to students. One could argue that this reframing might be effective in the short term, but over time, students may realize that the evaluation criteria remain the same, with only the framing of the scoring being altered.

In this paper, we address three key issues related to the impact of loss-framing in scoring. Firstly, we investigate whether loss-framing positively influences students' performance. To test this, we conducted a field experiment in which students were graded under one of three scoring methods: (1) control (or *Gain*) treatment, where scoring followed the standard procedure and students gained points for correct answers; (2) *Loss* treatment, where students were initially awarded the maximum possible points and subsequently lost points for incorrect answers; and (3) *Hybrid* treatment, where students earned points as usual during the semester but were subject to the loss treatment for the final exam. Our results indicate that loss-framing has a statistically significant positive effect on learning outcomes, as measured by the points scored. This effect is evident both during the semester and in the final exam. Specifically, the treatment effect ranges from 2.5 to 5 percentage points during the semester and from 7 to 9 percentage points in the final exam. These findings align with prior studies on loss-framing (see McEvoy, 2016, Smith et al., 2019). We argue that these are significant effects, especially considering the low cost of implementing the treatment, since there is no need to alter the curriculum or the tests, only the feedback system.

Secondly, we examine whether the effect of loss-framing diminishes over time. One could argue that the impact of loss-framing might be temporary. Early in the semester, students might be motivated by the fear of losing points, but over time, they could recognize that the treatment is merely a reframing of the scoring system or adapt to the loss-framing as the new status quo. To test for this potential 'novelty effect', we leveraged the *Hybrid* treatment group, where students experienced loss-framing only for the final exam. If the novelty effect exists, students encountering loss-framing for the first time in the final test should outperform those subjected to it throughout the semester. Our results show that during the semester, students in the *Loss* treatment group performed about 11 percentage points better than those in the *Hybrid* group, a statistically significant difference. However, for the final test, this difference diminishes and is no longer statistically significant. After controlling for students' performance during the semester, the effect decreases further. This finding challenges the idea that students will adapt to loss framing and become less responsive to it as time passes and supports our broader conclusion that the effectiveness of loss-framed grading does not depend on its novelty; rather, it has a sustained impact on performance that is consistent over time.

Thirdly, we explore heterogeneous effects of loss-framing, which is critical for policy implementation. Ideally, we aim for a treatment that produces a Pareto improvement, meaning no subgroup is negatively affected. We first examine gender differences. While female students generally score higher during the semester regardless of the treatment, we find no statistically significant gender-specific effects of the treatment itself. This aligns with the findings of McEvoy, 2016, and challenges the results by Apostolova-Mihaylova et al., 2015 who reported heterogeneous effects by gender. Next, we test for heterogeneity based on students' prior skills. Comparing students who had taken advanced math classes in secondary school with those who had not, we find that the treatment effect is greater for students without advanced math training. Nonetheless, the effect remains positive for both groups. Finally, we examine quantile treatment effects. While the treatment is not statistically significant across all quantiles, the positive sign of the effect suggests no adverse outcomes for any subgroup. These results are particularly encouraging as they provide reassurance about the equity of such interventions.

The remainder of the paper is organized as follows: Section 2 describes the experimental design; Section 3 presents the data and discusses the main results of the analysis; and finally, Section 4 concludes.

1.2 Experimental Design

To carry out our study, we conducted a field experiment involving first-year business students enrolled in a mandatory Macroeconomics for Business course at Corvinus University of Budapest, Hungary. The course was taught weekly, consisting of one 90-minute lecture and one 90-minute practical session, over 13 weeks during the spring semester of the 2022/2023 academic year.

We divided the participants into three groups as follows: (1) *Gain* group: students earned points throughout the semester by completing tasks. For each correct answer, they gained points; (2) *Loss* group: students began the semester with the maximum possible score (100 points) and lost points for each incorrect answer. In both tests and the final exam, results were framed as losses (e.g., ”-5 points” instead of ”5 out of 10”); and (3) *Hybrid* group: students earned points throughout the semester similar to the *Gain* group. However, for the final exam, they started with the maximum achievable score and lost points for each incorrect answer.

The effects of the first two treatments are well-documented in the literature, as they have been utilized in prior experiments (see for example Apostolova-Mihaylova et al., 2015 or McEvoy, 2016). The third treatment, however, introduces a novel approach that allows us to explore whether there is a novelty effect associated with loss aversion. In the *Loss* treatment, students encounter their first losses during the third week of the semester when they receive feedback on their performance. Over time, these students might adapt to losing points rather than gaining them, perceiving the system as merely a reframing of their efforts. This shift in the status quo could potentially lead to a decline in the effectiveness of loss-framing as the semester progresses. The *Hybrid* treatment provides a unique opportunity to examine this phenomenon. Students in this group only experience loss-framing during their final exam, which accounts for 40% of their overall grade. By comparing the performance of students in the *Loss* treatment with those in the *Hybrid* group, we can assess whether the effects of loss aversion diminish over time or remain consistent.

The course included 14 practical sessions, offered in various time slots from Tuesday to Friday, with each session taught by the same professor throughout the semester. The course lectures, on the other hand, were led by a single professor. To ensure fairness in the treatment assignment, we employed a stratified randomization approach: within each professor’s set of practical sessions, the treatment conditions (*Gain*, *Loss*, and *Hybrid*) were assigned randomly.

Students’ grades were determined based on the following components:

1. Weekly Homeworks (not framed): Starting from the second week of the semester, students completed weekly homework assignments, each worth one point, for a total of 12 points. These assignments were submitted through the course’s online learning platform. However, since the platform did not support negative points, the homework scoring remained the same for all groups.

2. Semester Tests: Four tests were held during the practical sessions throughout the semester, each worth 16 points. The tests included calculations, open-ended theory questions, and graph analysis. Only the three highest scores out of the four tests

were counted, contributing a maximum of 48 points. Retakes were not allowed for these tests.

3. Final Exam: At the start of the exam period, students took a final exam worth 40 points, covering material from the entire semester. This exam consisted solely of multiple-choice questions. Like for the semester tests, there were no retakes allowed for the final test. The final exam was conducted in larger auditoriums at the same time for all treatment groups, with multiple standardized versions of the test. Students who failed the final exam had the option to take a comprehensive exam; however, our analysis does not include this.

During their first lecture, students were informed that they would participate in an experiment designed to evaluate a new teaching methodology, and detailed explanations of their specific treatment were provided during their practice group sessions. Additionally, treatment-specific syllabi were uploaded to the course's online learning platform, where all other study materials were made available to the students. To ensure students understood the treatment they were assigned to, we uploaded Excel files to each practical session's online learning platform at the beginning of the semester. These files presented examples of how their grades would be calculated. Figure 1.1 shows an example for the *Loss* treatment, where points were framed as losses. Similar examples were provided for the *Gain* and *Hybrid* treatments. Students received their results in the same Excel file, where they were identified solely by their University IDs (i.e. 'Neptun-codes').

	A	B	C	D	E	F	G	H
1	**Name**	**Neptun-code**	**1.test (max: 16 points)**	**2.test (max: 16 points)**	**3.test (max: 16 points)**	**4.test (max: 16 points)**	**40 point exam**	**Total**
2	**John Doe**	**AAA123**	-2	-3	-1	-13	-10	72
3								
4			John Doe during the 16-point seminar tests, lost 2 points on the first test, 3 points on the second, 1 point on the third and did not prepare for the fourth, losing 13 points. Thus, counting the three best tests, he lost a total of 6 point from his seminar tests. Finally, on the 40 point exam, he lost 10 points, leaving him with 72 points excluding homework.					
5								
6								
7								
8								
9								

Figure 1.1: "Loss" treatment example from the Excel file provided to students to clarify how their grades were being calculated

The tests administered during the practical sessions were standardized across all sessions, with only the framing of points differing by treatment. The evaluation process was also standardized to ensure uniform grading. Professors were not blind to the treatment conditions, as they needed to adjust their feedback to match the treatment. For example, in the loss-framing group, professors would write -5 for deducted points instead of presenting the score as $11/16$. Despite this, all evaluations followed a standardized grading approach to ensure fairness. Professors coordinated with each other to ensure that partial answers were graded consistently across all practical sessions, regardless of the treatment condition.

The practical sessions were held on different days, which could have influenced student's performance. Students taking tests on Fridays may have had more time to prepare compared to those who took tests earlier in the week, such as on Tuesday. Additionally, the structure of the tests could have been shared among students, potentially giving an advantage to those in later sessions. To account for this, we controlled for the day of the week in our regression analysis.

At the beginning and end of the semester, students completed background questionnaires which allowed us to account for factors that might influence performance when comparing treatment and control groups. The first questionnaire was completed during the first practical session. It asked students about their demographic information and their university studies, with a focus on Macroeconomics for Business, a course that typically produces a wide distribution of grades. This questionnaire also included a consent form for using their data in our research.[2] The second questionnaire was administered after students took the final exam, during the examination period. It included questions about their feedback on the grading system, the amount of time dedicated to studying the course, and whether they were aware of the other treatment conditions. Although only a limited number of students responded, it allowed us to verify that the treatments did not create feelings of unfairness and that there was no significant contamination (i.e., students knowing about other treatments).[3] Both questionnaires were administered online via Qualtrics.

1.3 Data and Results

A total of 461 students enrolled in the course, of which 370 consented to the use of their data by completing our first questionnaire. The analysis is based on 321 observations

[2] A copy of the first questionnaire can be found in Section 9 of the Appendix.

[3] The analysis of the second questionnaire is provided in Section 6 of the Appendix.

after excluding those with missing data. These 321 students completed the first questionnaire and provided sufficient information for filtering out undesired mechanisms. For robustness checks, we conducted our analysis on all test score data, which includes 408 observations (see Appendix 4).

Table 1.1: Summary Statistics

Statistic	N	Mean or %	Median	St. Dev.	Min	Max
Test 1 (16 pts), %	321	63.094	68.750	24.584	0.000	100.000
Test 2 (16 pts), %	321	58.411	62.500	25.937	0.000	100.000
Test 3 (16 pts), %	321	64.272	68.750	27.603	0.000	100.000
Test 4 (16 pts), %	321	59.112	62.500	32.318	0.000	100.000
Best 3 Tests (48 pts), %	321	68.627	73.958	21.421	2.083	98.958
Final Test (40 pts), %	321	63.723	70.000	25.097	0.000	100.000
Homeworks (12 pts), %	321	87.583	90.250	12.878	22.917	100.000
Total Score (100 pts), %	321	68.940	72.250	18.696	3.750	96.250
Gain	321	0.340	0	0.474	0	1
Loss	321	0.396	0	0.490	0	1
Hybrid	321	0.265	0	0.442	0	1
Mother's Educ. University	321	0.754	1	0.431	0	1
Knows to take derivatives	321	0.595	1	0.492	0	1
Female	321	0.526	1	0.500	0	1
Does not work	321	0.439	0	0.497	0	1
Classes this semester	321	6.869	7	1.076	4	14
Credits this semester	321	33.822	34	3.719	0	42
Tuesday Session	321	0.287	0	0.453	0	1
Wednesday Session	321	0.287	0	0.453	0	1
Thursday Session	321	0.047	0	0.211	0	1
Friday Session	321	0.427	0	0.495	0	1

Summary statistics are shown in Table 1.1. Generally, the test scores appear left-skewed, with the median higher than the average. The average test scores are around 60% of the maximum achievable points (between 9.3 and 10.2 out of 16), while the average for the final test is around 64% (25.5 out of 40). Regarding homework, no student scored zero, indicating that all students put at least minimal effort into the course throughout the semester. However, since the homework assignments were simple and unframed, they are not a reliable measure of students' time investment in the course. As expected, differences in homework completion were not statistically significant between treatments.

Although the treatment assignment was balanced across groups, the distribution of questionnaire responses varied: 34% were in the *Gain* treatment, 39.6% in the *Loss* treatment, and 26.5% in the *Hybrid* treatment. Additionally, 75% of students reported that their mother had at least a university degree, which is somewhat expected given that Corvinus University of Budapest is Hungary's most prestigious business school.

The gender distribution in the sample is balanced, with 52.6% of the participants being women. On average, students earned around 34 credits during the semester, slightly exceeding the required 30 credits. Students could choose from various time slots for the practical sessions, with Friday offering the most options and Thursday only one. Additionally, based on prior experience, we hypothesized that first-year students who had learned how to take derivatives in secondary school might achieve better grades. This could also serve as a proxy for their mathematical knowledge, as in the Hungarian education system, only students who took higher-level math classes learn about derivatives. To account for this factor, we asked students if they had learned to take derivatives before university. In our sample, approximately 60% of students had this prior knowledge.

Table 1.2 presents the covariate balance table to assess the random assignment of treatments. It includes p-values from t-tests comparing the means of each group. The results show a significant difference between groups regarding the day of the practical sessions. This is important because students who took their tests later had more time to prepare and might have received information about the test structure from students in earlier sessions, as the tests were standardized. To account for this, we include the day of the practical session and the instructor as control variables in our analysis. Additionally, we observe that the *Gain* group has a lower proportion of students who do not work, suggesting that the differences between groups may be due to the students' work habits rather than the treatment itself. Therefore, we control for employment status in our regression analysis.

As mentioned earlier, the grading structure of the course was such that, out of the four tests taken throughout the semester, only the three best scores contributed to the final grade. [4] As a result, the observed shift to the right in the *Loss* group's scores for the second test may be due to strategic decision-making rather than behavioral changes. To address this, we focus our analysis on the three best test scores, as this is the score students aim to maximize, thereby eliminating any influence of strategic decision-making.

Figure 1.2 presents boxplots showing the total points earned from the three best tests, broken down by treatment group. The distributions for the *Hybrid* and *Gain* groups are similar, while the interquartile range (from the 25th to the 75th percentile) for the *Loss* group is narrower. Additionally, the median score for the *Loss* group is higher — 37 points — compared to 34.5 points for the *Hybrid* group and 32.0 points for the *Gain* group.

[4] This approach was adopted for administrative simplicity; it allowed the course coordinator to avoid organizing separate re-take exams.

Table 1.2: Balance of Dataset

	Treatment:			P-value of T-test	
	Gain (N=109)	Loss (N=127)	Hybrid (N=85)	Gain vs. Loss	Gain vs. Hybrid
Mother's Educ. University	0.734 (0.444)	0.748 (0.436)	0.788 (0.411)	0.806	0.384
Knows to take derivatives	0.569 (0.498)	0.630 (0.485)	0.576 (0.497)	0.341	0.915
Female	0.477 (0.502)	0.559 (0.498)	0.541 (0.501)	0.210	0.378
Does not work	0.358 (0.482)	0.512 (0.502)	0.435 (0.499)	0.017	0.275
Classes this semester	6.789 (1.010)	6.874 (0.943)	6.965 (1.322)	0.505	0.295
Credits this semester	33.688 (3.385)	33.827 (3.288)	33.988 (4.656)	0.750	0.604
Tuesday Session	0.000 (0.000)	0.606 (0.491)	0.176 (0.383)	<0.001	<0.001
Wednesday Session	0.404 (0.493)	0.205 (0.405)	0.259 (0.441)	<0.001	0.035
Thursday Session	0.000 (0.000)	0.000 (0.000)	0.176 (0.383)	-	<0.001
Friday Session	0.596 (0.493)	0.189 (0.393)	0.565 (0.499)	<0.001	0.660

During the final exam, the *Hybrid* treatment was adjusted: students, similar to the *Loss* treatment, began with a maximum of 40 points and lost points for incorrect answers. This change was clearly communicated to the students and was emphasized in the syllabus provided at the start of the semester. [5] Figure 1.3 shows the box-plot distribution for the final exam. Comparing it to Figure 1.2, we observe a noticeable shift in the *Hybrid* treatment. The median score for the final exam is around 30 for both the *Loss* and *Hybrid* treatments, where students could only lose points, while it is around 26 for the *Gain* treatment. Additionally, 30 students scored zero on the final exam. Once again, the data suggests that students in the *Loss* and *Hybrid* treatments performed better, on average, than those in the *Gain* treatment.

While these descriptive statistics support our hypotheses, it is essential to account for other factors that could influence the results. In the following, we present our regression analysis to examine these findings in greater detail.

[5] We assume that, if regulations allowed, the recency effect would be more pronounced, leading to higher scores on the final exam.

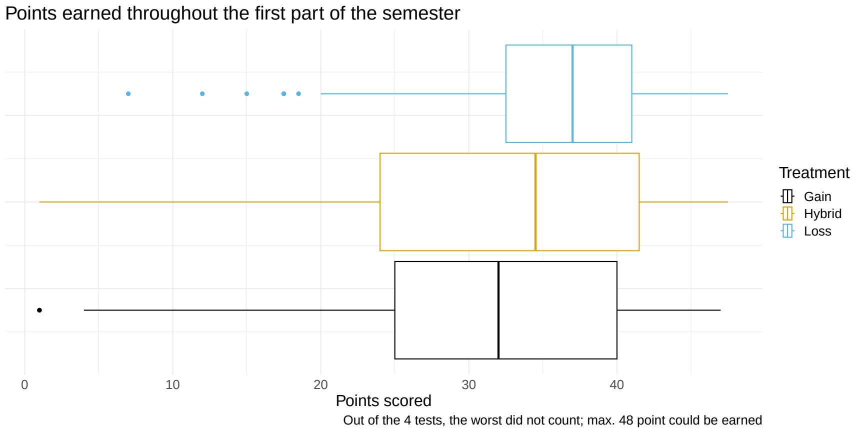

Figure 1.2: Box-plots for the best three tests by treatment groups. At this stage, the *Hybrid* treatment is the same as the *Gain* treatment.

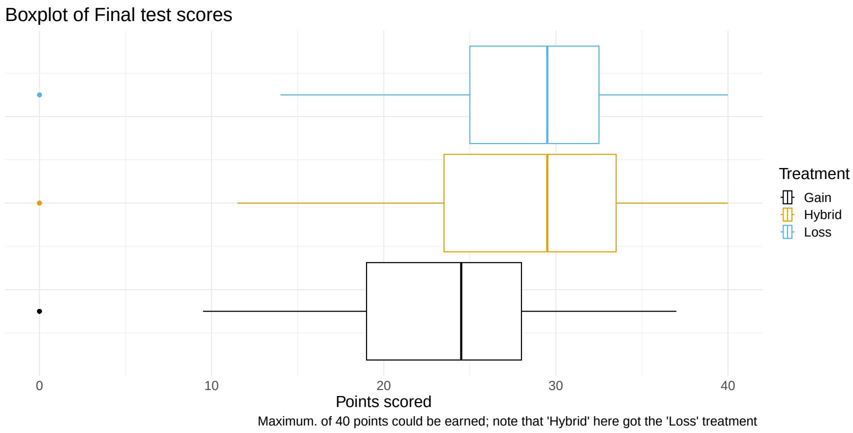

Figure 1.3: Box-plots for the Final test scores for each treatment group. At this stage, the *Hybrid* treatment is changed to correspond to the *Loss* treatment.

1.3.1 The Impact of Loss Aversion on Learning Outcomes

As shown in Table 1.2, the dataset was balanced across treatments, except for two variables: the time slot of the practical sessions and part-time work. Controlling for the time slot is essential for analyzing the tests conducted during the semester. While the tests were standardized across sessions, students who wrote them on Fridays might have had an advantage by obtaining information from earlier sessions. Similarly, working part-time alongside university studies could also influence academic performance.

Table 1.3: Regression results for tests written throughout the semester and the Final Exam - In percentages

	Test 1	Test 2	Test 3	Test 4	Best 3 Tests	Final Test	
Losing Points	-1.689	6.661***	4.375	7.578***	4.912**	7.909**	7.820***
	(3.472)	(0.775)	(2.618)	(1.229)	(1.932)	(2.856)	(1.583)
Female	5.485**	5.778**	7.388*	4.481	5.953**	2.418	-0.974
	(1.826)	(1.882)	(3.879)	(3.608)	(2.543)	(2.973)	(1.765)
Best 3 Tests							0.558***
							(0.113)
Constant	63.63***	66.73***	64.58***	59.73***	69.65***	58.28***	16.75*
	(3.508)	(2.999)	(6.366)	(5.485)	(3.595)	(6.382)	(8.450)
Observations	321	321	321	321	321	321	321
R-squared	0.154	0.172	0.124	0.094	0.162	0.160	0.351
Residual Std. Error	23.084	24.092	26.374	31.412	20.023	23.488	20.673

Standard errors in parentheses, clustered by seminar group;* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Control variables include: mother's university education, knowing how to take derivatives, working or not, and teacher, practice-session and campus fixed effects. Note that *Losing points (!)* correspond to students losing points during the semester tests and during the final test, respectively.

First, we examine whether losing points affects learning outcomes. [6] We compare students who gained points to those who lost points for each test. [7] Table 1.3 presents the regression results for each test, as well as the final exam. The dependent variable is the percentage of points scored on each test. For regressions (1)-(4), we observe that the treatment effect varies. In regression (1), there is a negative effect, while for all subsequent tests, the effect is positive. Overall, the impact on scores across the four tests is inconclusive.

[6] In our analysis, we use clustered standard errors at the practical session-level. This approach is appropriate because treatments were assigned at the session level following stratified randomization by professors. Recent econometric research emphasizes the importance of using clustered standard errors in such cases (Abadie et al., 2023).

[7] Note that one of the treatment groups, the *Hybrid* group, only lost points during the final exam. As such, in that regression, they are considered treated, while in regressions (1)-(5), they are not.

Since only the three best test scores out of the four tests written during the semester count toward the final grade, analyzing individual tests may be biased due to strategic decision-making. Students might aim to maximize their best three scores instead of performing well on each individual test. To address this, in regression (5), we test the effect on the Best 3 Tests, eliminating the issue of strategic decision-making. [8] Our results show a positive and significant effect: on average, students who were losing points throughout the semester performed 4.9 percentage points better on the Best 3 Tests compared to students in the control group, who were accumulating points. This is a substantial effect, especially given that the treatment was a simple framing change and no other teaching methods were altered.

One potential concern is the demand effect on the professors, due to the open-ended nature of the questions and the teachers' role in grading. Although the tests were standardized and teachers communicated on how to grade, we cannot entirely eliminate the possibility of bias in the grading process. However, we can rule out this demand effect for the Final Exam, as it consisted of standardized multiple-choice questions and was taken by all students in the same time slot.

When analyzing the treatment effect on the Final Exam, we find that, on average, students who were losing points performed 7.9 percentage points better than their peers in the control group. One might argue that this difference in final test scores is not solely due to the treatment but could also be influenced by the students' performance throughout the semester. Specifically, students in the *Loss* treatment may have performed better on the semester tests, which in turn contributed to their better final test performance. However, when controlling for performance during the semester (Best 3 Tests), we find that this does not significantly alter the results. This suggests that the observed treatment effect on the final test is primarily driven by the loss-framing rather than prior performance. [9]

For robustness checks, we ran the same regressions but excluded students who likely dropped out during the semester—specifically, those who scored zero points on the Final Exam. Since dropouts in our sample are treatment-dependent (out of 30 dropouts, 25 were in the *Gain* treatment), excluding these students would likely reduce the observed treatment effect. However, it is important to exclude them to ensure that our results are not driven solely by dropout behavior. After omitting these students, the effect remains

[8] Note, however, that this grading method reduces the differences between students. If we compare the total score from all four tests to the score from the best three tests, the standard deviation of the latter is smaller.

[9] We ran these regressions without control variables in Appendix 2 and find an even greater treatment effect.

Table 1.4: Running the regressions of Table 1.3 with only students in the *Gain* and *Loss* treatments.

	Test 1	Test 2	Test 3	Test 4	Best 3 Tests	Final Test	
Loss	-3.837*	7.232***	1.296	4.163***	2.652***	8.994***	7.200***
	(1.692)	(0.482)	(1.241)	(0.750)	(0.598)	(0.997)	(0.880)
Female	4.529*	3.704	5.364	5.311	4.780	2.690	-0.543
	(2.347)	(2.793)	(4.818)	(3.998)	(3.128)	(3.852)	(2.029)
Best 3 Tests							0.676***
							(0.135)
Constant	66.88***	64.50***	72.43***	62.54***	72.77***	55.46***	6.237
	(3.316)	(4.131)	(5.985)	(5.931)	(4.277)	(5.915)	(10.54)
Observations	236	236	236	236	236	236	236
R-squared	0.208	0.168	0.126	0.099	0.163	0.199	0.451
Residual Std. Error	21.766	23.974	25.556	30.502	19.152	23.074	19.137

Standard errors in parentheses, clustered by seminar group;* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Control variables include: mother's university education, knowing how to take derivatives, working or not, and teacher, practice-session and campus fixed effects.

statistically significant for the Final Exam, with students who lost points scoring at least 5.2 percentage points higher compared to the control group. A detailed analysis of the dropouts can be found in Section 3 of the Appendix.

Another way to analyze the treatment effect is to compare students who were losing points throughout the semester with those in the control group (excluding the *Hybrid* treatment). Table 1.4 presents these results. Overall, the effect size for the semester tests decreases significantly (from 4.9 percentage points to 2.6), but it remains statistically significant at the 1 percent level. For the Final Exam, after controlling for performance in the semester tests, the effect size increases to 7.2 percentage points, which is highly statistically significant.

1.3.2 Testing the Novelty Effect

Our findings suggest that loss-framed grading improves students' performance, with stronger evidence when using standardized multiple-choice questions. Next, we aim to test whether the effect of loss-framing diminishes over time. One potential reason for this is that students may initially struggle with the idea of losing points, as they are not accustomed to this form of grading. The treatment could have a stronger effect at the beginning, as students realize they are losing points. However, over time, they may realize that nothing has truly changed – they are still earning points, only the feedback is different. Although our study has a relatively short time-frame (with the semester starting in February and the Final Exam taken at the end of May), it is plausible that students may come to understand that the loss-framing is merely a re-framing of scores. Therefore, we would expect the effect to be weaker by the end of the semester.

To test whether a novelty effect is present in loss-framed grading, we make use of our treatment group where students were losing points only once – during the Final Exam. If a novelty effect is indeed at play, then comparing the group who only encountered losing once to the treatment where students were losing throughout the whole semester, we would expect that: (i) there is a significant difference in scores for tests taken throughout the semester, with those students who are losing points earning more points; and (ii) for the Final Exam, if the loss-framing effect diminishes over time (i.e., the novelty of the grading system drives the effect), students who experience losing points for the first time should perform better.

Table 1.5: Testing the Novelty Effect: Comparing students in the Loss (treatment estimate) and Hybrid (control) groups

	Test 1	Test 2	Test 3	Test 4	Best 3 Tests	Final Test	
Loss	7.783**	8.116***	9.951**	15.10***	11.32***	5.458	2.128
	(2.872)	(0.400)	(3.854)	(1.970)	(2.504)	(4.939)	(5.407)
Female	4.955**	6.072**	7.882*	2.165	5.156**	0.584	-0.932
	(1.930)	(2.156)	(4.102)	(2.055)	(1.512)	(2.145)	(1.877)
Best 3 Tests							0.294*
							(0.155)
Constant	52.31***	63.89***	55.01***	53.99***	61.96***	61.57***	43.36***
	(4.331)	(4.048)	(7.225)	(4.611)	(3.255)	(7.193)	(9.291)
Observations	212	212	212	212	212	212	212
R-squared	0.184	0.163	0.151	0.118	0.192	0.069	0.128
Residual Std. Error	22.303	23.421	25.810	30.423	18.648	21.709	21.059

Standard errors in parentheses, clustered by seminar group;* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Control variables include: mother's university education, knowing how to take derivatives, working or not, and teacher, practice-session and campus fixed effects.

Table 1.5 presents the results using the same specifications as before. In this analysis, the control group consists of students who lost points only on the Final Exam, while the treated group includes students who lost points throughout the semester. For the tests taken throughout the semester, the difference in performance is large and statistically significant: students who were losing points scored, on average, 11 percentage points higher than those in the *Hybrid* group (who lost points only on the Final Exam). This effect is even larger than what we observed in Table 1.4.

When examining the Final Exam, where the control group also lost points, we find that the difference between the two groups decreases and is no longer statistically significant. Taking into account student's performance throughout the semester further reduces this effect. This suggests that the effect is not due to novelty, as students who encountered the loss-framing for the first time did not outperform those who experienced it throughout the semester, and the difference between the two treatments is not statistically significant. [10]

Comparing these treatment effects with earlier results, we find that the effect for the tests written throughout the semester is significantly higher than in Tables 1.3 and 1.4. In Section 4 of the Appendix, we explore this phenomenon further. We hypothesize that this could be due to a selection effect in the *Hybrid* treatment, which might cause larger differences between the two treatment groups. As a result, this leads to a lower estimate when testing for the diminishing effect (i.e., the difference between the two groups is greater than in the full sample).

1.3.3 Tests of Heterogeneity

Overall, we find that losing points positively impacts study outcomes, with minimal cost – only a change in how students perceive their scores. However, it is essential to examine heterogeneous treatment effects, as one group may benefit while another is harmed. Specifically, we look at gender differences, as prior studies show mixed results on how loss aversion affects males and females. Additionally, we assess whether low-performing students are disadvantaged by the intervention.

Gender Heterogeneity

Throughout the analysis, we found a significant positive difference between females and males, with all other factors held constant. As shown in Table 1.3, female students scored, on average, 6 percentage points higher during the semester compared to their male peers, but there were no significant differences on the Final Exam. This suggests that females tend to perform better on tests throughout the semester, but it does not indicate whether the treatment effect is heterogeneous.

The literature provides ample examples of gender differences in loss aversion. Schmidt and Traub, 2002 found that females are more likely to choose loss-framed lotteries com-

[10] As a robustness check, we ran these regressions without control variables in Appendix 2 and find that the difference between the two treatments is even smaller – and once we control for semester performance, the effect is essentially zero.

pared to males. Females also tend to be more loss-averse in financial investments (Arora and Kumari, 2015; Hassan et al., 2014). However, recent research shows that the gender difference in loss aversion may depend on how loss aversion is defined (Bouchouicha et al., 2019).

Table 1.6: Testing gender-heterogeneity of losing points

	Test 1	Test 2	Test 3	Test 4	Best 3 Tests	Final Test	
Losing Points	1.169	11.11***	6.230	9.029**	8.150**	10.28**	9.102***
	(4.335)	(2.956)	(5.134)	(3.292)	(3.317)	(4.134)	(2.645)
Female	7.591**	9.052**	8.755	5.550	8.338*	5.540	0.726
	(2.545)	(2.921)	(5.544)	(5.704)	(4.181)	(7.104)	(3.702)
Female x Losing Points	-5.262	-8.182	-3.416	-2.672	-5.960	-4.782	-2.590
	(2.949)	(5.113)	(7.442)	(5.462)	(4.901)	(6.971)	(3.763)
Best 3 Tests							0.556***
							(0.115)
Constant	62.77***	65.39***	64.02***	59.29***	68.67***	57.06***	16.19*
	(3.845)	(3.322)	(7.086)	(6.270)	(4.162)	(7.208)	(8.028)
Observations	321	321	321	321	321	321	321
R-squared	0.157	0.178	0.125	0.094	0.166	0.162	0.352
Residual Std. Error	23.085	24.047	26.404	31.457	20.002	23.499	20.698

Standard errors in parentheses, clustered by seminar group;* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Control variables include: mother's university education, knowing how to take derivatives,working or not, and teacher, practice session and campus fixed effects.

Table 1.6 presents the regression results for gender heterogeneity. We observe that across all tests, the interaction term between being female and losing points is negative but not statistically significant. When tested together with the Female variable, the interaction term slightly reduces the effect. Overall, gender differences appear more significant throughout the semester, which may be due to the first four tests being open-ended, as suggested by Cole, 1997. For the Final Exam, neither the gender variable nor the interaction term is statistically significant, and they cancel each other out.

As a robustness check, we compare the control group to students who were losing points throughout the entire semester. Table 1.7 shows these results, where neither the gender dummy nor the interaction term is statistically significant, and they work in opposite directions. Thus, similar to McEvoy, 2016, we find no evidence of heterogeneous treatment effects based on gender.

There may also be gender heterogeneity related to dropouts, as we observed that females performed better throughout the semester. Since males lost more points in the treatments, they might have been more discouraged and prone to dropping out. We explore this

Table 1.7: Testing gender-heterogeneity of losing points

	Test 1	Test 2	Test 3	Test 4	Best 3 Tests	Final Test	
Loss	-0.759	10.37***	1.802	7.154	5.751	11.52**	7.636**
	(3.517)	(2.712)	(5.986)	(4.515)	(4.048)	(3.869)	(2.695)
Female	7.703	6.946	5.886	8.396	7.978	5.298	-0.0913
	(4.892)	(4.223)	(9.258)	(8.686)	(7.043)	(7.117)	(3.074)
Female x Loss	-5.894	-6.019	-0.968	-5.730	-5.937	-4.841	-0.831
	(5.295)	(5.372)	(10.73)	(8.694)	(7.681)	(7.294)	(4.188)
Best 3 Tests							0.676***
							(0.139)
Constant	65.67***	63.26***	72.23***	61.36***	71.55***	54.46***	6.127
	(3.782)	(4.655)	(7.587)	(7.322)	(5.459)	(6.999)	(10.16)
Observations	236	236	236	236	236	236	236
R-squared	0.211	0.171	0.126	0.101	0.168	0.201	0.451
Residual Std. Error	21.765	23.980	25.612	30.537	19.138	23.094	19.179

Standard errors in parentheses, clustered by seminar group;* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Control variables include: mother's university education, knowing how to take derivatives, working or not, and teacher, practice-session and campus fixed effects.

further in Section 3 of the Appendix. Our findings indicate that gender is not a significant predictor of dropout behavior. Additionally, omitting dropouts from our sample does not affect the results, confirming that the treatment effect is gender-neutral.

Heterogeneity by Student Performance

To examine whether there are heterogeneous treatment effects based on students' prior performance, we first conduct a quantile regression on the Final Exam scores. This allows us to assess whether the treatment effect varies across different performance levels. Specifically, we test whether the average treatment effect is positive across the entire distribution. To identify unconditional quantile treatment effects, we use the Residualized Quantile Regression (RQR) framework (Borgen et al., 2024).

Figure 1.4 presents the estimated treatment effects by quantiles. There are two key takeaways. First, we observe statistically significant positive effects between the 30th and 65th-70th percentiles (the regression results are included in Section 5 of the Appendix). The lack of significance in other quantiles is likely due to low statistical power, as our sample only includes 321 observations. Second, and more importantly, the treatment effect is positive across most quantiles, with the exception of the lower quantiles, which is likely explained by the presence of dropouts in these groups. This suggests that, although the size of the effect may vary, none of the students are worse off compared to the control group.

Additionally, we analyze students by stratifying them based on whether they learned how to take derivatives in secondary school, which serves as a proxy for their mathematical ability. [11] We then run the previously specified regressions separately for these groups. Table 1.8 presents the results: for students who did not learn how to take derivatives in secondary school, the treatment effect is larger (around 15%), while for those who did, the effect is smaller (around 4.5%) and becomes statistically significant only after accounting for their performance during the semester. It is important to note that the effect is non-negative for both subgroups, meaning that, in absolute terms, no one is worse off. [12]

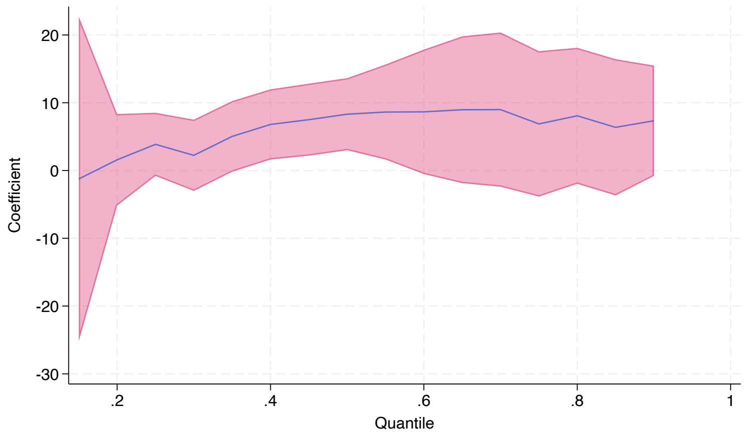

Figure 1.4: Quantile Treatment Effects for Losing Points at the Final Test

While we do not find any evidence of heterogeneity in students' performance, the lack of data limits the conclusiveness of these findings. Specifically, we do not have access to students' GPAs, which would have served as an important robustness check for our results. It is also important to note that we analyze the effects on scores in absolute terms, as grading on the curve is not commonly used in Hungary. Therefore, treatment heterogeneity becomes significant when there is a sign difference in the effect. The implications for

[11] In the Hungarian school system, secondary school students only learn how to take derivatives if they take advanced math classes.

[12] Another way to stratify students is by dividing them into the top and bottom 50% based on their scores. While we observe a positive and statistically significant effect for both groups, endogeneity might be a concern in this case.

Table 1.8: Regression results for Final Test scores - regressions for proxy of student's math skills

	Final Test Scores			
	No Derivatives		Knows Derivatives	
Losing Points	15.20**	14.94***	4.519	4.369**
	(5.421)	(3.885)	(3.162)	(1.746)
3 Best Tests		0.482**		0.645***
		(0.180)		(0.164)
Constant	46.91***	8.709	67.46***	19.67
	(8.240)	(11.93)	(7.306)	(12.05)
Observations	130	130	191	191
R-squared	0.272	0.418	0.126	0.365
Residual Std. Error	23.919	21.481	23.070	19.723

Standard errors in parentheses, clustered by seminar group;* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Control variables include: mother's university education, knowing how to take derivatives, working or not and teacher, practice-session and campus fixed effects.

grading on the curve, however, are different: in such cases, unequal effect sizes should be considered.

1.4 Conclusions

In this paper, we have examined the effect of loss-framing on student performance. The literature on loss aversion in educational settings has shown mixed results, with some studies suggesting a negative impact on learning outcomes, while others have indicated positive effects. Additionally, we have investigated whether differences in student performance could be attributed to a novelty effect, where students encounter the concept of losing points for the first time. If this has been the case, we would expect the effect of loss-framing to diminish over the course of the semester. To our knowledge, this has been the first study to test this hypothesis.

Our findings have shown that loss-framing positively affects student outcomes: students who have been losing points have earned, on average, 7 to 9 percent more points on the Final Exam. The effect has also been evident throughout the semester, with students who have lost points earning between 2.5 to 5 percent more points compared to their peers. The results remain statistically significant even after controlling for performance throughout the semester and student dropouts. Given the low cost of implementing this grading feedback, we argue that the size of the effect is rather substantial.

We have found that throughout the semester, students who have lost points performed significantly better than those who have only lost points during the Final Exam. However, the difference has diminished and become statistically insignificant on the Final Exam. After controlling for semester performance, the gap has narrowed further, suggesting that the differences are not due to the novelty effect. This indicates that the treatment has not lost its effectiveness over the semester. Analysing whether these effects persist in the long run should also be explored in future research.

We have also examined heterogeneous treatment effects to ensure the intervention has led to a Pareto improvement, benefiting all students equally. First, we have investigated potential gender differences in response to the treatment. Previous literature has presented mixed evidence on gender heterogeneity in loss-framing, but we have found no statistically significant gender-specific treatment effects. Next, we have tested whether the treatment has impacted high- and low-performing students differently. We have stratified students based on whether they have taken advanced math classes in secondary school, using this as a proxy for mathematical ability. While the effect size has been smaller for students who have taken advanced math classes, it has remained positive for both groups. Additionally, quantile regression analyses have shown that the treatment effect has increased with performance percentiles.

By showing that loss-framed grading can produce sustained improvements in student performance without negatively affecting specific subgroups, our findings suggest that this approach could be a scalable and cost-effective intervention for enhancing educational achievement. Ultimately, our research highlights the potential of loss aversion as a tool for motivating academic performance, suggesting opportunities for its consideration in educational policy and practice.

This study does have limitations. It was conducted at one of the top business schools in Hungary. Previous research has suggested that student motivation plays a critical role in the effectiveness of interventions similar to ours (as in the case of Czibor et al., 2020), and in our case, students were highly motivated – only two out of 321 students have indicated they were merely aiming to pass. Additionally, financial incentives, such as scholarships for good grades, likely played a role. We also acknowledge that the implications of loss-framed grading might differ when grading on the curve. Additionally, we do not consider broader implications of loss-framing, for example, psychological impact on students due to the increased emphasis on ”losing”. Further research should also explore these implications.

One assumption in our approach has been that students feel endowed with points when they are losing them. The effect could have potentially been enhanced if students had first earned their points and then lost them due to incorrect answers. Future research could explore how task complexity and the type of test (e.g., multiple-choice vs. open-ended) influence the effectiveness of loss-framing. Finally, while our study has found that the effect of losing points persisted throughout the semester, further research should test this effect over longer periods.

Chapter 2

Fair and Unfair Differences in Individual Decision-making[1]

2.1 Introduction

Although mainstream economics models assume homo economicus' rational and self-interested behaviour, the development of behavioural economics in recent decades has brought to light a number of other factors that are important in individual decision-making. These include a number of institutional, decision environmental and psychological factors. For example, whether an individual perceives the transaction as fair or perceives the difference as fair relative to the observation on which the comparison is based plays an important role in the outcome of transactions. It is therefore important to examine and understand the rationality (or irrationality) behind fair or equitable behaviour.

When making decisions, we are constantly comparing different alternatives: in the shop we choose products of different price and quality, in the job market we choose jobs with different payments, different responsibilities and different skills. But often we make decisions not rationally, but interpersonally - comparing our situation with that of another person - and from this comparison can come the development of malicious envy.

According to Leon Festinger's theory of interpersonal comparison, people like to compare their own opinions, abilities, talents and general situation with those of others (Festinger, 1954). This is because there is a general motivation in everyone to get an accurate picture of themselves, which they try to achieve by comparison. The importance of

[1]This chapter was originally published as ”Méltányos és méltánytalan különbségek az egyéni döntéshozatalban. *Közgazdasági Szemle*, 69(10), 1170-1194.”

comparison and fairness in decision making and welfare state policy has been discussed by Rawls, 1971 and Harsanyi, 1955, 1975a, 1975b. While John Harsanyi examined the question of justice in terms of a fundamentally utilitarian equality, John Rawls emphasized much more the issue of equity itself. The relationship between comparative equity and envy was also explored in Varian, 1973; his analysis showed that in the eyes of people, an equitable distribution of resources precludes the development of envy. The integration of fairness into decision theory has become important since the 2000s (Bolton and Ockenfels, 2000; Fehr and Schmidt, 2000), but little has been said in economics about the negative effects of envy – the harms of malicious envy. Malicious envy refers to the behaviour whereby, as a result of envy, we damage the payoff of another economic agent in order to reduce the payoff gap between us (Bedeian, 1995; Cohen-Charash, 2009).

Several previous studies have looked at the impact of fairness on decision-making, but envy as a cause has been identified to a limited extent. While Güth et al., 1982 declared the importance of fairness in game-theoretic payoffs, they did not address the nature of payments and their fairness-implications. It is also important to clarify what information was available to decision makers when they declared a distribution to be unfair. Kahneman et al., 1986a, 1986b investigated the reason for identifying something as fair, while other psychological studies have examined the relationship between fairness and envy and counterproductive workplace behaviour (Cohen-Charash and Mueller, 2007; Khan et al., 2009; J. Kim and Park, 2018; Marescaux et al., 2021). However, these studies did not measure the effect unfairness has on payments, or how much of a pay difference decision makers are willing to accept as a fair in the workplace. Furthermore, the studies on envy have not addressed how and in what direction new information about the nature of inequality affects perceptions of fairness. For example, without any information, it may be easy for one actor to perceive a specific wage distribution as unfair, and to reduce the other person's pay out out of envy, but by shedding light on the reason for the discrepancy, rationalising it may eliminate the harmful behaviour.

This paper investigates the impact of envy through comparison and fairness judgments using decision-theoretic tools. As we will see below, envy as a decision factor has been studied only marginally from a decision-theoretic perspective, but as a harmful behaviour it can significantly influence individual payoffs. Another important question is how counterproductive behaviours can be counteracted, or what information can be used to counteract them, and how this new information can be incorporated into the decision mechanism. We use a questionnaire approach to explore the importance of fairness in counterproductive behaviours and how some new information about inequality affects the perception of fairness. We show that a non-negligible proportion of people would reduce perceived inequalities and that respondents are more likely to appreciate physically visible effort rather than previously acquired knowledge or experience. In the first part of the paper, we summarise the relevant literature on fairness and envy and reference point decision theory, and after formulating the hypotheses, we turn to the questionnaire data on which the analysis is based. Finally, our hypotheses on the effect of fairness are analysed using statistical methods.

2.2 Literature Review

Traditionally, two dominant models have been used in economic thinking to analyse decisions under uncertainty: the expected utility model (Von Neumann and Morgenstern, 2007), associated with János Neumann and Oskar Morgenstern, and the prospect theory model defined by Amos Tversky and Daniel Kahneman (Kahneman and Tversky, 1979). Varian, 2014 argues that the theory of expected utility is realistically simplistic, since in weighing our options we can indeed choose only one option, and this choice depends on our individual risk preferences, but Varian emphasises that our choices should not depend on a preference for another, imagined state of nature (Varian, 2014,pp. 229-248).

Simon, 1957, 1966 criticised the theory of expected utility and the profit maximisation axiom itself from the firms' point of view. Simon argues that, on the one hand, the rationality of the decision-maker is not trivial, and, on the other hand, given his payoffs, it is quite possible that the individual decision-maker is not maximizing profit but simply seeking to obtain an income that satisfies him. In analysing satisficing behaviour, Simon points out that the decision always revolves around some goals or values, or a perception of the environment and particular values and facts – so that it can be easily influenced by an external reference point, for example.

The alternative to expected utility is the prospect theory presented by Kahneman and Tversky, 1979, which uses the as the basis for utility. Figure 2.1 illustrates the value function: the horizontal axis indicates positive and negative deviations from the reference point, while the vertical axis indicates increases and decreases in utility from the reference point.

The methodology can be used to measure preferences for different perspectives, but the actual scaling is more difficult due to the decision weights defined by individual risk perception (Kahneman and Tversky, 1979), but not only the relativity of the environment can be used to evaluate our decisions. For example, Kőszegi and Rabin, 2006 reference-based

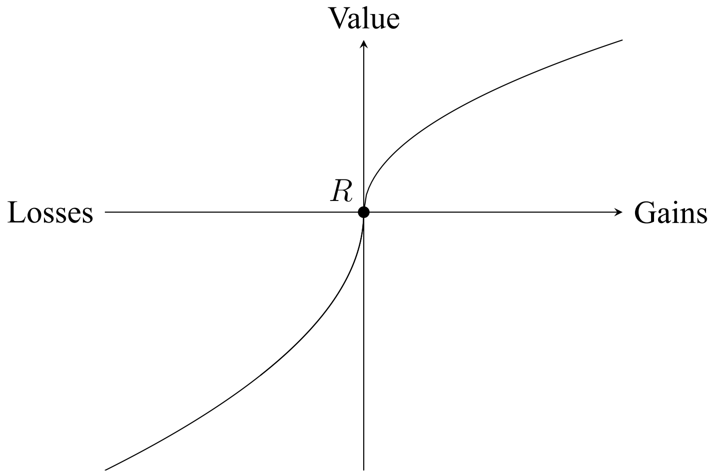

Figure 2.1: Prospect Theory based on Kahneman and Tversky, 1979

utility function started from an endogenous reference point derived from expectation, i.e., she assumed that the decision maker makes his/her decision based on his/her own past payoffs.

When examining the issue of fairness, the models of Akerlof, 1978, 1980, 1982 and Okun, 1981 have an important difference from standard models: firms do not simply seek to maximise profits, but to maximise profits in the long run. Okun found that unfair behaviour can cause disturbances in consumer markets, which can lead to consumers being unable to consume their full (potential) consumption. Looking at pricing issues, he concluded that consumers react hostile to price increases that are not justified by some cost increases and are therefore perceived as unfair. At the same time, they tend to accept fair price increases when demand is stagnating. In sum, an unfair supply price may create a situation in which consumers are willing to seek alternatives to the firm where they perceive unfairness, even at a cost (thereby seeking a form of revenge against the firm).

Fair behaviour was first investigated experimentally Güth et al., 1982 and Binmore et al., 1985, who defined the one- and two-round ultimatum games. In contrast to standard game-theoretic models, in this experiment the players strove for an approximately equal, i.e., fair distribution, and unfair offers were generally rejected by the second player.

Kahneman et al., 1986b focused on the nature and economic effects of unfair behaviour. The authors' research, conducted in the form of a questionnaire, focused on two main questions.

- What is society's perception of the fairness of the prices, wages and rents set by firms, and what impact might this perception have on pricing?
- How does fairness affect the outcome of transactions?

In their study, the authors used reference-transactions to model the impact of reference points on decisions. In evaluating the results, they analysed the extent to which the transaction deviated from the 'norm' and the influence and impact of the action itself on the parties involved. Kahneman et al., 1986b concluded from the reference transactions that the deviation between transactions that consumers perceived as fair and transactions that they expected to be fair was small, i.e., deviations from the reference point were considered unfair for individual decision making. The authors did not, however, address intra-firm differences, interpersonal reference points – i.e., comparisons between individuals. The present study examines a subset of these interpersonal differences, the perception of intra-firm differences and the perception of fairness differences between employees

Fairness is essential in our discussion because it can be understood as the opposite of envy from a decision-theoretic point of view. Varian, 1973defined fairness and envy as mutually exclusive phenomenae. According to his definition, if the allocation of goods is such that no individual prefers another's basket to his own and the allocation is Pareto-efficient, then the allocation is equitable. In this definition, each economic agent puts herself in the place of the other individual in her valuation and evaluates her own payoff accordingly. Varian notes that this definition is very model-based, since it is only the physical quantity of goods that the agents are comparing. Similarly, the exclusion of envy by equity is also a model-like but efficient approximation, and so in this paper we use this approach to explain the elimination of malicious envy. Furthermore, Varian pointed out that potentially other factors may be important in the comparison: for example, the strength of individuals' preferences, the importance of morality for decision-makers, and the size of contributions and investments of effort and their evolution over time.

2.3 Hypotheses and Empirical Background

In the following, we present our hypotheses on the impact of fairness in decision theory. As already stated in prospect theory, the reference point, or deviation from it, has a significant function in decision making. Also important is the Aristotelian idea that in interpersonal comparisons we compare ourselves to persons close to us (and although we may compare ourselves to persons much more powerful than ourselves, this comparison is not reflected in our decisions). According to one possible definition:

"(...)envy occurs when a person lacks another's superior quality, achievement, or possession and either desires it or wishes that the other lacked it" (Parrott and Smith, 1993, pp. 906).

Based on this, envy appears from the decision-maker's point of view in the negative, loss-indicating domain of the reference point. There are basically three scenarios in decision-making:

1. the decision-maker is envious of the reference person but does not do him/her any harm,

2. the decision maker is envious of the reference person, wants to harm him, but cannot create a change in the ultimate utility of the reference person, has no control over it, or

3. the decision-maker is envious of the reference person, wishes to harm him and ultimately does so (malicious envy).

In each of these cases, the decision maker tries to reduce the distance between the reference person and their payoff by reducing the reference point to avoid losses. In the first two cases, the decision maker may choose an alternative reference point (i.e., compare himself to another person), but in the third case, the decision maker has the possibility to negatively influence the payoff of the reference person, i.e., to show "malicious envy". It is important to note, however, that several things can influence whether a decision-maker uses malicious envy. These include, for example, the personality (Bedeian, 1995) and emotional state (Loewenstein, 2000) of the decision-maker, the information available, the relationship with the reference person, and the magnitude of the deviation itself.

Based on the reference point decision theory, we formulate the following five hypotheses on the consequences of comparison and envy.

Hypothesis 1: In general, we consider people at nearly the same level as us as reference points.

Hypothesis 1 examines the identity of the reference person, the identification of the reference point. As discussed earlier, this hypothesis has been observed in antiquity, but we would like to confirm it by empirical investigation. The existence and purpose of inter-personal comparisons, as well as the existence of malicious envy as formulated in Hypothesis 4, is supported by the experiment of Hoffman et al., 1954. In the experiment, three players played a special performance-based game in which one player, pre-designed, performed significantly better than his other two peers. In the case where they had no opportunity to cooperate, the two lower-scoring players in the experiment competed with each other rather than with the much higher-scoring player. However, as soon as they had the opportunity to cooperate, the weaker players teamed up against the stronger player, thus reducing the score of the player with the advantage. The result of this experiment provides support for the hypothesis that if a person is envious and able to reduce inequality, he will take advantage of this opportunity. As for the object of comparison itself, Festinger, 1954 argues that, consistent with earlier assumptions, we tend to compare ourselves to and essentially compete with persons closer to us.

Hypothesis 2: When comparing, (especially in the case of envy), we do not take into account the effort or energy of the person, we only form envy based on the end result.

Hypothesis 3: Following from Hypothesis 2, envy may disappear if, when comparing with the reference person, we draw attention to differences in effort.

Hypotheses 2 and 3 highlight the importance of available information. According to Hypothesis 2, when developing malicious envy, we do not take into account the effort of the reference person, partly because we have no information about it, only on their pay (for example, we do not take into account how much more experience the other person's job requires, we only perceive the pay difference). If we have no information about what causes the pay difference between the reference person and the decision-maker, we are more likely to judge the difference as unfair, and thus the likelihood of malicious envy may increase. However, if new information is brought to light about the reason for the difference (e.g., the person has been working longer, possibly has higher education and expertise), the decision-maker can rationalise the difference, so that even if his/her feelings of envy do not disappear, malicious envy will not occur because he/she will already consider the perceived difference to be fair

Hypothesis 4: Malicious envy exists, i.e., persistent negative deviation from the reference point results in harm to the reference person, which reduces the distance from the reference point for the decision maker doing the harm.

The hypothesis follows from the basic idea that with a change in the reference point the preferences of the decision maker may change. Kahneman and Tversky, 1979 notes that a negative evaluation of a choice problem may in some cases increase risk-seeking behaviour. An example is when a consumer has failed to adapt to a new situation that is not the result of old losses. A person who has not come to terms with his losses is more likely to accept options that he would not otherwise accept. For example, it is possible to imagine a morally unacceptable action, such as harming the reference person, which could be defined as malicious envy.

Lin and Bates, 2021 found a positive correlation between envy and redistribution support for reallocation, an effect that was further strengthened when malicious envy was in the regression. Thus, it is possible that malicious envy may lead decision makers to support redistribution.

Kahneman et al., 1986a conducted a two-phase experiment: in the first phase, players played an anonymous version of the dictator game. The anonymity, and the fact that the second player could not punish the first player lead to 76% of the participants in the experiment chosing equal distribution. In the second half of the experiment, the second player had the choice of choosing \$1 (i.e., at his own expense) to reward a fair dealer and punish an unfair one. 74% of the participants exercised this option. The unfair punishment of a dictator can be understood as malicious envy, which is completely irrational from a traditional economic point of view: on the one hand, the person has to pay, thus reducing his own utility, and on the other hand, he should not take into account the utility of other persons utility of others. If, however, it is understood that it is not absolute utility that is looked at but are trying to minimize the difference in payments, then the decision can be seen as legitimate and justified.

Hypothesis 5: Perceptions of fairness and consequent malicious envy may differ in different societies.

According to hypothesis 5, the perception of fairness is not universal and may depend on other factors, such as cultural ones. Hundley and Kim, 1997 investigated the determinants of perceptions of fairness of wages and wage differentials among American and Korean workers. The authors wanted to find out whether there are factors that influence the perception of fairness in wages and whether there are differences in the importance of each factor across countries due to cultural differences. They used regression methods to analyse worker characteristics such as family and educational background and effort at work. The results showed that Korean subjects' perceptions of the fair wage gap were significantly influenced by the size of the reference person's family, educational difference, and age difference. In contrast, for Americans, individual work performance and the amount of effort devoted to the job were more influential. A similar result was found by Kim and Leung, 2007, who discovered a link between the materiality of cultures and fair distribution: interpersonal fairness was more influential in the perception of intra-firm fairness for American and Japanese workers than for Korean or Chinese workers. These findings suggest that social and cultural influences may be important in the perception of the fairness gap.

2.4 Data

The empirical research was conducted using a questionnaire method. The questionnaires were distributed online in two languages: one in Hungarian and one in English via university channels abroad (see Appendix for the questionnaire). The questionnaire questions were pre-tested with several native speakers of Hungarian and English before finalisation, and we made sure that the questions were clear, neutral and that the translation did not affect the interpretation or objectivity of the question. There are several arguments in favour of the questionnaire method: firstly, it allows more people to be interviewed, thus making the sample more representative. On the other hand, since we are comparing the attitudes of Hungarians and foreigners, it is easier to carry out a questionnaire survey than to conduct experiments in different locations due to a lack of resources. Further experiments are worthwhile to investigate the perception of envy and the respondents' choice of reference persons (e.g., Hoffman et al., 1954, Güth et al., 1982).

Respondents were asked to either provide specific figures or rate the likelihood of a decision on a Likert scale of 1 to 7. After providing the standard socio-demographic data (gender, age, nationality, place of residence, education, work experience), respondents were asked to answer questions that tried to encourage them to think in terms of reference points or reference persons. For example, we asked them to place themselves in society in terms of livelihood and wages, and asked them who or what type of person they liked to compare themselves to – while also allowing us to test Hypothesis 1.

The sample consists of 306 observations, 246 Hungarian respondents and 70 foreign respondents. In terms of gender distribution, 200 women (65.3 percent) and 106 men (34.7 percent) completed the questionnaire. Figure 2.2 shows the distribution of respondents by place of residence. Almost half of the respondents, around 48%, live in the capital city and 19% in a county town. Only 22% of the respondents live in a smaller town or village, so the sample is skewed towards individuals living in bigger cities.

Figure 3 shows the age distribution of respondents, with the youngest respondent aged 18 and the oldest 62. The figure clearly shows that those under 30 are over-represented in the sample, while middle-aged respondents are under-represented. The sample has a right skewed distribution with an average age of 30.72 years and a median age of 27 years.

The sample is under-represented by lower educational attainment: only 20 percent of

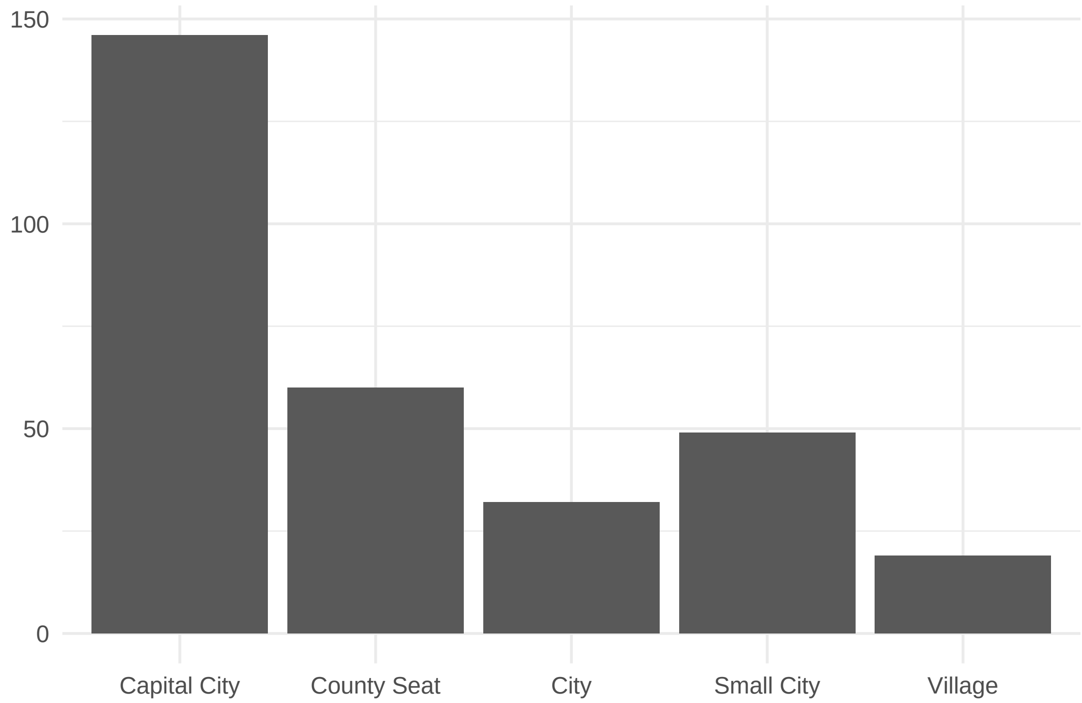

Figure 2.2: Distribution of the sample based on place of living

respondents have a secondary or lower education attainment level. For work experience, only 3 percent of respondents do not have and only 15 percent have less than one year of work experience. This is important in the sense that most respondents can draw on their own experience to judge fairly the pay gap, especially if they have experienced a perceived pay gap and possibly pay tension in their workplace.

All this said, the sample is not representative of the whole (Hungarian) society, however, significant conclusions can be drawn, especially for inequality attitudes of younger employees with university degrees in the business sector.

2.5 Results

The evaluation of the questionnaires are presented below. First, the identity of the reference persons is examined, followed by an analysis of the perception of fairness and the factors that influence fairness. We then analyse the probability of reducing the difference between the reference person and the decision-maker, i.e., the existence of malicious envy

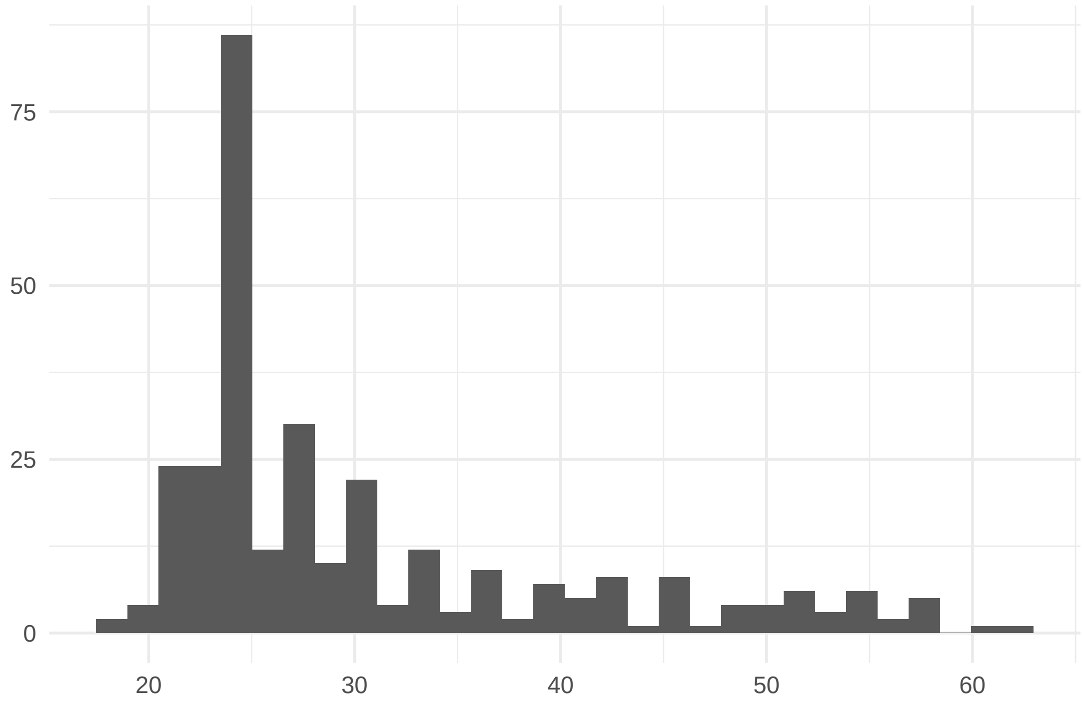

Figure 2.3: Distribution of the sample based on age

in the present reading. Finally, we examine how different types of information about the reason for the difference affect the perception of the fairness of the difference.

Figure 2.4 shows the choice of reference point for decision-makers according to the direction of deviations. 72 percent of respondents gave answers that suggest that, in general, the reference point and the decision-maker's payment are close to each other. 6 per cent gave an answer that they prefer to deviate from the reference point in a strictly positive way: this can be understood as a prospect-theory loss aversion motive. The remaining 23 per cent prefer to compare themselves to people with better living conditions. It is worth noting, however, that this is much higher among foreign respondents: 47.2 per cent, which is the same as the rate for comparing themselves with people in the same circumstances. However, a higher proportion of foreign respondents mentioned persons in a higher position than themselves (immediate superior, senior manager, famous person) when identifying the reference point. Thus it is possible that motivational factors may also play a role in comparisons with better living conditions.

The next question asked for the identity of the reference person: when answering, respondents were given the option to name more than one reference person. As can be seen

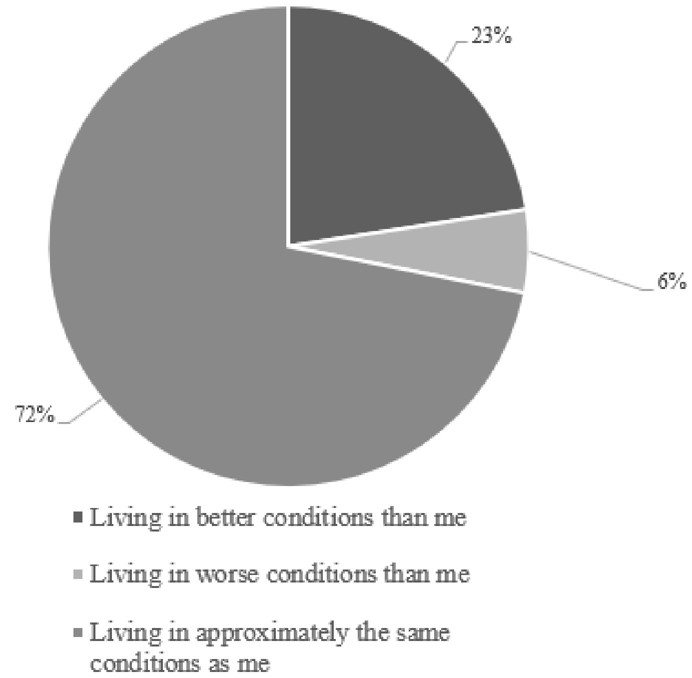

Figure 2.4: Direction of comparison in the sample

in Figure 5, respondents mainly indicated persons with whom they have a close relationship and meet frequently. The top four most common reference persons were a co-worker (73 per cent), a close friend (52 per cent), a friend (44 per cent) and a close relative (37 per cent). These are mainly persons with whom we have daily contact. Comparisons with famous people or senior executives or distant family members show significantly low scores, which also supports the argument that we tend to compare ourselves with those close to us (socially and economically). 23 per cent of respondents tend to compare their salary to that of their direct manager, which may also provide a motivational incentive to perform better.

In conclusion, we cannot reject hypothesis 1, i.e., no answer refuted that we compare ourselves to people close to us. Therefore, it is rare that the perceived glaringly high pay gap plays a role in our decisions. The result is consistent with the Aristotelian findings mentioned earlier, as well as with previous psychological experiments. For example, in the psychological experiment of Major and Forcey, 1985, subjects also sought to compare their own pay with similar pay, both in terms of gender and the nature of the work performed.

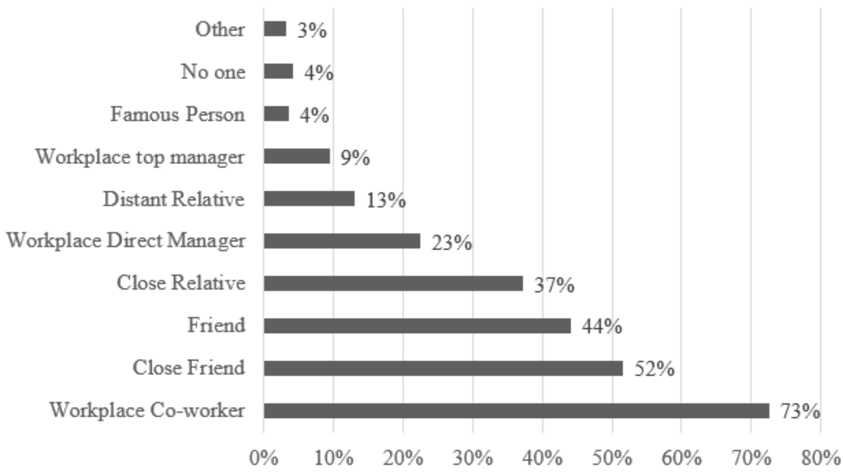

Figure 2.5: Proportion of reference person types in the sample (percentage). Note: multiple reference persons could be named in this question

2.5.1 Boundaries of Fairness

After the introductory questions, we asked respondents to indicate how much of a difference they considered fair between themselves and their co-worker, themselves and their direct manager, and themselves and the top manager of the company, in a situation where their salary was equal to the average salary in society (200,000 HUF or 1,000 USD, depending on the questionnaire). Here, we wanted to investigate both how much difference they allow between themselves and their co-worker (if any) and how much more of a difference in pay they allow between themselves and their supervisor, and how much of a spread this difference has in the sample. Here we have assumed that if there is a large pay differential allowed, respondents do not compare themselves to that person. The responses may also reveal how sensitive a society is to inequalities and how egalitarian it views itself. These can be important for welfare economics (e.g., Hundley and Kim, 1997, Alesina and Giuliano, 2011).

In the questionnaire data, answers where there was a clear local value deletion were corrected (for example, if their direct manager was allowed a maximum wage premium of HUF 500,000 while their colleague was allowed a maximum of HUF 1 million; in this case, I corrected the latter by one decimal place). Observations where the respondents allowed more than five times their own salary to their colleague and more than 25 times their own salary to their top manager were considered as outliers. Thus, the analysis was finally carried out on a sample of 292 items.

The basic statistics for the percentage differences are shown in Table 2.1. The distribution of the differences does not show a normal distribution in either case, which can be attributed to two reasons. On the one hand, the average for co-workers and immediate managers is very close to the median, so that the majority of observations are concentrated in this area. On the other hand, the non-normal distribution is also due to the fact that people like to give whole, round numbers when they have to give continuous numbers as answers, especially when asked about monetary values (so, for example, HUF 50 000 might be a sufficiently common 'rule of thumb' answer). Two extreme cases occurred when examining differences that are considered fair: there are employees who (by their own admission) are willing to tolerate any wage difference and those who want a completely equal salary. The former were treated as outliers and were excluded from the analysis. It is clear, however, that the tolerance of pay differentials varies between different levels of seniority: while the vast majority of responses for co-workers at the same level are between 0 and 100 per cent, the vast majority of responses for those in the same level of seniority show greater flexibility in dealing with pay differentials; this can be inferred from the variance. This is also the perfectly logical explanation that people will be able to articulate much more clearly that what is a fair difference at their own level rather than at the level of a very distant, leading person in a very distant position. This is also due to the abstract nature of the question, after all it is hard to imagine what kind of work a billionaire gets paid for, and what kind of effort.

Table 2.1: Descriptive Statistics

	N	*Mean*	*Std. Dev.*	*Min.*	*25th Pctl.*	*75th Pctl.*	*Max.*
Co-worker	292	0.409	0.71	0	0.1	0.5	5
Direct Manager	292	1.23	1.711	0	0.35	1.25	10
Top Manager	292	3.95	5.523	0	0.8	4.0	25

In the following, we analyse the fairness gap between co-workers and the respondents, taking into account the available socio-demographic data. As can be seen from the descriptive statistics, the fairness gap in percentage form is significantly skewed to the right, with the majority of cases between zero and 100 percent. In order to analyse the direction of change using linear regression, we chose to normalise the distribution of the dependent variable. The results of the linear regression are shown in Table 2.2: the model includes dummy variables for gender (reference group: male), being a Hungarian citizen (reference group: foreign citizen), of the type of municipality of residence of the respondent (reference group: capital city residence), and work experience (reference group: less than 1 year of work experience); finally, we controlled for respondents who had only a high school diploma or less.

Table 2.2: Regression Results

Variable	Coefficient
Gender (ref.: male)	
Female	-0.325^{***}
	(0.110)
Nationality (ref.: foreign)	
Hungarian	-0.350^{*}
	(0.179)
Residence (ref.: capital city)	
Village	0.146
	(0.221)
City	0.038
	(0.056)
Work Experience (ref.: $<$ 1 year)	
1–3 years	-0.365^{**}
	(0.160)
3–5 years	-0.367
	(0.234)
More than 5 years	-0.461^{***}
	(0.152)
Education (ref.: higher education)	
High school diploma or lower	-0.277^{**}
	(0.128)
Constant	0.806^{***}
	(0.159)
Observations	292
R^2	0.156
Adjusted R^2	0.126
Residual standard error	0.889 (df = 283)
F-statistic	6.711^{***} (df = 8; 283)

Note: Standard errors in parentheses.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

In the regression, the constant was found to be significant, i.e., the average respondent still considered a 20% lower salary to be fair. The model did not show a significant effect of place of residence. However, there is a significant difference between the sexes: women tolerated a substantially smaller difference in the survey. There was also a significant difference between Hungarian and foreign respondents: Hungarians tolerated less variation. Based on the latter result, we cannot reject Hypothesis 5 about the perception of a fair difference between cultures; further research may be worthwhile to analyse the deeper reasons for this. This result is consistent with Hundley and Kim, 1997 and Kim and Leung, 2007, where significant differences in perceptions of fairness across countries. Furthermore, a significant effect was found for education level: those with high school diploma or lower education tolerated a smaller difference than those with higher education.

When looking at those with less than one year of work experience, there was also a significant negative difference between respondents with between one and three and more than five years of work experience, but those with between three and five years of work experience did not show a significant difference compared to the reference group. Thus, those with more work experience tolerated a smaller wage gap compared to the other groups.

The R-squared value of the regression is relatively low (around 16 percent), due to the small sample size and the unobserved variables such as income, cognitive and non-cognitive abilities of decision-makers, ideological views (e.g., conservative, liberal, etc.). These questions were not included in the questionnaire due to the heterogeneity of the sample (i.e., foreign and Hungarian respondents). Other variables not observed are the personality traits of individuals. This issue may also be worth including in statistical analysis in further research.

Following the questions on fairness, a distribution problem was presented to the respondents. Assuming that they perceive a significant pay gap between themselves and their colleagues, how likely would they be to distribute this pay gap between themselves and their colleagues? In essence, the question asked whether there was malicious envy: the decision would imply that one's own utility would increase while the utility of the reference person would decrease, thus setting up what Simon, 1966 calls a sustainable equilibrium. The consequences of the decision are illustrated in Figure 2.6 using the prospect theory function, where R is the original reference point, E is the decision maker's original payoff, and R* is the new reference point with which the decision maker's new payoff is equal. The question assumption is framed in such a way that it rationalizes the redistribution and also exonerates the decision maker from an ”emotional side” by doing good to his/her co-workers, creating a much more equal working environment by redistributing. On the one hand, this can be equated to rationalising malicious envy, but on the other hand, it can also be understood as a non-cooperative dictator game, where a rational decision-maker will always take advantage of the redistribution option, since it a) increases his own payoff and b) reduces the loss of utility due to cognitive dissonance caused by the difference

The distribution of responses to the question on equal distribution is shown in Figure 2.7. Slightly less than half of respondents, 48.6 percent, would have preferred to take up the redistribution option; the highest proportion of respondents would not have taken up the redistribution option at all, more than twice as many as would have definitely taken up such an option.

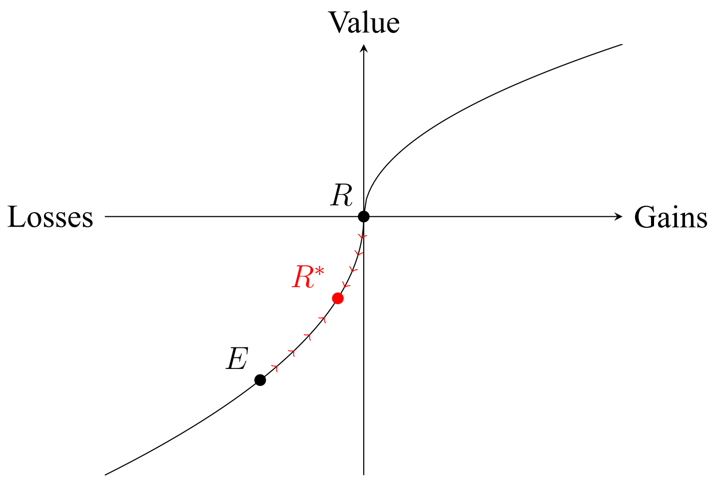

Figure 2.6: Illustration of the effect of redistribution with Prospect Theory

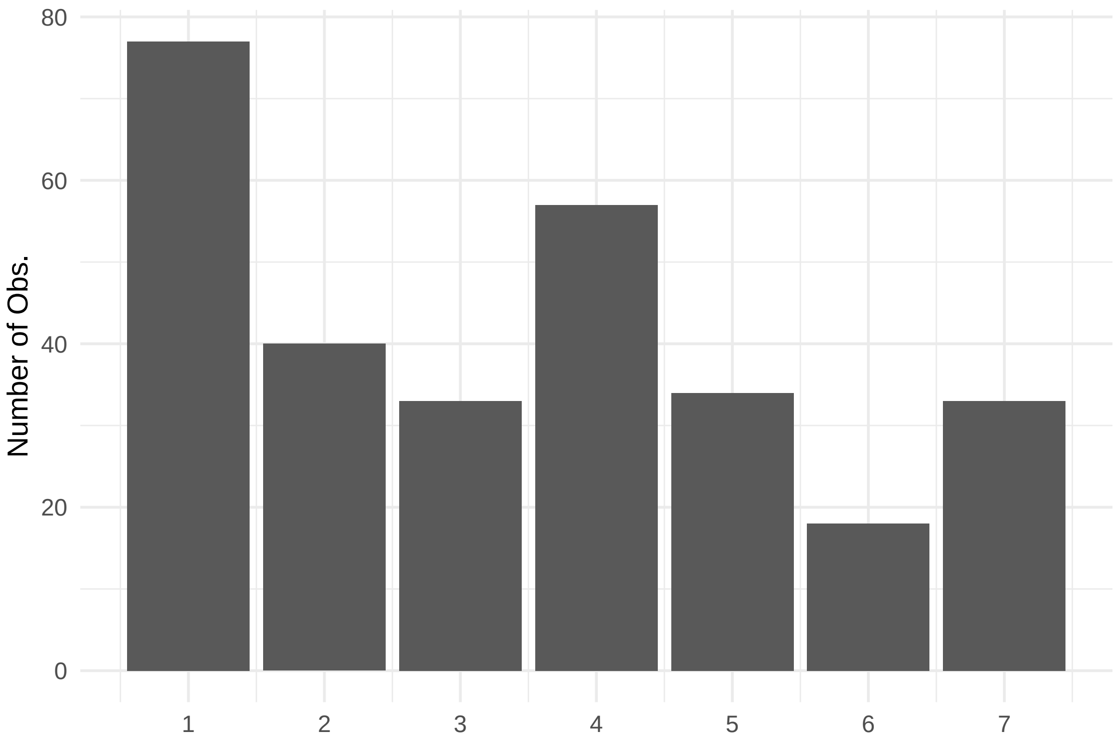

Figure 2.7: Answers to the question ”How likely would you redistribute unequal payoffs among employees?” on a 7-point Likert scale

If decision-makers are only interested in material well-being, a self-interested, rational decision-maker would essentially take the option, as in the case of the dictator game, in which he would maximise his own payoff. It is possible, however, that the decision-maker does not consider the wage differential outlined to be unfair, so that – if the fairness-seeking hypothesis of Fehr and Schmidt, 1999 is correct – he will not take advantage of redistribution, even if he himself would be objectively better off. To test this, we ran a logistic regression, controlling for demographic variables, work experience and whether the difference of 60,000 forints (or 200 USD) outlined in the redistribution question was considered fair by respondents. As a target variable, we included the logistic variable of whether respondents would take advantage of the redistribution or not. In Table 3, this was true for regression (1) if they marked a 3 and for regression (2) if they marked a value higher than 4 on the 7-point Likert scale, so that in the latter case we could also analyse those who were more clearly inclined towards redistribution. In both cases, the difference in fairness was significant, increasing the probability that respondents would favour redistribution by around 21 and 23 percent respectively. In contrast to the question on equity, there was no significant difference in the distribution either between genders or

Table 2.3: Logistic Regression Results

	$Likertscale > 3$	$Likertscale > 4$
Gender (ref.: male)		
Female	0.01444	-0.05513
	(0.232)	(-0.995)
Nationality (ref.: foreign)		
Hungarian	-0.11848	0.01351
	(-1.19)	(0.152)
Residence (ref.: capital city)		
Village	-0.12869	-0.06099
	(-1.048)	(-0.557)
City	-0.01698	0.02066
	(-0.545)	(0.744)
Work Experience (ref.: $<$ 1 year)		
1–3 years	0.15062*	0.10959
	(1.696)	(1.384)
3–5 years	0.21355	0.23076*
	(1.641)	(1.989)
More than 5 years	0.06939	0.02607
	(0.82)	(0.346)
Education (ref.: higher education)		
High school diploma or lower	0.05718	0.10587
	(0.805)	(1.672)
Unfair difference	-0.23351**	-0.2646***
	(-2.98)	(-3.788)
Constant	0.55069***	0.27057**
	(6.017)	(3.319)
Observations	292	292
Log-Likelihood	-204.31	-170.747
AIC	428.62	361.493

Note: In case (1) the respondents indicated a value of 3 on a 7-point Likert scale, whereas in case (2) the respondents indicated a value higher than 4.
Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

between Hungarians and foreigners, but a few years of experience showed a significant difference compared to less than one year of experience. Among those who would have clearly divided the wage gap - i.e., who marked a 5 or higher on the Likert scale for this question - there was a significant difference between those with tertiary education and those with only a secondary school leaving certificate or less. On the basis of the regression results, we cannot reject the null hypothesis for Hypothesis 4, which states that if we perceive a payoff above the reference point, we will try to make the reference point closer to us. Whether or not we do so to the detriment of another person is significantly affected by whether or not the perceived difference is fair, as shown in the results. The result is consistent with the findings of Lin and Bates, 2021, where a significant relationship was found between redistribution and (malicious) envy. The results also support the theory of

Marescaux et al., 2021 on workplace comparisons, according to which perceived unfair differences may encourage decision makers to engage in harmful behaviour.

2.5.2 The role of information in fairness

We also assumed that the information available at the time would have a significant impact on the assessment of fairness, and thus on the actual implementation of the malicious action. That is, in the case of malicious envy, we do not take into account the cause of the differences we experience (i.e., for example, differences in expenditure), but if we have information about them, we can judge the difference as fair, so that if we are not driven by extreme egalitarianism, malicious envy may disappear. Following on from the previous question, we asked how likely they would be to take up the redistribution option if their higher-earning colleague a) worked more, b) studied more or c) had been with the company longer. Here, we hypothesise that different results may emerge: while the first piece of information is concrete evidence of a greater effort on the part of the work partner, the other two reasons – possibly greater professional or in-house experience – give a much more nuanced picture of the effort involved.

Figure 2.8 shows the degree of change on the Likert scale from the original distribution depending on the information received, while Figure 2.9 shows the direction of the shift. The figure clearly shows that the greatest impact on perceptions of fairness was for tangible effort: in this case, 61 per cent of respondents shifted the probability of redistribution in a negative direction, compared to 45.5 per cent for more experience and only 37.7 per cent for more learning. In the case of more hours worked, only a third of respondents remained indifferent, while for experience and learning it was 45.5 and 50 per cent respectively. It is also worth noting that, in the latter two cases, one in ten respondents increased their likelihood of redistributing the pay gap, i.e., a shift in the opposite direction to that expected.

Overall, therefore, it can be said that new information has an impact on the perception of fairness, but different types of information have different effects. For example, some information may have the potential to increase feelings of envy, thereby increasing the tendency for the decision-maker to further reduce inequalities. An alternative explanation for the emergence of comparison and fairness and malicious envy can be provided based on the results of J. Kim and Park, 2018, who found a positive relationship between envy and fairness, but when the difference was accompanied by a loss of self-esteem, decision-makers became more prone to counterproductive, harmful behaviour.

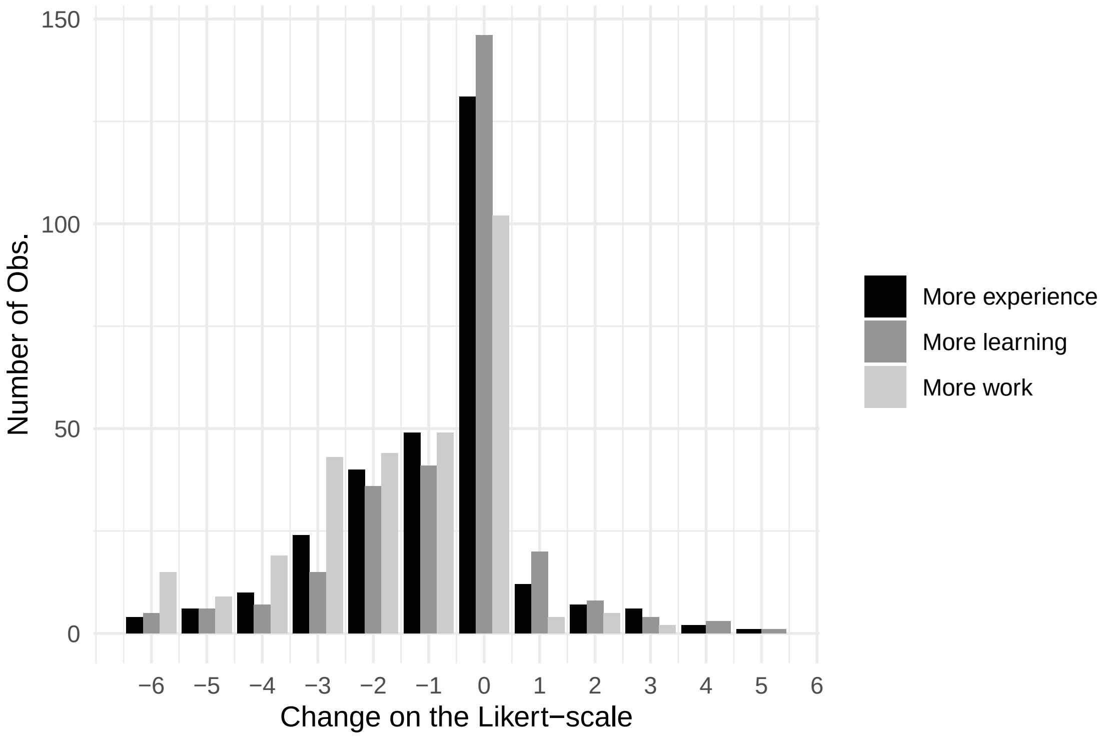

Figure 2.8: Change in the probability of redistribution on a 7-point Likert scale for different types of information

It can also be argued that, based on the distributional issues, we cannot reject hypotheses 2 and 3 for the comparison with the reference population. However, it is worth pointing out the fact that in the present study we present simplified versions of the different types of information, and it may be worthwhile to consider in more depth the rationalisation of the differences in the future. For example, it may well be that the actual effort of the other person makes the difference more acceptable than simply the number of hours worked - i.e., there may be an even closer link between a quantifiable difference in performance and equity.

The results could prove useful for companies where managing wage tensions is a problem. On the one hand, excessive wage tensions can lead to malicious behaviour within the firm, which should be addressed (an important topic in management science, see for example Marescaux et al., 2021, Kim et al., 2020). The questionnaire suggests that a wage differential of 40-50% may be perceived as fair by peers (not counting respondents with extreme egalitarian views), but wage differentials beyond this level bring the possibility of wage redistribution to the fore. However, if there is some justification for the wage difference - whether it is greater work experience, higher education or greater effort - the

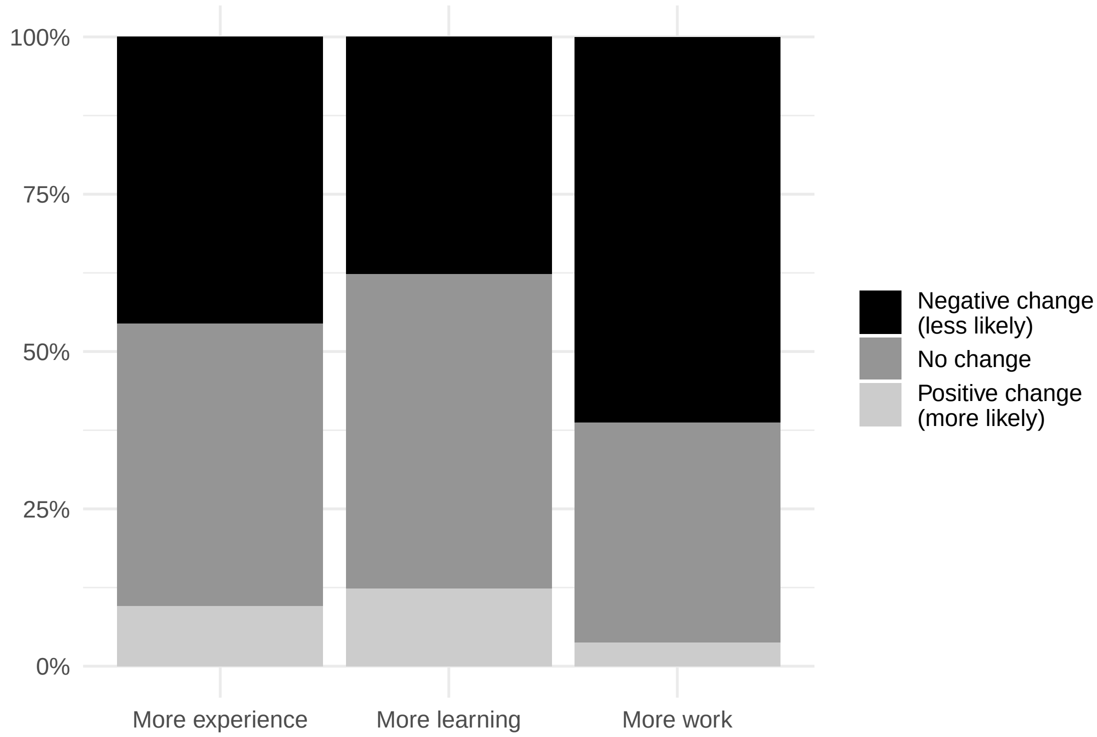

Figure 2.9: Direction of change on Likert scale by type of information (percentage)

risk of malicious envy can be significantly reduced. Among these, we should highlight the case where the pay differential is actually due to effort or performance: when this information was emphasised in the questionnaire, the risk of malicious envy was significantly reduced. Thus, for example, the evaluation of effort or performance in a company performance appraisal can also mitigate pay tension.

2.6 Conclusions

This study presented the importance of reference points, fairness and envy in decision theory. The identity of the referent and the formation of the reference point are of primary importance in the comparison, and were therefore examined first. Our analysis confirmed previous findings in economics, psychology and philosophy that comparisons with people far away from us serve only a long-term motivational purpose, and that we actually compare ourselves to people close to us in our decision-making.

When we talk about comparison with others, what actors consider fair is a very important factor. To paraphrase: what is the difference between the decision-maker and the reference person that the decision-maker still considers acceptable and fair? This question can be of great importance in transactions, as an unfair difference can have negative behavioural consequences. For example, unfairness may lead the decision-maker to withdraw from the transaction, i.e., both parties may end up worse off (e.g., in a consumer market, the seller may not sell his goods and the buyer may either withdraw from the transaction altogether or incur transaction costs by seeking other options). In such cases, there may also be a sense of envy or, less frequently, malicious envy, where the decision-maker reduces the utility of the other party, even to the extent of reducing his own utility.

The results, confirming previous literature, highlight the important role of equity in decision-making, especially on distributional issues. However, if more information is provided about the differences, a significant proportion of decision-makers can rationalise the difference and accept it. It is important to note that the type of information plays a large role in the degree of variation: for example, a clear distinction in effort is accepted as fair by decision-makers, but a difference in pay due to more nuanced differences, such as experience or skills, is less accepted by respondents. Another important finding is that, although there was a difference between Hungarian and foreign respondents in their perception of the extent of the fair pay gap, there was no significant difference between Hungarian and foreign respondents on the question of how to reduce the gap, i.e., redistribute it. These results are limited to being correlations rather than causality due to the nature of the methodology. Future studies could try to disentangle these country-specific effects in a controlled lab-experiment setting.

Further research should explore the factors that influence perceptions of the fair pay gap more broadly, for example, which factors are more likely to influence specific domestic thinking about fair differentials. While the study shows that time spent at work is clearly one such factor, there may be other factors, such as social factors, that can influence decision-making processes. On the one hand, such knowledge could help public policy or economic policy making, for example on the introduction of different taxes and tax schemes, as discussed by Devos, 2013. On the other hand, a more accurate picture of equity could help firms design their wage systems – reducing the risk of wage tensions.

Chapter 3

Stability of Economic Preferences

Evidence from a Representative Survey

3.1 Introduction

Several recent studies have examined the evolution of preferences and preference measurements over time. The stability of preferences in both the short and long run has been analyzed across various dimensions, including measurement methods (e.g., surveys vs. experimental approaches (Chuang and Schechter, 2015)), temporal changes (Chuang and Schechter, 2015; Dasgupta et al., 2017), external economic shocks (Carlsson et al., 2014; Hardardottir, 2017; Krupka and Stephens Jr, 2013), and specific disruptions such as the COVID-19 pandemic (Alsharawy et al., 2021; Bokern et al., 2021; Harrison et al., 2022; Shachat et al., 2020). The stability and consistency of economic preferences are critical to microeconomic theory, as these concepts underpin fundamental assumptions about individual decision-making (Arrow, 2012; McFadden, 2001; Sugden, 1985).

Despite substantial research on individual preference stability, there remains a gap in studies addressing the stability of economic clusters. This gap may be due, in part, to the relatively recent application of clustering methods in behavioural economics and the growing use of surveys and experiments to measure multiple preferences. The key question I raise are three-fold. First, whether we can cluster individuals into meaningful clusters using only measured economic preferences. Second, whether these clusters are stable over-time, i.e., whether they occur when we measure these preferences in a later period. Third, whether the predictions on other individual characteristics and background variables (such as: age, income, gender, etc.) based on the clusters defined remain similar in both measurements. Conducting this research allows us to learn more about the non-linear connection between economic preferences and outcome variables. [1] Additionally, we may get a better of understanding of ”economic personality types” similar to psychology.

In this paper, I analyse the stability of economic preferences using data from a Hungarian representative survey conducted in the second half of 2020. The survey measured several preferences and personality traits, including time preferences, risk aversion, co-operation, altruism, competitiveness, and locus of control, across two waves: the first in June and the second in November. The dataset primarily consists of pooled cross-sectional data, with approximately 200 individuals participating in both waves. To examine the stability of individual preferences, I focus on this panel sub-sample. For clusters of economic preferences, I apply partitioning clustering methods to identify distinct groups of preferences and compare their stability between the two waves. As an alternative approach, I use the cluster results from the first wave as a benchmark to classify the observations from the second wave; then, I analyse whether external variables not used for the clustering - such as age, gender, earnings - differ in the same way as observable for the first wave. To my knowledge, this is the first study to analyse the stability of economic preferences using these methods.

I connect to the literature on two strains: the first is the grouping of preferences. The classification of preferences may happen on the level of measured preferences, using methods such as the Principal Component Analysis (Chapman et al., 2023; Lades et al., 2021). Alternatively, researchers can use clustering algorithms to group individuals together based on their similarities regarding preferences. This methodology is relatively new in behavioural and experimental economics, and it emerged with the opportunity to measure multiple preferences at once, such as in Chowdhury et al., 2022; Epper et al., 2024; Fehr and Charness, 2023. Grouping individuals addresses whether ”profiles” can be constructed based on economic preferences and whether there are—potentially non-linear—connections between these profiles and economic outcomes (Chowdhury et al., 2022). In other words: we can test whether there These findings are similar to the Big Five personality traits commonly used in psychology.

I find that economic preferences remained relatively stable within the sample. In the panel subsample, only trust and altruism changed significantly. Analyzing the pooled cross-sectional data, time preferences—specifically time inconsistency and the discount

[1] Important to note, however, that the analysis does not allow us to detect causal effects – we can only identify correlations with such methods.

factor—along with willingness to cooperate and external locus of control, were found to be stable. Turning to clusters of economic preferences, I find that when running the clustering algorithms separately for the two waves using k-medoid clustering, a group consistently emerges, consisting of individuals who are more cooperative, competitive, relatively less patient, and exhibit a more internal locus of control. Members of this group tend to be younger, more educated, and have higher incomes compared to the other cluster.

As a robustness check, I ran the clustering on the first wave and used those results to predict cluster membership for individuals in the second wave. Comparing the variables not used in the clustering—age, gender, wage, and education level—I find that these variables differ in the same direction and are statistically significant. Thus, there is evidence of stability in economic clusters, and these preference groups yield consistent implications for socio-economic variables. It is important to note that throughout the paper, any connections between clusters and economic variables should be understood as correlations, not causal relationships.

Following the introduction, the paper proceeds as follows: in Section 3.2, I summarise the relevant literature. In Section 3.3, I introduce the data used for the analysis with the relevant summary statistics. In Section 3.4 I show the results for the clustering analyses; finally, in Section 3.5 I conclude the paper.

3.2 Literature Review

As mentioned in the introduction, this paper relates to two bodies of literature: first, the literature on the clustering of preferences, and second, the stability of economic preferences. The clustering of economic preferences is a relatively new area of study within economics. [2] The use of clustering algorithms offers the advantage of addressing potential issues of multi-collinearity and mitigating them in analyses. Moreover, these algorithms can account for possible non-linearity

Lades et al., 2021 used Principal Component Analysis (PCA) on 20 pro-environmental behaviours, utilising the day reconstruction method (Kahneman et al., 2004). They identified four distinct factors: eco-shopping, electricity- and water-saving behaviours, awareness, and waste- and consumption-reducing behaviours. They then analysed the predictive properties of seven economic preferences, based on the measurements from the Global Preference Survey (Falk et al., 2018), and found that altruism, positive reciprocity, and pa-

[2] Although clustering algorithms are frequently used in other fields of economics, such as international trade (Diaz-Bonilla et al., 2000; Disdier and Van Tongeren, 2010; Vahalík and Staníčková, 2016

tience were strong predictors of the identified factors. In a more general study, Chapman et al., 2023 also applied PCA to 21 different preference measurements on a representative sample of the US population. They found that the measured preferences could be grouped into six clusters: Generosity, Risk Aversion (Willingness to Accept), Willingness to Pay/Inequality, Overconfidence, Impulsivity, and Uncertainty. Additionally, they found that these factors correlate with cognitive abilities.

For clustering at the individual level, Chowdhury et al., 2022 used k-medoid clustering on risk, time, and social preferences for Bangladeshi families. They identified two large clusters: the first, more patient, risk-tolerant, and prosocial, while in the other cluster, families were more impatient, risk-averse, and spiteful. In a similar vein, Epper et al., 2024; Fehr and Charness, 2023 used Dirichlet Process Mixture algorithms on individuals' decisions in twelve money-allocation tasks, where the experimenters varied the benefits and costs of redistribution. They identified altruistic, selfish, and inequality-averse groups within the general population; for university students, they also identified altruistic and selfish groups, but the third, smaller group could not be categorised in a straightforward manner. One of the shortcomings of these studies is the use of a single clustering algorithm, an important factor when considering the robustness of the results (Ertl et al., 2024).

Economic preferences were also found to be connected with other outcome variables. For example, the preference clusters identified in Chowdhury et al., 2022 were linked to income and household size, with relatively more patient, risk-tolerant, and prosocial families having higher incomes and larger households. Alternatively, economic preferences were analysed for their predictive power on behaviour (Breitkopf et al., 2024), norm-enforcing behaviour (Friehe and Schildberg-Hörisch, 2018), and labour market occupational choice (Vaaramo et al., 2024).

There are studies that examine the long- and short-run stability of such preferences, may that be risk (Chuang and Schechter, 2015; Dasgupta et al., 2017; Salamanca, 2018; Schildberg-Hörisch, 2018),time (Chuang and Schechter, 2015; Hardardottir, 2017; Meier and Sprenger, 2015; Salamanca, 2018), competitiveness and confidence (Dasgupta et al., 2017) or social preferences (Bruhin et al., 2019; Carlsson et al., 2014; Chuang and Schechter, 2015; Lotz et al., 2013). These studies find varying stability regarding preferences, with risk preferences perceived to be more unstable, while social preferences are more stable (Chuang and Schechter, 2015). The (in)stability of preferences may depend on context (De Oliveira et al., 2012; Lotz et al., 2013) or group-dependence of preferences (Böhm et al., 2021), while time-preference instability was found to be in connection with macroeconomic and financial outlooks (Hardardottir, 2017; Krupka & Stephens Jr, 2013). While these differences might be expected, there is some evidence that instability of risk preferences can also be connected to subjects being confused by the questions (Chuang and Schechter, 2015), or the measured risk preference may depend on the method it is being measured (Reynaud and Couture, 2012).

When analysing stability of preferences, Dasgupta et al., 2017 grouped the literature to four parts: a) stability of choices in different domains (i.e., contributing to a public good in a laboratory experiment vs. contributing to building a bridge, as in Carlsson et al., 2014); b) stability of personality traits (i.e., whether Locus of Control, or the Big Five personality types are stable throughout a person's lifetime) c) stability of economic preferences, (i.e., whether someone behaves in experiments similarly over-time), and d) state-dependent economic preferences (i.e., whether one becomes more risk-averse after experiencing the Great Financial Crisis). This paper mainly covers the second and third points.

Chuang and Schechter, 2015 provides an extensive survey of papers related to the stability of risk, time, and social preferences. For risk preferences, out of 19 papers, only two (plus one inconclusive) found that risk preferences were not stable over time. For time preferences, out of eight papers, only one was inconclusive. For social preferences, out of the four surveyed papers, two showed partial evidence of non-stability. Overall, they found that survey-based observations were more stable compared to experimentally measured observations.

The stability of economic preferences during the COVID-19 crisis has also been explored in the literature. Using pooled cross-sectional data, Alsharawy et al., 2021 measured multiple preferences in three waves at the start of the pandemic. They found that economic preferences varied in the short run, with the variation explained by individuals' fear of COVID-19. Harrison et al., 2022 conducted online experiments between May and October 2020 and found significant differences in temporal risk preferences compared to pre-pandemic measurements. Shachat et al., 2020 measured preferences in Wuhan over multiple waves for six weeks, starting from the imposition of the lockdown. They found that during this period, measured altruism, cooperation, trust, and risk tolerance preferences changed significantly. Altruism and cooperation increased overall, but willingness to cooperate and trust decreased among those who remained in Wuhan during the lockdown. Contrary to these findings, Bokern et al., 2021 measured economic preferences before and during the pandemic and found economic preferences measured through incentivised tasks to be stable.

3.3 Data

I used survey data from representative samples (in terms of gender, age, education, and settlement type) conducted by Tárki, a Hungarian polling company (on behalf of the HUN-REN CERS Institute of Economics). The surveys were conducted in two waves: the first in June 2020, and the second in November. Due to the pandemic, the surveys were carried out using phone interviews. In the first wave, 1,025 people participated. In the second wave, 1,013 people took part, with 204 individuals participating in both waves.

The surveys consisted of three parts: in the first part, respondents were asked about their socio-demographic data. In the second part, they were surveyed about various economic preferences, which were not incentivised. Notably, some questions were asked in the first wave but not in the second. [3] To compare the results of the two waves, I only used economic preferences measured in both. Finally, in the third part, participants were asked about their financial status.

Table 3.3 shows the summary table for the preferences used in the analysis. The measured preferences were: Time preference, Risk, Cooperation, Altruism, Competitiveness and (external) Locus of Control.Histograms for these variables are found in section 1 of the Appendix. The questions were asked in the order of listing. The answers were validated and were found to be consistent with the literature (as shown in Khayouti et al., 2021).

Most of the variables could be used for the analysis as is, with the notable exception of the locus of control measurements, consisting of seven separate questions to determine whether an individual perceives internal or external locus of control. To effectively measure locus of control, seven questions were asked, changing between the internal and external-oriented wordings. One method to identify the type of locus of control is by scoring the answers to these questions appropriately. However, I opted for running a Principal Component Analysis (PCA) and extracting the first factor, similar to Piatek and Pinger, 2010. Thus, the measurement is standardised with a mean of zero; based on the factor loadings I obtained, the measure corresponds to external locus of control. It is important to note that I only ran PCA on Wave 1 observations; Wave 2 data was projected onto these factor loadings. [4]

[3]For example, in the first wave, cooperation was measured using three questions, while in the second wave, only one of these questions was asked.

[4]This was done to maintain consistency across the sample. A slightly different distribution of the variables in Wave 2 would mean that the factor loadings would also be different, thus there could be a slight distortion in the resulting aggregated variable

For the analysis, I restrict the data by filtering out all missing values. [5] Thus, sample size is reduced to 742 in the case of the first wave, and to 772 in the case of the second wave. [6]

[5]Respondents could answer ”I do not know” or ”I do not want to answer” as well during the interview. While from a clustering point of view, there are methods to use missing values as well (k-medoid, for example, as shown in Chowdhury et al., 2022, due to the mixed nature of the data, I decided to only concentrate on respondents who were sure about their decisions.

[6]To test whether this restriction of the dataset significantly alters the average values of the variables, I ran tests on the differences between the unrestricted and restricted databases; the results are shown in Section 2 in the Appendix; there were marginal decreases in age and household size, while the average net wage increased in the restricted sample.

Table 3.1: Measurements of economic preferences

Variable	Measurement	Wording of the question
Time Preference 1	Stairways method similar to Falk et al., 2018 - inter-temporal decisions between getting paid now vs. getting paid a higher amount one month from now	Would you rather have 10,000 HUF now or 12,500 HUF in one month?
Risk preference	Out of 10,000 HUF, how much do respondents are willing to take on a bet of heads or tails - based on Sutter et al., 2013	How much money would you bet on a heads or tails bet?
Cooperation	Likert scale from 0-5 based on Kasik, 2015	How typical are the following statements for you? I like solving problems in a group.
Trust	Likert scale from 0-5	In general, what would you say most people can be trusted (5), or that you can't be too careful (0)?
Altruism	Donation question	Imagine the following situation: Today, unexpectedly, you receive 100 000 HUF. How much of this would you donate to a good cause?
Time Preference 2	Stairways method similar to Falk et al., 2018 - inter-temporal decisions between getting paid one year from now vs. getting paid a higher amount one year and one month from now	Would you rather have 10,000 HUF in one year or 12,500 HUF in one year and one month?
Competition	Likert scale ranging from 1-5 based on Fallucchi et al., 2020	Please tell me on a scale of five how true the following statement is for you. Competition brings out the best in me.
(External)Locus of Control	Based on Pearlin and Schooler, 1978, seven questions were asked on feeling in control. Note: for the analysis, I used the first component of a Principal Component Analysis to measure Locus of Control.	Please tell me on a scale of five how much you agree with the following statements. Sometimes I feel like I'm just drifting along in life. (first question)

Note: From Time preference 1 and Time preference 2 , I calculated the variables $delta$ and $beta$.

Table 3.3 shows the summary statistics for both Wave 1 and Wave 2, with the appropriate t-tests in the last column. All variables used in our clustering are above the dashed line. In the empirical analysis, we will use groupings of preferences based on clustering algorithms to examine whether significant differences exist based on age, household size, wage, education level, and gender.

Table 3.2: Summary Statistics for Waves 1 and 2, with the p-values of the appropriate t-tests shown in the last column

Variable	Wave 1 N = 742	Wave 2 N = 772	P-value of Appropriate Tests
Risk	3,557	3,329	0.0604
	(3,421)	(3,469)	(Wilcoxon rank sum test)
Trust	2.19	2.23	0.7289
	(1.49)	(1.51)	(Wilcoxon rank sum test)
Cooperation	3.58	3.48	0.1057
	(1.40)	(1.39)	(Wilcoxon rank sum test)
Altruism	25,614	32,281	<0.0001
	(24,733)	(27,568)	(Wilcoxon rank sum test)
Competition	3.32	3.19	0.0361
	(1.28)	(1.32)	(Wilcoxon rank sum test)
Time-inconsistency	1.04	1.04	0.6772
	(0.24)	(0.25)	(Wilcoxon rank sum test)
Discount Factor	14,148	14,231	0.7768
	(3,675)	(3,704)	(Wilcoxon rank sum test)
Locus Of Control	0.00	0.00	0.9722
	(1.00)	(1.00)	(Wilcoxon rank sum test)
Age	51	51	0.5877
	(17)	(17)	(Wilcoxon rank sum test)
Household Size	2.55	2.73	0.0165
	(1.28)	(1.40)	(Wilcoxon rank sum test)
Net Wage	181,746	191,570	0.0439
	(206,758)	(158,646)	(Wilcoxon rank sum test)
Unknown Wage	269	235	
Education level:			0.2469
Less than secondary school	270 (36%)	250 (32%)	(Pearson's Chi-squared test)
Secondary school	304 (41%)	331 (43%)	
University Degree	168 (23%)	191 (25%)	
Gender:			0.5227
Female	379 (51%)	407 (53%)	(Pearson's Chi-squared test)
Health Concerns	2.91	2.83	0.0054
	(0.66)	(0.71)	(Wilcoxon rank sum test)
Family Health concerns	2.99	2.84	<0.0001
	(0.87)	(0.97)	(Wilcoxon rank sum test)
Financial Concerns	2.66	2.56	0.0938
	(0.96)	(0.86)	(Wilcoxon rank sum test)

Looking at preferences, there are only a couple of differences: for the question ”out of 10,000 HUF, how much are you willing to take on a bet of heads or tails”, the amount decreased marginally, from 3,550 to 3,330, but the difference is not significant at the 5% significance level. The measured altruism was significantly higher in the second wave, while competition was lower. Testing separately for the seven locus-of-control questions, only one was significantly different - once aggregated, this difference disappeared. Overall, we do not find statistically significant differences for risk, trust, cooperation, time preferences and external locus of control. This initial analysis suggests stability for most of the measured preferences.

Additionally, under the dashed line, I included the summary statistics of important background variables available to us, namely: age, household size, net earnings, education level, and gender. There are differences in household size (statistically significant, but not economically), as well as in net earnings; the latter difference could be attributed to the fact that the first wave was conducted right after the first Covid wave, when many people lost their jobs, were moved from full-time to part-time, or were forced into unpaid leave Gáspár and Reizer, 2020. Reassuringly, after the data filtering, there are no differences in age, gender, or education level between the two samples.

The survey being conducted in 2020, individuals were also asked how vulnerable they feel to Covid-19. Respondents were asked to answer on a scale of 1-7 how dangerous they feel Covid is a) to their own health, b) to the health of family members, and c) to their financial situation. These questions cover most of the external shocks being present at the time of the survey; these being relatively unchanged between the two waves indicate that incidental instability is not rooted in external shocks.

Looking at the Covid-related variables, we see that generally, respondents felt less vulnerable to Covid-related shocks during the second wave of the surveying; the differences in health being significantly different. Thus, in the pooled cross-sectional data structure, the differences between the two waves due to Covid-exposure might play a role.

3.4 Results

3.4.1 Stability of preferences - panel subsample

As a first step, I analyse the stability of individual preferences. It is important to note that the comparisons of Table 3.3 cannot be interpreted causally. Differences might arise from three sources: first, the (in)stability of preferences; second, changes in preferences (due to the external shock of Covid-19); and finally, due to the sampling method. However, by analysing individuals who were surveyed in both waves, we mitigate the latter issue.

Table 3.3: Comparison of responses for respondents being present in both Wave 1 and Wave 2.

Variable	Mean (Wave 1)	Mean (Wave 2)	Pairwise diff.	P-value	Count
Risk	3157.46	3186.07	198.52	0.55	165
Trust	2.20	2.41	0.22	0.05	204
Cooperation	3.52	3.44	-0.08	0.51	203
Altruism	29112.50	33662.55	5253.49	0.02	189
Competition	3.36	3.41	0.04	0.64	202
Time Inconsistency	1.04	1.05	0.00	0.99	188
Discount Factor	14345.38	14084.69	-69.41	0.81	188
Locus of Control	0.00	0.00	0.00	0.78	199
Net Wage	177521.47	204804.05	25785.99	0.01	111
Health Concerns	2.81	2.72	-0.10	0.08	204
Family Health Concerns	3.00	2.83	-0.17	0.12	204
Financial Concerns	2.60	2.62	0.02	0.78	204

Note: Pairwise comparisons were only made where in both waves, respondent answered the question (excluding ”I don’t know” and ”not willing to answer” responses).
P-values indicate to pairwise test appropriate for the variable type. Paired t-tests were performed on numerical, Wilcoxon signed-rank test were performed on ordinal and Chi-Square tests were performed on categorical (binary) variables.

Out of our sample, we have 204 individuals who were both surveyed during the first and second waves. Thus we can analyse whether there were differences not on the sample, but on the individual level, eliminating uncertainty from the sampling process. A disadvantage, however, is the sample size being relatively small. Table 3.4.1 shows the mean values of the answers in both waves, and the t-test for the pairwise differences. In the last column, I also included the number of observations for the test of given variable. For example, for wages, 111 of the original 204 individuals answered in both waves.

Similarly to Table 3.3, significant difference is found for altruism, but not for risk. [7] Additionally, Table 3.4.1 shows a significant difference in trust.

Looking at the test results for Covid-exposure, people felt less vulnerable during the second wave (at the end of 2020; with the exception of financial vulnerability being marginally higher on average in the second wave). The difference, however, was only marginally significant in case of the responder’s own health (p = 0.084). Wage differences were also statistically significant between the two waves.

Overall, comparing pooled cross-sectional and panel structures, we see that time pref-

[7] While in Table 3.3 risk preferences showed significant difference at the 10% significance level, the opposite can be found in the calculations of Table 3.4.1.

erences, as well as willingness to cooperate and external locus of control, are stable, with mixed evidence on risk preference, trust, and competitiveness. To summarise, in the panel sub-sample, I find that only trust and altruism changed between the first and second waves; all other measured preferences remained relatively stable. Differences in reported Covid exposure were not statistically significant at the 5 percent significance level, suggesting that external shocks have a more moderate effect in this subsample. The difference in altruism aligns with the findings of Shachat et al., 2020, though they did not find a significant difference in trust. [8]

3.4.2 Stability of clusters of economic preferences

In the following sections, I analyse whether there are groups that can be defined based on the measured economic preferences, and whether they are stable. Additionally, I examine whether there is a connection between the defined clusters of preferences and background variables such as age, education, gender, household size, and earnings.

I argue that there are two ways to analyse cluster stability. First, we can run the same clustering algorithms (that is, the same algorithm with the same distance method and the same number of "k" clusters specified) and compare the results. Here, we can first analyse whether we can identify groups that are internally cohesive and externally isolated for each wave. Then, we can compare the results for the two waves and check whether similar groups emerged. I use this first method in Section 3.4.2. However, there might be some issues with this methodology. While it has been shown that Euclidean k-clustering is consistent (Yoshida & Ito, 2022), it is possible that we do not identify the same groups due to sampling.

Alternatively, one can define clusters based on the first wave; then, clusters for the second wave can be defined by using only the preference measures and projecting them onto the cluster structure of the first wave. By conducting the analysis this way, we mitigate the data sensitivity of clustering algorithms. Then, the background variables can be compared between the two groups across the two waves. If we find that, for example, the first cluster consistently exhibits higher average wages in both waves, that would provide suggestive evidence of cluster stability. In other words, if we assume the first clustering result to be accurate, we can use the classifications of individuals as a baseline; then, we

[8] However, it is important to note that Shachat et al., 2020 measured trust using the Trust Game, whereas in our data, a Likert-scale-based question was used to proxy general trust towards the populace.

can assign observations from the second wave to clusters based on the classification of their nearest neighbour specified in Euclidean space. Section 3.4.2 pursues this idea.

Comparing separately identified clusters

When using clustering methods, the researcher has to make decisions on the type of the clustering algorithm (hierarchical, partitioning , model-based; bayesian, non-bayesian, etc.), the distance measures to be used (for example, Euclidean or Manhattan-distance), and on the number of clusters to create (which has to be decided before running the algorithm, with a few exceptions such as in the case of hierarchical clustering). The challenge using such unsupervised machine learning algorithms is that there is no universally correct way of making these decisions, as we lack information on the ”true outcome” (i.e., the ”true” underlying groups based on preferences). [9]

The two most commonly used algorithm types are hierarchical and partitioning algorithms. Hierarchical clustering groups individual observations that are close to each other in order to find the next set of similar groups; it can either start from the individual observations (agglomerative) or begin by dividing all observations into larger, similar groups (divisive clustering). In the case of partitioning clustering, the algorithm decomposes the data into a set of groups based on their proximity to iteratively calculated cluster centres. The main difference between the two methods is that the former does not require a predefined number of clusters (k). For the main analysis, I use the k-medoid algorithm. One advantage of the k-medoid algorithm compared to other partitioning algorithms such as k-means clustering is that it is less sensitive to outliers in the data. This method was used in Chowdhury et al., 2022, the only paper that I know of that uses clustering algorithms to classify groups of economic preferences. [10]

Besides the mixed nature of the data, one of the challenges is handling ordinal variables. The treatment of ordinal data has been discussed both in the context of multivariate analysis (for example, Kampen and Swyngedouw, 2000; J.-O. Kim, 1975) and factor analysis (for example, Jöreskog and Moustaki, 2001). More recently, Robitzsch, 2020 shows that for factor analysis, treating these variables as either continuous or ordinal can lead to biased estimates depending on the data-generating process. In the main analysis, I chose to treat ordinal variables as continuous, allowing us to use Euclidean distance. [11]

[9] For a detailed overview on Cluster Analysis, see Everitt et al., 2011

[10] As a robustness test, I use hierarchical clustering. However, I find that the latter essentially captures outliers very well, making it harder to generalise the results. I include the results of that analysis in Section 3 of the Appendix.

[11] As a robustness test, I ran the the analysis using k-prototype algorithm which can handle the mixed

The analysis is threefold: first, we want to check whether the clustering results are meaningful, i.e., there are statistically significant differences between the variables used for the clustering itself (in other words: observations in different clusters are indeed different from each other). Second, whether there are any differences in the background variables not used during the clustering methods (such as age, gender, education, earnings, and Covid exposure). [12] Third, whether we see the same or similar clusters emerge in both waves.

In order to decide upon the number of clusters, I calculated the optimal number of clusters based on the Elbow - and Silhouette-methods, implemented in R via the ”NbClust” package (Charrad et al., 2014). For both methods, k = 2 were the optimal choice for both hierarchical and partitioning algorithms.

Table 3.4 shows the comparison of the two generated clusters in both waves. Above the dashed line are the variables used throughout the clustering, while under the dashed line are the background variables. Additionally, appropriate t-tests were calculated for the differences between the two clusters.

First, we can note that the clusters are internally meaningful, in the sense that all variables used for the clustering are significantly different in the two groups across both waves (with the exception of external locus of control in Wave 2, although the difference was highly significant for Wave 1). Looking at the background variables under the dashed line, we see that the created groups are also divisible by age, wages, and education level. In both waves, one of the groups was relatively younger, earned more on average, and had higher education (at both secondary and tertiary levels); these groups were also more balanced in terms of gender (although not significantly different in Wave 1).

Comparing Cluster 2 across the two waves, there are some similarities: Cluster 2 is consistently associated with higher levels of cooperation, more competitiveness, less time-consistency, and a higher discount factor; individuals are also more prone to an internal locus of control. In some cases, the direction of the difference shifted. Risk preference, trust, and altruism changed direction for Cluster 2: while in the first wave, Cluster 2 exhibited lower risk-taking preferences, relatively lower trust, and a lower level of altruism; in the second wave, the opposite is true compared to Cluster 1. It is worth noting, how-

nature of the data. While the results show a more modest difference between the clusters, the direction of the average differences is shown to be consistent between the two waves. The results for this analysis are included in Section 4 of the Appendix.

[12] It is very important to note that this analysis concerns mere correlations, not causality. For example, we do not yet know whether economic preferences cause changes in economic outcomes (such as wages), or vice versa (higher wages and socioeconomic status cause people to be more trusting). As such, I intend to refrain from any causal wording

ever, that of these three, trust and altruism changed significantly between the two waves, as observed in the panel sub-sample analysed in Table 3.4.1.

To summarise, using k-medoid clustering, we were able to identify meaningful clusters, and these clusters were found to be associated with higher earnings, relatively lower age, and higher educational attainment. There is some evidence for the stability of these clusters; however, external changes and their potential effect on the measurement of preferences limit our analysis.

Table 3.4: Comparison of clusters for Wave 1 and Wave 2, using K-medoid clustering. Similarities between the found clusters in the two waves are highlighted.

	Wave 1			Wave 2		
Variable	Cluster 1, N = 342	Cluster 2, N=400	p-value	Cluster 2, N = 488	Cluster 1, N=284	p-value
Risk	4,643	2,629	<0.001	4,060	2,073	<0.001
	(3,555)	(3,010)		(3,612)	(2,798)	
Trust	2.36	2.05	0.003	2.78	1.29	<0.001
	(1.53)	(1.46)		(1.35)	(1.28)	
Cooperation	3.14	3.96	<0.001	4.02	2.54	<0.001
	(1.47)	(1.21)		(1.10)	(1.35)	
Altruism	29,659	22,156	<0.001	35,904	26,056	<0.001
	(26,009)	(23,063)		(28,990)	(23,722)	
Competition	2.96	3.64	<0.001	3.58	2.50	<0.001
	(1.32)	(1.16)		(1.18)	(1.27)	
Time-inconsistency	0.95	1.13	<0.001	1.08	0.97	<0.001
	(0.19)	(0.24)		(0.26)	(0.22)	
Discount Factor	16,581	12,069	<0.001	13,534	15,430	<0.001
	(3,830)	(1,772)		(3,339)	(3,989)	
Locus of Control	0.22	-0.19	<0.001	-0.05	0.08	0.3
	(1.05)	(0.91)		(0.94)	(1.10)	
Age	53	48	<0.001	49	54	<0.001
	(17)	(16)		(16)	(16)	
Household Size	2.52	2.58	0.7	2.79	2.62	0.15
	(1.24)	(1.31)		(1.45)	(1.31)	
Net Wage	150,497	210,827	<0.001	208,626	160,671	<0.001
	(106,394)	(265,333)		(178,980)	(106,514)	
Unknown Wage	114	155		142	93	
Education Level:			<0.001			0.002
Less than secondary school	150 (44%)	120 (30%)		137 (28%)	113 (40%)	
Secondary school	122 (36%)	182 (46%)		216 (44%)	115 (40%)	
University Degree	70 (20%)	98 (25%)		135 (28%)	56 (20%)	
Gender:			0.3			0.033
Female	182 (53%)	197 (49%)		243 (50%)	164 (58%)	
Male	160 (47%)	203 (51%)		245 (50%)	120 (42%)	
Health Concerns	2.87	2.94	0.14	2.86	2.78	0.12
	(0.75)	(0.57)		(0.68)	(0.76)	
Family Health Concerns	3.03	2.96	>0.9	2.81	2.91	0.4
	(1.03)	(0.71)		(0.85)	(1.15)	
Financial Concerns	2.64	2.67	0.9	2.61	2.48	0.043
	(1.00)	(0.93)		(0.89)	(0.80)	

Alternatives of cluster stability

So far, we have run the clustering algorithms separately and checked whether the resulting clusters were similar to each other. This analysis allows us to assess whether the same patterns emerge in two samples (i.e., whether the same kind of observations are closer to each other in Euclidean space). One critique of this method, however, is that the clustering results might be driven by the sample; while the overall characteristics of the created clusters may be similar for both waves, the importance of certain variables for specific clusters might differ.

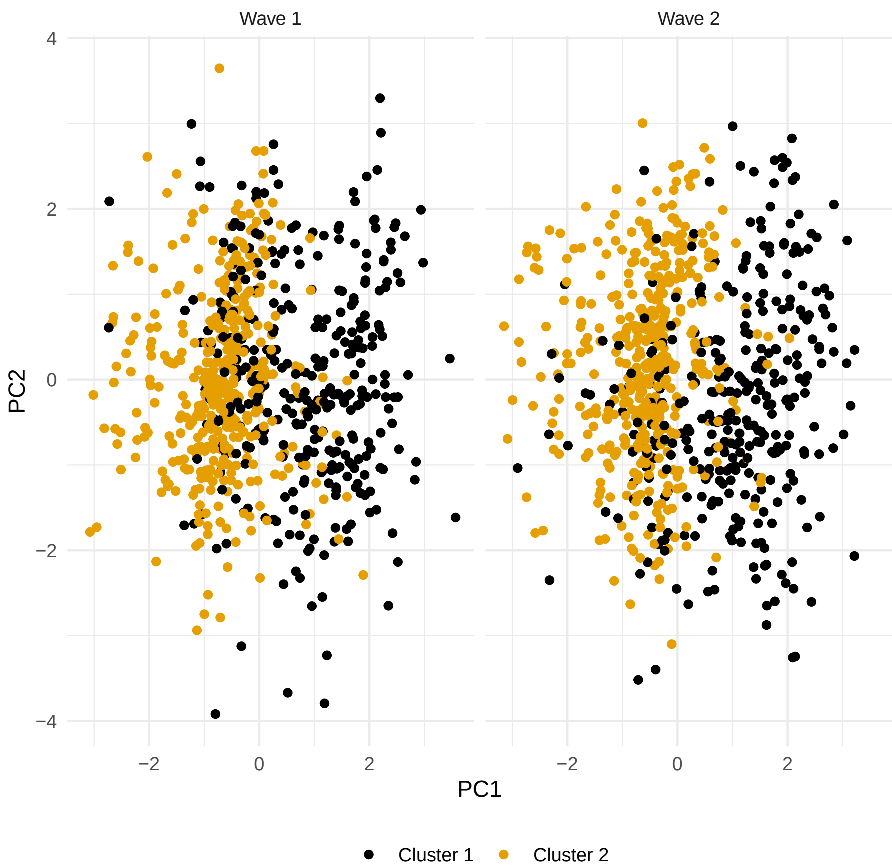

Figure 3.1: Visualization of clustering results using Principal Component Analysis. For Wave 1, K-medoid clustering was used; based on these results, K-nearest Neighborhood was applied to predict the clusters for Wave 2.

An an alternative way to analyse this is to run the clustering algorithm only the first wave, and predict cluster classification for the second wave using the results from the first. If the clustering is meaningful, we should still be able to find the patterns on the background variables shown in Table 3.4. Additionally, we can check whether the proportion of the two groups changed from the previous findings.

To predict cluster classification for the second wave, I run K-nearest Neighbours on the first clustering result, again only using the measured preferences. The idea is that for Wave 2, cluster assignment is based on the closest observations in Wave 1, defined in Euclidean space. For $k = 1$, only the closest neighbour would be taken into account; increasing k may improve the assignment mechanism up to a certain point.

To find the optimal k, I tested the accuracy of the assignments by dividing the data of Wave 1 into a training and testing sets, and tried to find the optimal number of neighbours to best predict cluster belonging for the test data. I chose $k = 3$, however there was not a great difference in accuracy with other parameters.

To visualise the assignment, Figure 3.1 shows the resulting classifications; the dimensions correspond to the first two factors extracted from the Principal Component Analysis of the data, a popular way to plot multiple dimensions into a single plot. The general pattern seems similar in the two databases, while the classification seems to be fuzzier in the second wave (which might be the artefact of both the prediction and the slightly changed preferences observed). [13]

Table 3.5: Statistical tests for K-nearest neighbor prediction for Wave 2 cluster placement

	Clustering (Wave 1)			Knn Prediction (Wave 2)		
Variable	Cluster 1, N = 342	Cluster 2, N = 400	p-value	Cluster 1, N = 333	Cluster 2, N = 439	p-value
Age	53	48	<0.001	53	50	0.013
	(17)	(16)		(17)	(16)	
Household Size	2.52	2.58	0.7	2.75	2.71	0.6
	(1.24)	(1.31)		(1.57)	(1.27)	
Net Wage	150,497	210,827	<0.001	184,206	197,128	0.059
	(106,394)	(265,333)		(169,711)	(149,797)	
Unknown Wage	114	155		102	133	
Education Level:			<0.001			0.006
Less than secondary school	150 (44%)	120 (30%)		123 (37%)	127 (29%)	
Secondary school	122 (36%)	182 (46%)		145 (44%)	186 (42%)	
University Degree	70 (20%)	98 (25%)		65 (20%)	126 (29%)	
Gender:			0.3			0.6
Female	182 (53%)	197 (49%)		179 (54%)	228 (52%)	
Male	160 (47%)	203 (51%)		154 (46%)	211 (48%)	
Health concerns	2.87	2.94	0.14	2.81	2.84	0.5
	(0.75)	(0.57)		(0.73)	(0.69)	
Family Health Concerns	3.03	2.96	>0.9	2.91	2.79	0.3
	(1.03)	(0.71)		(1.09)	(0.87)	
Financial Concerns	2.64	2.67	0.9	2.50	2.61	0.062
	(1.00)	(0.93)		(0.83)	(0.87)	

Table 3.5 shows the differences for the background variables. Similarly to Table 3.4, Age, and Education levels are statistically different in both cases, while the distributions of Gender is not significant for Wave 2 as found previously. Additionally, Household size continued to be insignificant. One thing to not that the predicted clusters for Wave 2 seem

[13] To account for the data sensitivity of the Principal Component Analysis, I ran the PCA only on data from the first wave; then, I projected the data from the second wave to the factor loadings, so the same values would correspond to the same coordinates in the graph.

to be more even (378 vs 394) compared to the k-medoid clustering results of Table 3.4 (284 vs. 488).

Differences in net earnings are smaller for Wave 2 compared to the separately ran cluster analysis, with it being statistically different only at the 10% significance level (p = 0.059). Education levels are again higher for Cluster 2 compared to Cluster 1, while gender contribution is balanced; Cluster 2 is also relatively younger on average, similarly to the results shown in Table 3.4. Worth noting that household-sizes while not statistically significant, but show similar values to those shown is the previous section. Also a consistent result is that perceived financial exposure due to Covid-19 seems to be higher for Cluster 2 (p = 0.062), despite the fact that these people had higher net wages on average.

Overall, we find that projecting the clusters found for Wave 1 on Wave 2 gives consistent results to the ones analysed in Section 3.4.2, related to the background variables.

3.4.3 Predicting factors for cluster groups

We observed in the previous sections that a cluster based on economic preferences consistently emerges, consisting of individuals who are relatively younger, earn more, and are more educated compared to the other cluster. In terms of preferences, this group includes people with higher levels of cooperation, greater competitiveness, less time-consistency, and a higher discount factor; they are also more prone to an internal locus of control. So far, we have only analysed these differences using t-tests; additionally, we can run regressions to identify which background variables are the most important factors in predicting whether an observation belongs to this cluster.

Table 3.6 shows regression results for the grouping with the higher outcome variables – Cluster 2 – separately for wave 1 and wave 2. Due to these clusters being defined by economic preferences, I did not include economic preferences in the regression. As it is expected from Table 3.4, the key regressors are age, education and wage - with place of living, marital status and gender being insignificant. [14] This findings are consistent with Chowdhury et al., 2022, where socio-economic status was found to be important for those who are more patient, risk-tolerant and pro-social. However, it is noteworthy that the R-squared is relatively low.

[14] With the additional note that gender was significant at the 10% level for the second wave, considering the OLS specification.

Table 3.6: Predictions on Cluster 2

Dependent Variable:	Cluster 2			
Model:	Wave 1 OLS	Wave 2 OLS	Wave 1 Logit	Wave 2 Logit
Variables				
Constant	-0.6535	-0.1034	-5.018**	-2.805
	(0.4529)	(0.4153)	(2.018)	(1.906)
Female	-0.0121	-0.0612*	-0.0526	-0.2722*
	(0.0367)	(0.0349)	(0.1567)	(0.1592)
Age	-0.0036***	-0.0040***	-0.0151***	-0.0178***
	(0.0011)	(0.0011)	(0.0048)	(0.0049)
Single	0.0189	-0.0110	0.0826	-0.0387
	(0.0376)	(0.0365)	(0.1609)	(0.1663)
Capital City	-0.0649	-0.0198	-0.2723	-0.0886
	(0.0478)	(0.0449)	(0.2038)	(0.2056)
Sec. School Educ.	0.1250***	0.1018**	0.5226***	0.4399**
	(0.0421)	(0.0408)	(0.1789)	(0.1825)
Univ. Educ.	0.0910*	0.1487***	0.3715*	0.6649***
	(0.0517)	(0.0484)	(0.2194)	(0.2223)
ln(Net Wage)	0.1092***	0.0762**	0.4737***	0.3486**
	(0.0372)	(0.0336)	(0.1660)	(0.1550)
Wage missing	1.238***	0.7827**	5.363***	3.571**
	(0.4093)	(0.3733)	(1.825)	(1.714)
No Wage	1.113**	0.8962**	4.851**	4.081**
	(0.4525)	(0.4129)	(2.015)	(1.896)
Fit statistics				
Observations	742	772	742	772
Squared Correlation	0.05842	0.05919	0.05901	0.05867
Pseudo R^2	0.04164	0.04415	0.04337	0.04579
BIC	1,094.0	1,086.4	1,045.8	1,037.7

IID standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

3.5 Conclusions

In this paper, I analysed economic preferences as well as groups of economic preferences defined by cluster analysis, and whether they change over time. To my knowledge, I am the first who ran analyses on stability of clusters of economic preferences. For the analysis, I used a Hungarian survey conducted in two waves: during the summer of 2020 and in the last months of 2020. Analysing the panel subsample shown stability in the measurements, with the exception of trust and altruism, both increasing during the second wave.

To identify groups of economic preferences, I applied k-medoid clustering to the data. To check whether the clusters are internally cohesive and externally isolated, I analysed the differences for preferences, as well as for background variables for each wave. One of the clusters were found to be more cooperative, competitive, relatively less patient and having a more internal locus of control (although the difference was not significant for Wave 2) consistently for both waves. This group was also found to be relatively younger, more educated, and had a more balanced gender distribution. Members of this clusters were found to be earning more relative to the other group. This result is similar to Chowdhury et al., 2022, who found that more patient, risk-tolerant and pro-social groups were better off financially.

As an alternative, I ran the clustering algorithm on only the first wave, and projected the cluster belonging to the sample of wave 2 using K-Nearest Neighbours, and only employing the economic preference variables. Analysing the background variables, a relatively younger, higher earning and more educated cluster emerged. Thus, with both methods I found that grouping solely on economic preferences, individuals with relatively higher socio-economic status could be identified.

As a robustness test, I ran the clustering using hierarchical clustering methods as well, as it is widely used for exploratory analysis. However, the results were inconclusive, as hierarchical clustering tends to pick outliers as a separate cluster. Similarly, I ran the analysis using the k-prototype partitioning method as well to account for the nature of the data. Here, while the differences were more nuanced than in the main analysis, the differences between clusters were still consistent — although not always statistically significant — between the two waves. It is worth noting that running the analysis with other alternative methods could be beneficial in future research.

One limitation of the study is the survey format, as the measurement of certain preferences are only asked through survey responses rather than through economic games (such is the case with trust, as mentioned in Section 3.4.1. Another limitation is that only a subsample of respondents were asked repeatedly: doing the analysis on a balanced panel data would be more ideal. Nonetheless, the dataset is still adequate to analyse whether similar clusters emerge in the population, and whether these groups have any indications for the background variables available.

It is also important to stress that the methodology is not capable of causal analysis – rather, we are able to make informed guesses about to connection between interconnected economic preferences and observed outcome variables. Finding out whether these preferences have causal effects to economic outcomes might have important policy implications. Finding out, for example, whether students with certain economic personality types have a higher chance of dropping out from school might lead to a more targeted education system.

Chapter 4

Heterogeneity of Economic Expectations

Dissecting the Role of Socioeconomic Status[1]

4.1 Introduction

Important economic decisions such as consumption, saving, and investment, are shaped by individuals' expectations regarding future macroeconomic conditions. A growing body of literature indicates a substantial heterogeneity in these expectations, which are closely associated with socio-demographic characteristics. For example, Dominitz and Manski, 2004 analyze the Michigan Index of Consumer Sentiment, and find that macroeconomic expectations correlate negatively with age, with males tending to be more optimistic, and higher levels of education being associated with more positive expectations. In the same vein, Das et al., 2020 report significant correlations between socioeconomic status (SES) and macroeconomic expectations, including economic outlook, business conditions, unemployment, and stock returns. Notably, higher income or higher education levels are generally associated with more favorable expectations. Similar patterns have been observed regarding inflation expectations where findings show that females, individuals with lower levels of education, and those with lower income tend to hold consistently higher inflation expectations (Angelico and Di Giacomo, 2019; Blanchflower and Mac-

[1] This chapter is a joint work with Hubert János Kiss.

Coille, 2009; Bruine de Bruin et al., 2010; D'Acunto, Malmendier, and Weber, 2021; Lombardelli, 2003).

Furthermore, differences in macroeconomic expectations contribute to disparities in investment and consumption patterns among individuals with different socioeconomic statuses, even after accounting for socio-demographic characteristics. Positive macroeconomic expectations are associated with a greater propensity to contemplate purchasing homes, durable goods, or cars (Carroll and Dunn, 1997; Das et al., 2020; Hanspal et al., 2021; Roth and Wohlfart, 2020). Higher inflation expectations often prompt individuals to advance their consumption (Bachmann et al., 2015; D'Acunto, Hoang, and Weber, 2022; D'Acunto, Hoang, et al., 2019), increase their expenditure on durable goods (D'Acunto et al., 2016; D'Acunto et al., 2018), and save less (Vellekoop and Wiederholt, 2019). These patterns highlight the impact of macroeconomic expectations on decision-making and the potential consequences for economic outcomes among different socioeconomic groups.

In this study, we build upon previous findings of the literature about the association between macroeconomic expectations and SES in three ways. First, in the study closest to ours, Das et al., 2020 find a sizable and persistent difference in macroeconomic expectations between individuals in the lowest and highest quintiles of the income distribution, as well as between those with and without a university degree.[2] Our analysis seeks to provide a more nuanced approach by examining all income quintiles to see if differences in the association between the quintiles and expectations are similar or uneven. In their regression analysis, Das et al., 2020 assume a linear relationship between income quintiles and economic expectations. However, our descriptive analysis indicates a possible non-linear association. In order to analyze this, and to conduct our analysis without the imposing of linear connection, we utilize quintile dummies to account for potential non-linearities. By allowing for a non-linear connection between SES and economic expectations, we are also able to conduct a more detailed analysis on heterogeneity of these effects, which might have important implications for policymakers to manage expectations in a more targeted way.

Regarding education, contrary to the binary distinction (university degree vs. no university degree) in Das et al., 2020, we provide a more detailed investigation by considering three education levels: individuals without a high-school degree, those with a high-school degree, and those with a university degree. Our aim is to find out where exactly on the education ladder the differences in expectations materialize. Second, while most previous

[2]Similarly, Bruine de Bruin et al., 2010 use a simple distinction based on the median split to investigate the relationship between inflation expectation and income/education, reporting a negative association.

studies, including Das et al., 2020, primarily focus on expectations at the macroeconomic level, we extend our analysis to thoroughly examine household-level expectations. Third, we examine the role of two factors identified in the literature through which SES may influence expectations: personal experience and optimism. We use respondents' assessment of their own household's financial situation over the previous 12 months as a proxy for personal experience. Additionally, we utilize household-level expectations as a proxy to capture optimism, which represents another potential factor underlying the relationship between SES and macroeconomic expectations.

Apart from gaining a deeper understanding of the relationship between socioeconomic status and macroeconomic expectations, in line with the existing literature, we also investigate whether these expectations influence economic decisions. Specifically, we examine the role of these expectations in shaping the intention to purchase durable goods such as homes and cars, as well as the decision to spend a substantial amount of money on home improvement.

In line with Das et al., 2020, we find that macroeconomic expectations differ significantly between the top and the bottom income quintiles, and also between individuals without a high-school degree and those with a university degree. In addition, we document important non-linearities. Regarding income quintiles, individuals in the upper (that is, fourth and top) quintiles hold significantly more positive macroeconomic expectations than those in the lower quintiles. The bottom two quintiles - and in some cases, the bottom three - however, are not significantly different from each other. Additionally, there is a significant difference in expectations between individuals in the fourth and the top quintiles. Imposing linearity yields that income quintile has a significant and positive coefficient, similarly to Das et al., 2020, but our analysis reveals that the picture is more nuanced, with no obvious differences in the lower quintiles, but clear disparities at the higher and lower end. This nuanced understanding is further validated by additional analysis, specifically running regressions on a more granular, decile-by-decile level.

Turning to education levels, we offer a more detailed analysis than Das et al., 2020 by considering three categories: individuals with less than a high-school degree, those with a high-school degree, and those with a university degree, as opposed to their binary classification of without/with a university degree. We find significant differences in economic expectations between those without and with a high-school degree in several cases, indicating that this distinction matters. Individuals with a high-school degree hold significantly more positive economic expectations than their counterparts without it. Differences in macroeconomic expectations between individuals with a high-school degree and with a university degree only materialize in inflation expectations, suggesting that disparities in macroeconomic expectations are more pronounced at the lower end of the educational spectrum.

Our study significantly contributes to the literature by offering a more comprehensive analysis of the link between socio-economic status and expectations. Since the introduction of the Phillips curve, economists have recognized the impact of expectations on various economic indicators, such as inflation. Consequently, understanding the heterogeneous effects of expectations on consumer sentiment holds substantial importance for policymakers. Furthermore, we broaden the scope of our analysis by incorporating household-level expectations. When considering income quintiles, we observe a similar pattern to macroeconomic expectations, but the differences in household-level expectations (especially, for saving expectations) seem to be more noticeable. In terms of education levels, there are clear differences in household-level expectations between education levels: higher levels of education are associated with more positive expectations.

In addition, we investigate the role of two factors that have been identified in the literature as potential determinants of macroeconomic expectations: personal experiences (specifically, experiences of recessions and events in the past year) and optimism (proxied by household-level expectations). Consistent with findings in Das et al., 2020, we observe that during recessions, the gaps in macroeconomic expectations decrease when considering education levels. However, we do not find a similar pattern when analyzing income quintiles. Macroeconomic expectations are positively associated with experiences during the past year and also with household-level expectations. Furthermore, when examining household-level expectations, we find that during recessions the differences in expectations diminish when considering income quintiles (but not when investigating education levels). Experiences from the past year have a significant and positive influence on household-level expectations.

Finally, our findings indicate a robust association between macroeconomic and household-level expectations and economic decisions, such as the intention to purchase a home or a car, as well as the intent to spend on home improvement. Importantly, these associations remain significant in most cases even after controlling for socioeconomic variables, suggesting that these expectations play a crucial role beyond their socioeconomic determinants. Therefore, the heterogeneity in expectations is relevant, because low-SES households make different choices compared to their high-SES counterparts. As a consequence, economic policy should take into account the heterogeneity of expectations and how those expectations shape economic decisions across the socioeconomic spectrum.

The remainder of the study is organized as follows. In section 4.2, we review the existing literature on expectations and their connection with SES, and summarize the mechanisms through which SES can affect macroeconomic and household-level expectations. In section 4.3, we present the data used for our analysis. Section 4.4 contains the results, and section 4.5 concludes.

4.2 Literature Review

In this section, first, we review the most relevant literature on how SES is associated with expectations and how those expectations shape economic decisions. Second, we briefly summarize the mechanisms behind the association between SES and expectations.

There is a growing body of literature documenting a significant relationship between SES and inflation expectations. Individuals with lower levels of education and income tend to have higher inflation expectations as supported by data from the UK (Blanchflower and MacCoille, 2009; Lombardelli, 2003), the US (Angelico and Di Giacomo, 2019; Bruine de Bruin et al., 2010; Bryan, Venkatu, et al., 2001), the European Union (D'Acunto, Malmendier, and Weber, 2022), or South Africa (Reid et al., 2021). [3] Moreover, individuals tend to act upon their inflation expectations. Higher inflation expectations predict higher current consumption (Bachmann et al., 2015; C. C. Binder and Brunet, 2022; Burke and Ozdagli, 2014; D'Acunto et al., 2021; Dräger and Nghiem, 2021; Ichiue and Nishiguchi, 2015). However, this relationship often holds only for specific subsets of individuals. Specifically, the link between inflation expectations and consumer spending is stronger for individuals with more accurate expectations (Bachmann et al., 2015), better cognitive abilities (D'Acunto et al., 2021), more assets (Ichiue and Nishiguchi, 2015), higher education (Burke and Ozdagli, 2014), more income (Coibion et al., 2022). Similarly, inflation expectations often correlate with savings (Arnold et al., 2014; D'Acunto, Malmendier, Ospina, and Weber, 2019; Premik and Stanisławska, 2017; Vellekoop and Wiederholt, 2019): higher inflation expectations are associated with lower levels of savings. In addition, individuals with higher inflation expectations tend to choose fixed-rate mortgage contracts over adjustable-rate ones (Botsch, Malmendier, et al., 2020). Experimental evidence (Armantier et al., 2015) also supports the notion that individuals act upon their inflation expectations, although this relationship does not hold for individuals with lower levels of education.

[3] The only exception is Jonung, 1981 which uses Swedish data and finds that individuals with higher income have higher inflation expectations (in an economy that experienced high inflation at the time).

Interestingly, there is a limited amount of literature available on the relationship between macroeconomic expectations (other than inflation) and SES. The study closest to ours is Das et al., 2020 that uses data from the Michigan Survey of Consumers from 1978 to 2014 with about 400 respondents each month to investigate how income rank and having a university degree are associated with different forms of macroeconomic expectations, including the probability of stock market gain, business conditions in the next 12 months or 5 years, unemployment. Through OLS regressions the study shows that both income rank and a university degree are highly significant predictors of all the studied expectations, even after controlling for factors such as age, gender, marital status, and recession. Additionally, instrumental variable regressions reveal that the optimism captured by the expectations is positively and significantly associated with household choices, such as investment decisions and intentions to purchase a home, durable goods, or a car, even when considering income rank and having a university degree. In line with Das et al., 2020, Dominitz and Manski, 2004 also document that respondents with higher education tend to have more positive expectations about the economic outlook. However, a related study by Roth and Wohlfart, 2020 does not find a significant association between recession expectations and education/income.

Macroeconomic expectations may be intricately related. According to the Euler equation, higher inflation expectations should lead to increased current spending. However, higher inflation expectations may make individuals more pessimistic about the overall economic outlook and their future income that, in turn, may result in precautionary savings and reduced current consumption, as shown in Coibion et al., 2019. This finding suggests that it is advisable to study macroeconomic expectations together (rather than solely focusing, for instance, on inflation expectations), as we do in this study.

We turn now to review the main mechanisms behind the relationship between SES and macroeconomic expectations. First, SES can be related to economic or financial optimism, which in turn may be associated with macroeconomic expectations. Evidence is provided by Brown and Taylor, 2006 who report a positive correlation between education and financial optimism, assessed through the question 'Looking ahead, how do you think you will be financially a year from now?'. This is an individual-level assessment that according to the authors synthesizes elements of individual factors (e.g. salary, job prospects) and also elements of a broader economic outlook, demonstrating the intertwined nature of economic/financial optimism and macroeconomic expectations. Experimental evidence also supports this mechanism. Studies by Kuhnen and Miu, 2017 and Das et al., 2020 indicate that individuals from lower socioeconomic backgrounds tend to exhibit more pessimism regarding the payoff distribution of risky assets. In this study, we proxy optimism by household-level expectations concerning the economic situation within the next 12 months.

Second, systematic differences in personal experiences and characteristics also contribute to the link between SES and the heterogeneity of macroeconomic expectations. Past experiences about unemployment, changes in net worth, or prices paid in the grocery store may shape macroeconomic expectations (D'Acunto, Malmendier, Ospina, and Weber, 2021; Kuchler and Zafar, 2019; Malmendier and Nagel, 2011, 2016). Additionally, personal characteristics including economic preferences, financial literacy, and the length of one's financial planning horizon can also influence macroeconomic expectations (Li and Huang, 2020; Lusardi and Mitchell, 2011; Van Rooij et al., 2012; Zikmund-Fisher and Parker, 1999). [4] As a proxy to account for past experiences, our data include self-assessments of changes in the household's economic situation in the past year. By using these data we can (at least partially) take into account personal experiences.

If, after accounting for household-level optimism and/or personal experiences the relationship between SES and macroeconomic expectations weakens or vanishes, it suggests that the related factor is behind the association.

Similarly to optimism and personal experiences, there may be other omitted variables that act as confounders in the relationship between SES and macroeconomic expectations. IQ may be such a confounder as it correlates with a host of factors such as financial decision-making (Agarwal and Mazumder, 2013; Grinblatt et al., 2011, 2012, 2016), or economic preferences (Burks et al., 2009; Falk et al., 2018) that are related to both SES and macroeconomic expectations. Moreover, IQ is directly associated to educational attainment (Herrnstein and Murray, 2010; Neisser et al., 1996), SES (Hackman and Farah, 2009; Larson et al., 2015) and expectations (D'Acunto et al., 2021). Exogenous shocks, such as the COVID-19 pandemic or recessions may also impact households of different SES differently. Furthermore, these shocks may also affect expectations in a diverse manners (C. Binder, 2020; Das et al., 2020). Overall, it is important to acknowledge that omitted variables remain a challenge. To the extent that these omitted variables correlate with optimism and personal experiences, we control for those omitted variables. It also implies that through the correlations, personal experience pick up the effect of the omitted variables.

[4] Similarly, media consumption may play some role in expectation formation, and if individuals with different SES have distinct news consumption habits, it could lead to heterogeneous expectations. However, the literature generally finds no (Coibion et al., 2020) or only a small effect 8Dräger, 2015), so media consumption is less likely to be the prime driver of divergent inflation expectations according to SES.

4.3 Data

In our study, we utilize survey data obtained from GKI Economic Research Co. GKI has been conducting monthly household surveys since 1993, employing EU methodology to analyze the economic expectations of the Hungarian population (GKI, 2022). The database contains pooled cross-sectional data. GKI provided monthly observations from June 2000 until the end of 2009.

The survey contains information on the expectations regarding three macroeconomic variables, as outlined in Table 4.2. The first variable is about the expected evolution of the general macroeconomic outlook of the country in the next 12 months (referred to as ECON-macro). The second variable concerns inflation expectations for the upcoming 12 months (denoted as INF). The third variable captures expectations about the evolution of unemployment over the next 12 months (referred to as UNEMP). In the survey, responses were coded on a scale ranging from -2 to +2, where -2 corresponds to ”it will be much worse”, 0 represents ”will remain approximately the same”, and +2 indicates ”it will improve significantly”.

The survey also includes household-level expectations. Respondents provide their expectations regarding the economic situation of their household in the next 12 months (referred to as ECON-hh). The survey also queries respondents about their household's ability to save during the upcoming 12 months (denoted as SAV). Furthermore, there is a question regarding the household's ability to purchase durable goods in the following year (referred to as DUR). For these questions, respondents were presented with various response options, including: 'will improve considerably' (+2), 'will improve somewhat' (+1), 'no change expected' (0), 'will worsen somewhat' (-1), and 'will worsen considerably' (-2).

Apart from the previous items, the survey also captures respondents' purchase intentions. Therefore, we know whether the household intends to purchase a car or a home, as well as whether they plan to make significant expenditures on their house (denoted as CAR/HOME/HOME-exp, respectively). Finally, the subjects were asked whether it is worth buying durables at the time of the question asked (DUR-worth). When inquiring about intentions, the available options were 'for certain' (+2), 'probably' (+1), 'probably not' (-1), and 'certainly not' (-2). The option of zero (0) was excluded from the choices by the pollster.

Similarly to Das et al., 2020, we calculate our own macroeconomic expectation index. We create an index for macroeconomic expectations (referred to as OPT-macro) by taking the average of the expectations regarding the change in the general economic outlook, unemployment, and inflation levels. Hence, OPT-macro = (ECON-macro + INF + UN-EMP)/3. We also compute a household-level expectation index (referred to as OPT-hh) based on the expectations concerning the household's economic prospect in the next 12 months (denoted as ECON-hh), the household's perceived ability to save in the upcoming 12 months (referred to as SAV), and the household's ability to purchase durables (denoted as DUR). Hence, OPT-hh = (ECON-hh + SAV + DUR) / 3.

Table 4.1: Summary statistics of key variables

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Age	75,713	46.285	17.132	17	32	46	59	99
Household Income	75,716	114,667	115,361	0	66,000	100,000	150,000	12,000,000
Quintile 1	14,647	32,092	23,042	0	1,000	38,000	50,000	92,000
Quintile 2	15,014	74,185	13,592	35,000	65,000	75,000	84,000	110,000
Quintile 3	15,187	101,406	19,067	56,000	90,000	100,000	116,000	150,000
Quintile 4	15,382	135,514	27,991	75,000	120,000	140,000	150,000	210,000
Quintile 5	15,486	224,313	206,371	100,000	170,000	200,000	250,000	12,000,000
Has university degree	8,437	12.1%						
Has High-School Degree	8,135	11.7%						
Less Than High School	53,126	76.2%						
ECON-macro	75,716	- 0.347	1.083	- 2	- 1	0	1	2
INF	75,716	- 1.167	0.745	- 2	- 2	- 1	- 1	2
UNEMP	75,716	- 0.797	0.946	- 2	- 2	- 1	0	2
ECON-hh	75,716	- 0.414	1.022	- 2	- 1	0	0	2
SAV	75,716	- 1.030	1.132	- 2	- 2	- 1	- 1	2
DUR	75,716	- 0.687	0.972	- 2	- 2	- 1	0	2
OPT-macro	75,716	- 0.770	0.722	- 2.000	- 1.333	- 0.667	- 0.333	2.000
OPT-hh	75,716	- 0.710	0.835	- 2.000	- 1.333	- 0.667	- 0.333	2.000
HH-Prev.Year	75,622	- 0.617	0.972	- 2	- 1	- 1	0	2
CAR	27,772	- 1.635	0.833	- 2	- 2	- 2	- 2	2
HOME	27,798	- 1.744	0.726	- 2	- 2	- 2	- 2	2
DUR.worth	72,346	- 0.857	0.355	- 1	- 1	- 1	- 1	1
HOME-exp	27,724	- 1.315	1.145	- 2	- 2	- 2	- 1	2

OPT-macro = (ECON-macro + INF + UNEMP)/3
OPT-hh = (ECON-hh + SAV + DUR) / 3

Table 4.2: Key questions used in the analysis

Variable	Range of answers	Wording of the question
Economic outlook of the country (ECON-macro)	(-2) to $(+2)$	In your opinion, how will the country's economic situation evolve over the next 12 months?
Inflation (INF)	(-2) to $(+2)$	In your opinion, how will inflation evolve over the next 12 months?
Unemployment (UNEMP)	(-2) to $(+2)$	In your opinion, how will unemployment evolve over the next 12 months?
Economic outlook of the household (ECON-hh)	(-2) to $(+2)$	In your opinion, how will your household's financial situation evolve over the next 12 months?
Economic change of the household (HH-Prev.Year)	(-2) to $(+2)$	How did the economic situation of your household change in the last 12 months?
Ability to save (SAV)	(-2) to $(+2)$	In your opinion, how will your household's savings change over the next 12 months?
Ability to buy durables (DUR)	(-2) to $(+2)$	Do you think your household will be able to save enough to buy high-value consumer goods in the next 12 months?
Intention to buy a new car (CAR)	(-2) to $(+2)$ (excluding zero as an option)	How probable it is that your household will buy a new car in the next 12 months?
Intention to buy a new home (HOME)	(-2) to $(+2)$ (excluding zero as an option)	How probable it is that your household buys or builds a house or apartment in the next 12 months?
Worth to purchase durables (DUR-worth)	(-2) to $(+2)$ (excluding zero as an option)	Do you think it makes sense to buy high-value consumer goods (furniture, washing machine, TV, etc.) these days?
Intended expenditure on housing (HOME-exp)	(-2) to $(+2)$ (excluding zero as an option)	How probable it is that your household spends more on your house or apartment in the next year or two?

The range of answers are coded from -2 to $+2$, with -2 meaning ”it will become a lot worse” and $+2$ meaning ”it will become much better” compared to last year. As such, a general rule for the analysis is the higher the value, the ”better” the expectation (for example: $+2$ of INF and UNEMP indicates that inflation will be much better (”lower”) compared to last year).

Following Das et al., 2020, throughout the analysis, we use household income levels and age (with roughly ten-year groups; between 18-30, 31-40, 41-50, 51-65, and 65 and above) to define income ranks for each month. [5] Descriptive statistics for the variables of interest are presented in Table 4.1. Note that for nearly all expectation-related questions, both the mean and the median values are negative, with a right-skewed distribution. The median is zero only in two cases: ECON-macro and ECON-hh. Overall, respondents exhibited a general pessimism regarding the future. This negative outlook is also evident in their intention to purchase durable goods, particularly when it comes to buying a car or a house. However, it should be noted that Table 4.1 provides pooled data spanning the whole period under consideration. Thus, it does not allow us to discern whether specific periods were characterized by generalized optimism or economic gloom. To address the external validity of these qualitative findings, we conducted an analysis using actual data, the details of which can be found in Appendix 3.

Regarding the SES variables, the increases between the lower quintiles appear to be of approximately the same magnitude, while a larger jump is observed when transitioning from the fourth to the fifth quintile. [6]

The share of respondents without high-school degree seems to be high, but it aligns with official statistics. In 2001 / 2011 (the two census years around our data range), the proportion of the population without a high-school degree was 67.5% / 56.9%. 20.5% / 25% of the population had at most a high-school degree, respectively. The share of those with a university degree was 12% / 18.1% in 2001 / 2011. Overall, our sample slightly over-represents individuals with lower educational attainment.

[5] In some cases, where an individual's income was greater than their indicated family income, observations were filtered out. If the family income was zero, but the individual's income was non-zero, we imputed that value as the family income.

[6] Note that since we have income data spanning 10 years and quintiles are formed based on each month, there may be instances where the upper percentiles in a lower income quintile are larger than the lower percentiles in an upper quintile. Therefore, there are overlaps between the income distributions of adjacent quintiles.

4.4 Results

4.4.1 Descriptive statistics

As a first step, we examine the correlations between our main variables in Table 4.3.[7] As expected, there is a positive correlation between income and holding a university degree. However, the correlation between income and expectations is relatively weak. The association between having a university degree and expectations is larger, but generally below 0.1, indicating a modest relationship. Furthermore, age does not exhibit a strong correlation with expectations. The negative sign suggests that higher age is associated with more pessimistic expectations. The correlation between economic expectations is positive and of considerable magnitude. We document the highest associations between OPT-macro and OPT-hh (in both the limited and full data it is approximately 0.6). This finding is consistent with the results reported in Dominitz and Manski, 2004 which also reports a strong correlation between macroeconomic and household-level expectations. It suggests that these two types of expectations are intertwined and difficult to separate.

As for consumption decisions, variables such as CAR, HOME, HOME-exp, and DUR-worth exhibit higher correlations with OPT-hh compared to OPT-macro (see Table 4.21 in Appendix 2), suggesting that while macroeconomic expectations are important, household-level expectations tend to have an even greater influence on these decisions. Finally, the self-assessed change in the household's economic situation in the last year (HH-Prev.Year) displays a high correlation with both macroeconomic and household-level expectations, as well as with consumption decisions. This indicates that individuals' perceptions of their own economic situation in the past year strongly relate to their expectations and subsequent consumption choices.

To see how expectations evolve over time, we plot the monthly average values of macroeconomic and household-level expectations, by quintiles based on household income. As in Das et al., 2020, quintiles are defined within year-age groups. However, while Das et al., 2020 focus solely on the top and bottom income quintiles, we also include the middle quintile to get a first impression of whether the relationship between the income rank and expectations is gradual.

In Figure 4.1, macroeconomic expectations are presented on the left, while household-level expectations are shown on the right. Shaded areas represent periods of recession,

[7] Note that Table 4.3 does not contain the variables for which we have considerably fewer observations (CAR, HOME and HOME-exp, see Table 4.1). For a comprehensive view including all variables of interest, please refer to Appendix 2.

Table 4.3: Correlation table for macroeconomic expectations and other relevant variables

	Income	Age	ECON-macro	INFL	UNEMP	ECON-hh	Hh.Prev.Year	SAV	DUR	OPT-macro	OPT-hh	University degree
Income	1											
Age	-0.024	1										
ECON-macro	0.004	-0.039	1									
INFL	0.041	0.013	0.352	1								
UNEMP	0.00001	0.012	0.469	0.381	1							
ECON-hh	0.018	-0.104	0.644	0.328	0.410	1						
HH.Prev.Year	0.065	-0.060	0.456	0.255	0.340	0.566	1					
SAV	0.102	-0.103	0.384	0.247	0.308	0.434	0.434	1				
DUR	0.049	-0.092	0.450	0.289	0.363	0.504	0.470	0.457	1			
OPT-macro	0.016	-0.010	0.826	0.686	0.803	0.614	0.464	0.412	0.483	1		
OPT-hh	0.073	-0.125	0.610	0.358	0.447	0.799	0.609	0.806	0.800	0.624	1	
University degree	0.187	0.010	0.051	0.070	0.063	0.057	0.088	0.150	0.083	0.077	0.123	1

OPT-macro = (ECON-macro + INF + UNEMP) / 3
OPT-hh = (ECON-hh + SAV + DUR) / 3

defined by two consecutive quarters of GDP decrease. Consistent with Das et al., 2020, differences between the top and bottom quintiles are clearly evident in most instances, with the former displaying greater optimism than the latter. When it comes to macroeconomic expectations, the disparity between these two groups is most notable in terms of inflation expectations. [8] At the household level, disparities are more pronounced, especially in savings expectations and the intention to purchase durable goods.

The picture becomes much less clear when we consider the middle income quintile. In some cases, the expectations of respondents in the middle quintile are clearly positioned between those of the top and bottom quintiles. However, expectations are often jumbled and difficult to distinguish. Generally, the middle quintile tends to be closer to the bottom quintile rather than the top one, suggesting that differences in expectations between income quintiles are not linear. Moreover, in certain instances (such as the case of inflation during 2007-2009), respondents belonging to the middle quintile seem to have even lower macroeconomic expectations than those in the bottom quintile.

In line with Das et al., 2020, differences in expectations tend to diminish and often disappear during recessions, which is clearly visible in Figure 4.1 during the Great Recession (and also when the austerity package was introduced in 2006). The only exception is savings expectations where recessions do not seem to cause as much turmoil as in the case of other expectations.

Figure 4.2 shows the analysis for the same variables, by education level. There is a clear distinction between the lower and higher ends, that is, between people without a

[8] We observe a significant decline in all macroeconomic expectations in 2006, which can be attributed to an economic austerity package announced in June of that year.

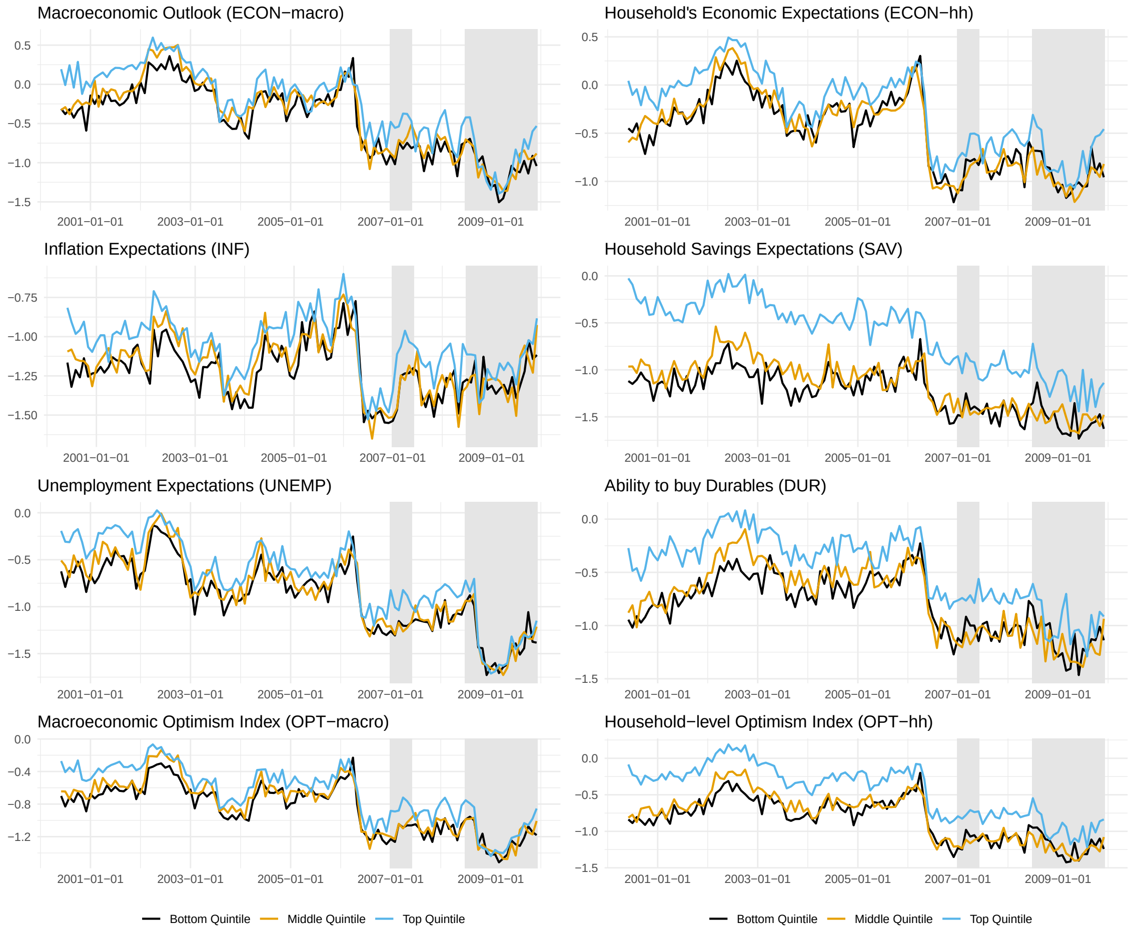

Figure 4.1: Average scores in the first, third and fifth income quintiles for macroeconomic (left) and household-level (right) expectations.

high-school degree and those with a university degree. This finding is consistent with Das et al., 2020, who only distinguish two categories by education level (those with and without a university degree). Expectations for people with high-school degree generally fluctuate between the two groups, as exemplified by inflation expectations. However, in general, the expectations of those with a high-school degree appear to be closer to those with a university degree. Similar to the case of income quintiles, differences in the expectations are clearly discernible in the case of household-level savings expectations. There seems to be an equal distance between the savings expectations of the different groups formed based on education level throughout the observation period. Similarly to what we have observed previously, differences in macroeconomic expectations diminish during recessions, particularly during the Great Recession. Household-level expectations based on education levels appear to be less sensitive to recessions. Note that similarities in expectations based on income quintiles and education level may stem from the strong correlation between the two factors.

4.4.2 Regression analysis

Non-linear associations

While the previous figures provide suggestive evidence of differences in expectations based on income rank and education, we now present a more formal and rigorous analysis. Table 4.4 displays the results of ordinary least squares (OLS) regressions, with the dependent variables being expectations at both the macroeconomic and household levels. Note that higher values in the table indicate a more optimistic expectation.

All regressions include quintile dummies (with the bottom quintile as the baseline) and education level dummies (with no high-school degree as the baseline). Therefore, while in the descriptive analysis income quintiles may have picked up the association between education and expectations (and vice versa), here we control for income and education as well. The use of dummies allows us to examine whether the relationship between the quintiles (or education levels) and expectations changes gradually as we move to higher quintiles (or education levels). All regressions include the following additional controls: dummies for year-month, age, gender, and marital status. Standard errors, shown in parentheses, are clustered at the individual level. The negative constants observed across the regressions reflect the predominantly pessimistic expectations, as already observed in Ta-

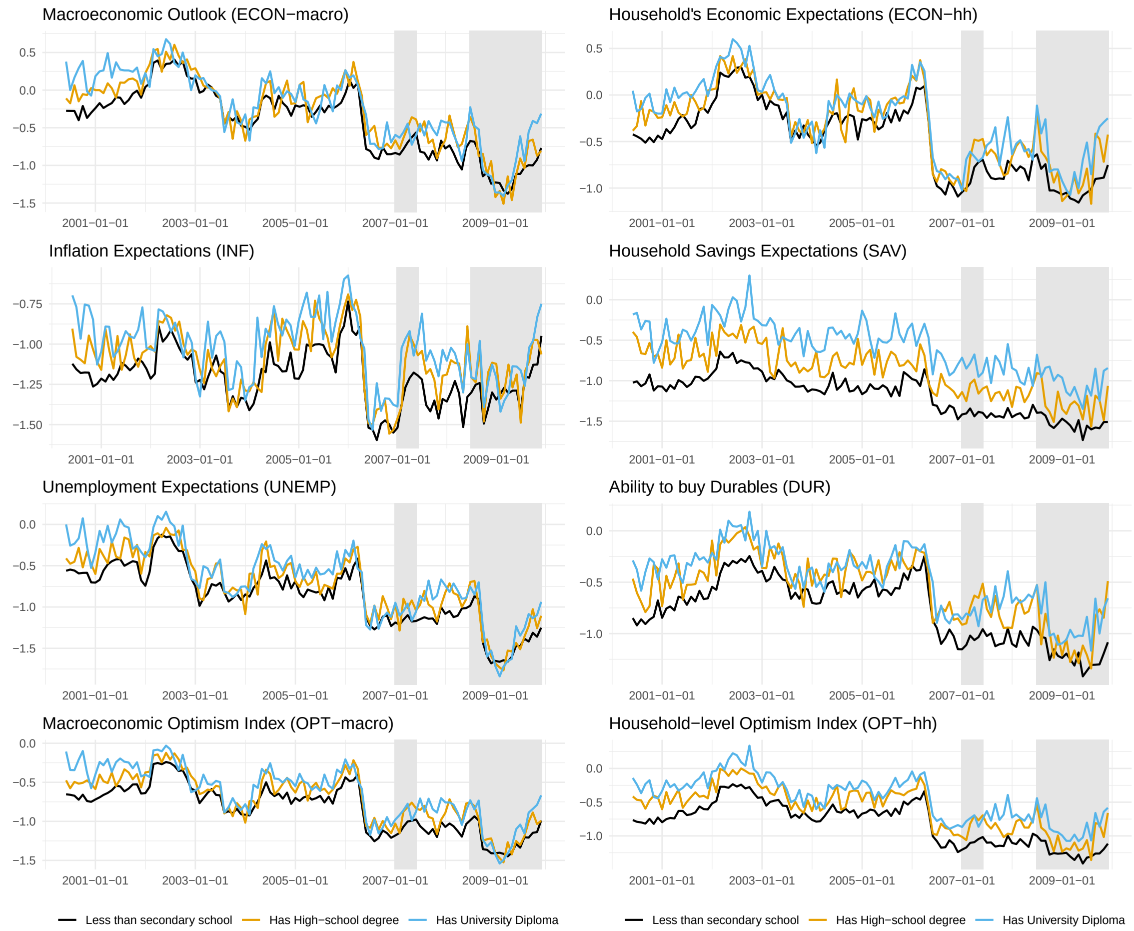

Figure 4.2: Average per education level for macroeconomic (left) and household-level expectations (right).

ble 4.1. In general, recessions tend to worsen expectations. However, this effect is not statistically significant when considering the economic outlook variables (ECON-macro and ECON-hh). Interestingly, during recessions, inflation expectations show a more optimistic trend. There was a noticeable shift towards greater optimism in 2007, following the significant decrease in general economic expectations in the latter half of 2006. [9] Another explanation could be that at the onset of a recession, expectations initially worsen, but as the shock subsides, people may become relatively more optimistic, giving rise to a ”the worst is over” sentiment.

Turning to the SES variables, when comparing the upper quintiles (3-5), we observe that respondents in these quintiles hold significantly more optimistic expectations com-

[9]The decline in 2006 is more likely to be attributed to political discontent against the government rather than actual macroeconomic foundations.

Table 4.4: Regression results for economic expectations based on separate quintiles

	Dependent variable:							
	ECON-macro	INF	UNEMP	OPT-macro	ECON-hh	SAV	DUR	OPT-hh
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Quintile 2	0.038***	0.0001	0.003	0.014*	0.030***	0.022*	0.028**	0.027***
	(0.012)	(0.009)	(0.010)	(0.008)	(0.011)	(0.013)	(0.011)	(0.009)
Quintile 3	0.090***	0.037***	0.045***	0.057***	0.067***	0.145***	0.091***	0.101***
	(0.012)	(0.009)	(0.010)	(0.008)	(0.011)	(0.013)	(0.011)	(0.009)
Quintile 4	0.157***	0.087***	0.105***	0.116***	0.139***	0.318***	0.185***	0.214***
	(0.012)	(0.009)	(0.010)	(0.008)	(0.011)	(0.013)	(0.011)	(0.009)
Quintile 5	0.220***	0.148***	0.193***	0.187***	0.232***	0.587***	0.328***	0.382***
	(0.012)	(0.009)	(0.011)	(0.008)	(0.011)	(0.013)	(0.011)	(0.009)
Has University Degree	0.122***	0.133***	0.142***	0.132***	0.133***	0.373***	0.173***	0.226***
	(0.012)	(0.009)	(0.011)	(0.008)	(0.011)	(0.013)	(0.011)	(0.009)
Has High-School Degree	0.101***	0.092***	0.113***	0.102***	0.074***	0.165***	0.137***	0.125***
	(0.012)	(0.009)	(0.010)	(0.008)	(0.011)	(0.013)	(0.011)	(0.009)
Recession	−0.007	0.326***	−0.200***	0.040	−0.026	−0.152**	−0.118*	−0.099**
	(0.066)	(0.048)	(0.058)	(0.043)	(0.063)	(0.071)	(0.061)	(0.050)
Constant	−0.929***	−1.469***	−1.312***	−1.237***	−0.697***	−1.571***	−1.042***	−1.103***
	(0.050)	(0.037)	(0.044)	(0.033)	(0.048)	(0.054)	(0.047)	(0.038)
Observations	75,713	75,713	75,713	75,713	75,713	75,713	75,713	75,713
R^2	0.169	0.066	0.161	0.193	0.162	0.124	0.124	0.195
Adjusted R^2	0.167	0.065	0.159	0.192	0.161	0.123	0.122	0.194
Residual Std. Error (df = 75588)	0.988	0.720	0.868	0.649	0.936	1.060	0.910	0.750
F Statistic (df = 124; 75588)	123.548***	43.262***	116.630***	145.980***	118.163***	86.452***	85.931***	147.601***

Standard errors in parentheses.
*/**/*** denotes significance at 1 / 5 / 10 % level.
All regressions include time-year dummies, age, gender, family status.
OPT-macro = (ECON-macro + INF + UNEMP)/3
OPT-hh = (ECON-hh + SAV + DUR) / 3

pared to respondents in the bottom quintile. Moreover, the coefficients in these quintiles show a clear upward trend, indicating that moving up a quintile is associated with increased optimism. For the second quintile, we also observe a positive deviation compared to the first quintile, however, in some cases, these differences are rather small, and even insignificant in the case of INF and UNEMP (and as a consequence, the coefficient of OPT-macro is only marginally significant).

For certain variables, there appears to be a linear relationship between income quintiles and the dependent variable. For instance, when considering ECON-macro, moving up a quintile from quintile 2 onwards is associated with an increase of approximately 0.06-0.07 points in optimistic views. However, for other variables, the changes between quintiles are more erratic. Taking UNEMP as an example, individuals in quintile 2 expect approximately the same levels of unemployment compared to those in the bottom quintile. However, the unemployment expectations of individuals in the top quintile is significantly better than those at the fourth quintile, doubling their differences compared to quintile 1. This is also the case with other variables as the difference relative to the bottom quintile becomes more pronounced in the upper quintiles. For instance, in the case of SAV or DUR, the coefficients of the top quintile are considerably larger than the coefficients of quintile 4, indicating a non-linear relationship between the income quintile and the dependent variable. This non-linearity may also arise from the right-skewed distribution of income, which results in income levels in the bottom quintiles being closer to each other (see Table 4.1). [10]).

To gain further insight into the differences between quintiles and education levels, we represent in Figure 4.3 the estimated coefficients for each quintile in Table 4.4 along with the corresponding 95% confidence intervals. Consistent with the findings in Table 4.4, we observe that expectations of individuals in quintile 2 are very close to bottom quintile in the case of macroeconomic expectations. It is notable that macroeconomic expectations tend to be increasing linearly with income quintiles (with the exception of UNEMP and quintile 5). As we move to the higher quintiles, the coefficients become significantly different from each other at the 1% level, as well as from the lower quintiles. Hence, at the upper end of the income distribution, higher quintiles are associated with significantly greater optimism. On the other hand, the increase tends to be non-linear in the case of ECON-hh, SAV and DUR (and OPT-hh, by extension). Note also that differences in expectations between quintiles 4 and 5 tend to be larger than the differences between the subsequent lower quintiles.

As a robustness check, we ran the regression using income decile dummies, which can be found in Appendix 8. This also confirmed the our results, with the additional information of the jump at the top quintile is not restricted to the top 10 percent of the income-distribution.

Turning to education levels, Figure 4.4 presents the estimates and confidence intervals of the education dummies derived from the regressions in Table 4.4. In all cases, individuals with a high-school degree exhibit noticeably more optimistic expectations compared to those without one. Additionally, the difference in expectations between individuals

[10] Our analysis assumes a linear relationship in the responses. That is, regarding expectations moving from -2 to -1 is the same as moving from 1 to 2. To allow for non-linear associations, we use ordinal logit models, see Appendix 6. The findings of this analysis are qualitatively similar to the results reported in Table 4.4.

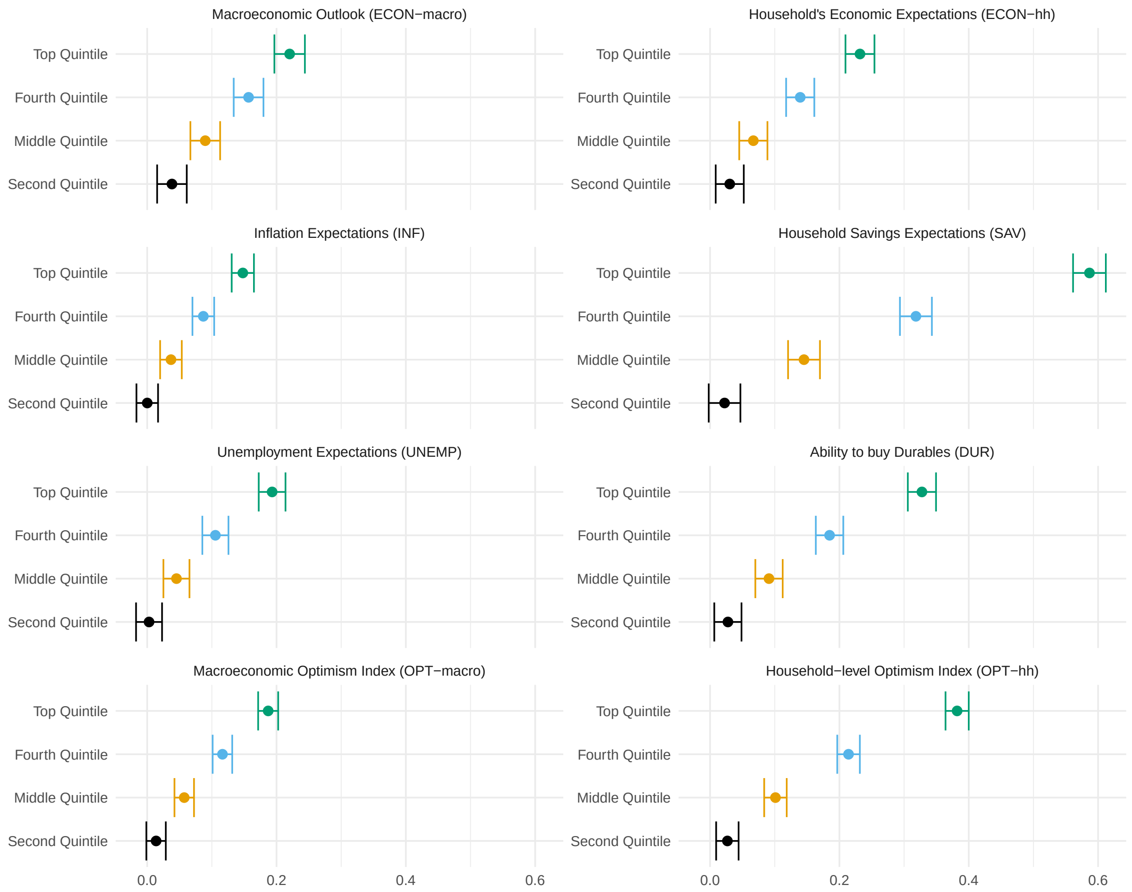

Figure 4.3: Estimates and corresponding 95% confidence intervals of the income quintile dummies compared to the bottom income quintile.

with a high-school degree and those with a university degree is also significant in most cases, with the exception of ECON-macro and UNEMP. Hence, differences in expectations do not only materialize if we use a binary classification based on a university degree but there are also clear disparities in expectations at lower education levels. Similar to income ranks, differences between the different education groups tend to be larger when considering expectations on the household level. It is evident that a higher level of education correlates positively with a higher income level (Bryan, Venkatu, et al., 2001), but by including both income quintiles and education levels in the regressions, we account for these correlations. Comparing the differences in macroeconomic expectations between individuals without a high-school degree and those with one, and the differences between the latter group and those with a university degree, we find that in most cases, the disparities at the lower end are greater than those at the higher end. That is, there are larger shifts in expectations when we move from no high-school degree to a high-school degree compared to when we move from a high-school degree to a university degree. This pattern holds for macroeconomic expectations and their components. For household-level expectations, we observe roughly equal increases for ECON-hh and OPT-hh, while for savings the jump between a high-school degree and a university degree is larger compared to when we move from no high-school degree to a high-school degree.

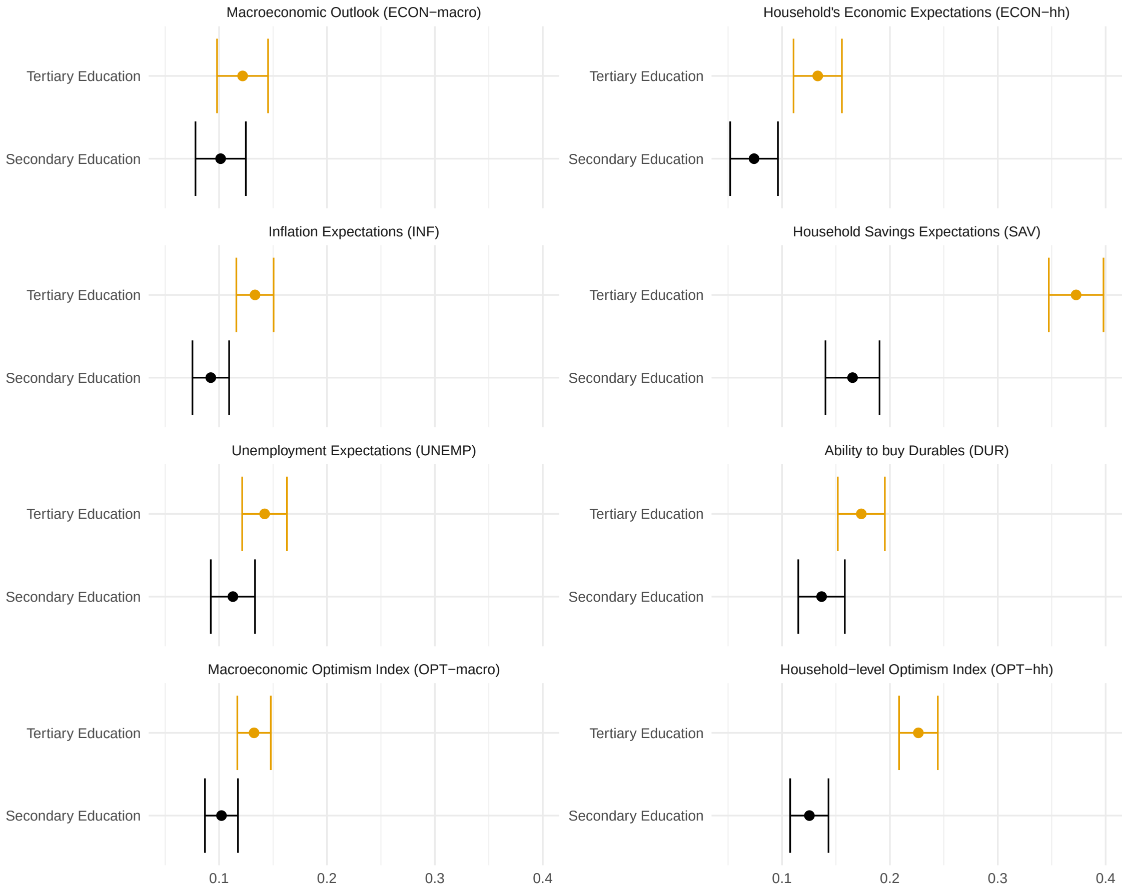

Figure 4.4: Estimates and corresponding 95% confidence intervals of education level dummies compared to education levels lower than secondary grade. Note: Higher values mean a more optimistic expectation

To compare the influence of each variable, we use the standardized coefficients presented in Table 4.24 of Appendix 5. When considering macroeconomic expectations, the difference in expectations between individuals with a university degree and those without a high-school degree is smaller compared to the difference between individuals in the bottom and the top income quintile. The same pattern holds for all other expectations, except inflation. If we take the influence of a recession as a reference, then in the case of ECON-macro, ECON-hh, SAV and DUR, we observe that for ECON-macro, ECON-hh, SAV, and DUR, the differences in expectations based on income quintiles or education level are often larger than the influence of a recession on expectations (except for INF and UNEMP). This suggests that the variation in expectations resulting from SES is significant.

Since income quintiles and education levels are positively correlated, there is a potential concern regarding multicollinearity. In Tables 4.22 and 4.23 of Appendix 4, we run separate regressions for all the expectations variables, considering the income quintiles and the education levels separately. The signs and the magnitudes of the coefficients change in the expected way due to the positive correlation between income and education. Nevertheless, the qualitative findings regarding the non-linearity of the coefficients for income quintiles and education levels still hold true. It is important to note that there was an increase in the proportion of individuals with a university degree in the population during the time period covered by our data, which suggests a composition effect.

Additionally, one might argue that taking the average of various variables on expectations might not result in proper ”optimism-indicies”. In Appendix 1, we rerun the most important specifications of our analysis for a dependent variable constructed using Principal Component Analysis (PCA). Our argument is that using all the variables, we can also construct a latent ”expectations” variable. however, using these specifications do not change our results.

Overall, in line with the study closest to ours (Das et al., 2020), we document significant differences in macroeconomic expectations based on income and education levels. However, our study also reveals novel findings. Importantly, the associations between the socioeconomic variables and expectations do not seem to be linear in all cases. Regarding income, we observe minimal or negligible differences in expectations within the lower quintiles, but substantial disparities between these lower quintiles and the upper two quintiles. There are also noticeable differences between the two highest quintiles. In terms of education levels, we observe not only between individuals with and without university degree differences, but also between those without and with high-school degree, which is a finding not previously documented in the existing literature. We also show that these patterns are more pronounced in the case of household-level expectations.

The role of past experiences and optimism

As a further step, we extend the analysis by considering potential factors identified in previous literature that may influence macroeconomic expectations. In this section, we focus on the optimism indices that we have constructed. Based on the literature review, we concentrate on three factors closely related to personal experiences and optimism. Two factors, recessions and experiences in the past year, capture personal experiences, while optimism is proxied by household-level expectations. Recessions generally result in negative personal experiences, leading to worse expectations as indicated in Table 4.4. Following Das et al., 2020, we study whether income quintiles and education levels are associated differently with expectations during recessions than in other times, using interaction terms. To account for past experiences in general (not only focusing on recessions), we also explore the role of the self-assessed change in the household's economic situation in the last year (HH-Prev.Year). According to the literature, it may be correlated with macroeconomic expectations (D'Acunto, Malmendier, Ospina, and Weber, 2021; Kuchler and Zafar, 2019; Malmendier and Nagel, 2011, 2016). To capture optimism that may be related to more optimistic macroeconomic expectations we include the household-level expectation. More optimistic households may have more positive macroeconomic expectations. Conversely, when the dependent variable is household-level expectation, then we add macroeconomic expectations as the latter may influence the former: gloomier macroeconomic expectations may cast a shadow on household-level optimism. We study the role of these variables separately and then jointly, as shown in Table 4.5. [11]

Starting with macroeconomic expectations, in line with Das et al., 2020, figures 4.1-4.2 suggest that during recessions, the difference in macroeconomic expectations narrows. However, for the income quintiles, we do not observe such a pattern as the coefficients of the interaction terms involving the upper quintiles are not significantly negative. In fact, once we control for past experiences in specifications (2), (3), and (4), the interactions become positive and significant, indicating that differences in macroeconomic expectations increase during recessions. Regarding education levels, the interaction terms are negative and significant, indicating that during recessions the difference in macroe-

[11] We acknowledge that including both macroeconomic and household-level economic expectations, as well as previous experience, in the model might reveal a multicollinearity issue. However, we have included these specifications because the literature remains unclear on the causal direction of expectations. Arguments can be made for both macroeconomic expectations causing household-level expectations and vice versa. Additionally, past experiences strongly influence our current outlook. Nonetheless, it is possible that a latent variable, such as 'general optimism,' affects all these factors. Therefore, to gain a better understanding of the effects, we controlled for possible unwanted mechanisms.

conomic expectations diminishes between individuals without a high-school degree and those with one. Interestingly, once we take into account past experiences and household-level expectations, the effect of these interactions changes very little. Experiences of the past year (HH.Prev.Year) and household-level optimism have a consistently positive and significant coefficient, indicating that more optimistic households and households with better past experiences have more positive macroeconomic expectations, *ceteris paribus*. Note that when household-level optimism is included in the regression, the significance of the income quintiles almost vanishes, while the education dummies remain significant (though the magnitude of the coefficient decreases considerably). This finding suggests that household-level optimism reflects to a large extent the income ranking of the household, and once we take it into account, income quintiles do not play a role anymore.

Turning to household-level expectations, we observe that differences in expectations decrease during recessions when considering the income quintiles. However, we do not find the same for education levels, as during recessions, the differences in expectations between individuals without a high-school degree and those with a university degree actually become larger in some specifications. Similar to the previous findings, experiences of the past year and macroeconomic expectations show a consistently positive and significant coefficient. However, even after including past experiences and macroeconomic expectations, both income ranks and education levels retain their significance. [12]

One pertinent question is whether past experiences differentially influence economic expectations across various income levels. To explore this, we conducted regressions on $Opt - macro$ and $Opt - HH$, using interactions with income quintiles. The findings, detailed in Table 4.30 in the Appendix, present a nuanced picture. For macroeconomic expectations, the interaction effects are somewhat inconsistent, initially increasing and then decreasing. Notably, at higher income levels, a positive previous experience correlates with a slight decrease in expectations, though the effect is marginal. Conversely, for household-level expectations, the relationship is more straightforward: individuals with higher income levels and favorable experiences from the previous year exhibit a progressively positive effect on their expectations.

[12] We also replicate the regression specification of Das et al., 2020 (see Table 4.28 in Appendix 7). When we do not control for experiences during the past year, similar to their results, we observe a significant linear association between quintiles and macroeconomic expectations. We obtain the same results for household-level expectations, even after accounting for experiences during the past year and macroeconomic expectations. The interaction terms related to recessions reproduce Das et al., 2020's findings concerning income rank, but not for education.

Expectations, SES, and economic decisions

So far, we investigated the relationship between SES and macroeconomic and household-level expectations, and we studied the role of potential factors. However, it is natural to ask whether these expectations have an impact on economic decision-making. While we cannot test the direct effect of expectations on actual purchasing decisions, following Das et al., 2020, we can assess whether there is a connection between macroeconomic expectations and the intention of the household to purchase a car, or a home, or make major expenditures related to the home. In the regression analysis, we also include $DUR-worth$, as respondents' subjective evaluation of whether it is worth buying durable goods can be informative[13]. Table 4.6 contains the results of OLS regressions.

Similarly to our previous results, the coefficients of income quintiles (particularly for the top quintiles) and education levels (primarily for individuals with a university degree) are significant, indicating that SES is associated with these economic decisions. Experiences in the past year show a consistently positive and significant association with the dependent variable in all specifications. It is crucial to highlight, however, that the belief channel—namely, the OPT indices and the confounding past experiences—exerts a substantial influence on economic decisions. The magnitude of their coefficients is comparable to the impact of being in the top 20 percent of the income distribution.

Moreover, when macroeconomic expectations are included separately (specifications (1), (4), (7), and (10)), they are significantly related to the intention to purchase big-ticket items, even after accounting for income rank and education level. Hence, having more positive macroeconomic expectations are positively associated with the purchase intent, beyond the influence of socioeconomic variables. We observe a similar pattern when considering household-level optimism. In specifications (2), (5), (8) and (11), where it is included solely, household-level optimism shows a significant and positive relationship with the dependent variable. Additionally, the coefficient for household-level optimism appears to be substantially larger than the coefficient of macroeconomic expectations. When both expectation measures are included (specifications (3), (6), (9) and (12)), household-level optimism remains consistently positive and significant. However, we do not see a consistent pattern when considering macroeconomic expectations. Since the expectation variables are highly correlated, there seems to be a multicollinearity issue, as the sum of the coefficients of the expectation variables approximately equals the coefficient

[13] We acknowledge that $DUR - worth$ may be also related to respondents' ability and not only their intent.

of the household-level expectation when included separately. The main message from Table 4.6 is that macroeconomic and household-level expectations do not merely reflect SES, but they are also closely related to economic decisions beyond their relationship with SES.

Table 4.5: The relationships between past experiences (recession and self-assessed change in economic situation), optimism and macroeconomic expectations

	Dependent variable:							
	OPT-macro				OPT-hh			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Quintile 2	0.008	-0.011	-0.002	-0.007	0.021**	-0.009	0.016**	-0.004
	(0.008)	(0.008)	(0.007)	(0.007)	(0.010)	(0.008)	(0.008)	(0.007)
Quintile 3	0.063***	0.022***	0.008	0.001	0.114***	0.048***	0.073***	0.038***
	(0.008)	(0.008)	(0.007)	(0.007)	(0.010)	(0.008)	(0.008)	(0.007)
Quintile 4	0.111***	0.043***	0.004	-0.005	0.224***	0.116***	0.152***	0.096***
	(0.008)	(0.008)	(0.007)	(0.007)	(0.010)	(0.008)	(0.008)	(0.007)
Quintile 5	0.190***	0.071***	-0.003	-0.018**	0.401***	0.212***	0.279***	0.179***
	(0.009)	(0.008)	(0.007)	(0.007)	(0.010)	(0.008)	(0.008)	(0.008)
Rec. × Quintile 2	0.031	0.036**	0.018	0.021	0.029	0.036*	0.009	0.019
	(0.020)	(0.018)	(0.017)	(0.016)	(0.023)	(0.019)	(0.019)	(0.017)
Rec. × Quintile 3	-0.034*	-0.008	0.004	0.008	-0.080***	-0.037*	-0.058***	-0.033*
	(0.020)	(0.018)	(0.017)	(0.016)	(0.023)	(0.019)	(0.019)	(0.017)
Rec. × Quintile 4	0.027	0.041**	0.057***	0.058***	-0.063***	-0.041**	-0.080***	-0.060***
	(0.020)	(0.018)	(0.017)	(0.016)	(0.023)	(0.019)	(0.019)	(0.017)
Rec. × Quintile 5	-0.018	0.010	0.037**	0.038**	-0.114***	-0.068***	-0.103***	-0.073***
	(0.020)	(0.019)	(0.017)	(0.017)	(0.024)	(0.020)	(0.020)	(0.018)
Has university degree	0.142***	0.096***	0.035***	0.033***	0.222***	0.150***	0.131***	0.105***
	(0.009)	(0.008)	(0.007)	(0.007)	(0.010)	(0.008)	(0.008)	(0.008)
Has High-School Degree	0.107***	0.081***	0.048***	0.047***	0.122***	0.081***	0.054***	0.044***
	(0.009)	(0.008)	(0.007)	(0.007)	(0.010)	(0.008)	(0.008)	(0.007)
Rec. × Univ. degree	-0.055***	-0.060***	-0.067***	-0.066***	0.024	0.014	0.059***	0.042**
	(0.021)	(0.019)	(0.018)	(0.017)	(0.024)	(0.020)	(0.020)	(0.018)
Rec. × High-School Degree	-0.026	-0.037*	-0.033*	-0.035**	0.015	-0.003	0.032	0.014
	(0.020)	(0.019)	(0.017)	(0.017)	(0.024)	(0.020)	(0.020)	(0.018)
HH.Prev.Year		0.283***		0.093***		0.451***		0.320***
		(0.002)		(0.003)		(0.003)		(0.002)
OPT-hh			0.481***	0.420***				
			(0.003)	(0.003)				
OPT-macro							0.643***	0.466***
							(0.003)	(0.003)
Recession	0.045	0.095**	0.072*	0.085**	-0.056	0.023	-0.085**	-0.021
	(0.045)	(0.042)	(0.038)	(0.037)	(0.052)	(0.044)	(0.043)	(0.039)
Constant	-1.238***	-0.930***	-0.704***	-0.670***	-1.109***	-0.620***	-0.313***	-0.186***
	(0.033)	(0.031)	(0.028)	(0.028)	(0.038)	(0.032)	(0.032)	(0.029)

Standard errors in parentheses.
*/**/*** denotes significance at 10 / 5 / 1 % level.
All regressions include year-month dummies, age, gender and family status.

Table 4.6: Expectations, SES, and economic decisions

	Dependent variable:											
	DUR-worth			HOME			CAR			HOME-exp		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Quintile 2	−0.001	−0.001	−0.001	0.001	0.002	0.002	−0.004	−0.002	−0.002	0.049**	0.053**	0.053**
	(0.004)	(0.004)	(0.004)	(0.014)	(0.014)	(0.014)	(0.015)	(0.015)	(0.015)	(0.021)	(0.021)	(0.021)
Quintile 3	0.016***	0.014***	0.014***	0.005	−0.00004	−0.00002	0.018	0.008	0.008	0.091***	0.079***	0.078***
	(0.004)	(0.004)	(0.004)	(0.014)	(0.014)	(0.014)	(0.015)	(0.015)	(0.015)	(0.021)	(0.021)	(0.021)
Quintile 4	0.037***	0.030***	0.030***	0.022	0.010	0.010	0.056***	0.030**	0.030**	0.182***	0.147***	0.147***
	(0.004)	(0.004)	(0.004)	(0.014)	(0.014)	(0.014)	(0.015)	(0.015)	(0.015)	(0.021)	(0.021)	(0.021)
Quintile 5	0.064***	0.051***	0.051***	0.151***	0.129***	0.129***	0.270***	0.224***	0.223***	0.362***	0.299***	0.298***
	(0.004)	(0.004)	(0.004)	(0.014)	(0.014)	(0.014)	(0.016)	(0.016)	(0.016)	(0.022)	(0.022)	(0.022)
Has university degree	0.038***	0.030***	0.029***	0.107***	0.091***	0.091***	0.136***	0.102***	0.103***	0.126***	0.080***	0.081***
	(0.004)	(0.004)	(0.004)	(0.014)	(0.014)	(0.014)	(0.016)	(0.016)	(0.016)	(0.022)	(0.021)	(0.021)
Has High-School Degree	0.009**	0.007	0.006	0.001	−0.005	−0.005	0.066***	0.052***	0.052***	0.034	0.014	0.015
	(0.004)	(0.004)	(0.004)	(0.014)	(0.014)	(0.014)	(0.016)	(0.015)	(0.015)	(0.022)	(0.021)	(0.021)
OPT-macro	0.047***		0.012***	0.058***		0.001	0.103***		−0.017**	0.133***		−0.031***
	(0.002)		(0.002)	(0.007)		(0.008)	(0.008)		(0.009)	(0.011)		(0.012)
OPT-hh		0.081***	0.076***		0.128***	0.128***		0.264***	0.271***		0.357***	0.370***
		(0.002)	(0.002)		(0.007)	(0.007)		(0.007)	(0.008)		(0.010)	(0.011)
HH.Prev.Year	0.034***	0.011***	0.010***	0.062***	0.021***	0.021***	0.110***	0.020***	0.022***	0.156***	0.033***	0.036***
	(0.002)	(0.002)	(0.002)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.008)	(0.008)	(0.008)
Constant	−0.897***	−0.891***	−0.883***	−1.234***	−1.208***	−1.207***	−1.248***	−1.180***	−1.192***	−0.819***	−0.720***	−0.742***
	(0.018)	(0.018)	(0.018)	(0.033)	(0.033)	(0.033)	(0.037)	(0.036)	(0.037)	(0.052)	(0.050)	(0.051)
Observations	72,264	72,264	72,264	27,767	27,767	27,767	27,742	27,742	27,742	27,694	27,694	27,694
R^2	0.074	0.088	0.089	0.072	0.082	0.082	0.117	0.150	0.151	0.108	0.142	0.142
Adjusted R^2	0.073	0.087	0.087	0.071	0.081	0.081	0.115	0.149	0.149	0.107	0.140	0.140
Residual Std. Error	0.342	0.339	0.339	0.699	0.695	0.695	0.786	0.770	0.770	1.081	1.060	1.060
F Statistic	47.669***	57.881***	57.651***)	42.951***	49.742***	48.786***	73.377***	99.063***	97.261***	67.379***	92.199***	90.629***

Standard errors in parentheses.

*/**/*** denotes significance at 10 / 5 / 1 % level.

All regressions include year-month dummies, age, gender, family status and recession dummy.

4.5 Conclusion

In this study, our aim is to disentangle the relationship between SES and macroeconomic/household-level expectations, as well as explore the implications of expectations on economic decision-making. To achieve this, we use a sample of approximately 80,000 observations from Hungary, covering the period from 2000 to 2009. We focus on how two aspects of SES (income rank and education level) are associated with the expectations reported by the respondents.

Our study makes several contributions to the existing literature. First, in addition to examining macroeconomic expectations, we also investigate household-level expectations. We find a strong correlation between these two types of expectations. Moreover, when we account for household-level expectations, the significance of income rank diminishes, suggesting that household-level expectations reflect the household's income situation. Second, in contrast to Das et al., 2020, we document that the relationship between SES and macroeconomic expectations is not linear. While notable differences exist between the lower and upper income quintiles, within the lower quintiles, the differences are rather small or non-existent. Third, our analysis reveals that a more nuanced examination of education levels enhances our understanding. Differences in macroeconomic expectations are not only observed between individuals with and without a university degree but also between individuals without a high-school degree and those with one. The patterns observed in macroeconomic expectations are mirrored in household-level expectations. Fourth, we highlight the importance of past experiences and optimism in shaping macroeconomic expectations. Including these factors in the analysis reduces the influence of SES variables. Last, our findings demonstrate that both macroeconomic and household-level expectations significantly impact economic decisions, as captured by purchase intentions. Even after controlling for SES variables, these expectations remain relevant, underscoring the need for a comprehensive understanding of these expectations.

We acknowledge two limitations in our research: first, we have relatively short data ranging from the middle of 2000 until the end of 2009 due to availability issues, and within this time frame, we only have limited data during recession. [14] Additionally, although we have data on the qualitative assessment of various economic expectations, we are unable to determine whether these expectations turned out to be correct or not, except in terms of their directional accuracy (i.e., whether respondents correctly predicted the

[14] In some cases, economic expectations were very heavily affected by political issues, as we noted about the sharp decline observed in 2006.

sign of the change, such as inflation, in the next period, as examined in Appendix 3). It would be valuable for future research to investigate the factors contributing to having a ”correct expectation” and explore whether socioeconomic factors have an impact on this. Additionally, alternative estimation strategies such as propensity score matching might be beneficial to further explore to magnitudes of the effects.

Conclusions

In each of the previous four essays, I present research rooted in behavioral and experimental economics. While the essays come from different topics, they share the methodological and theoretical approaches offered by behavioral economics. Chapter 1 investigates the influence of loss aversion framing on students' test performance. Employing three different approaches — giving points, granting students a perfect score with deductions for incorrect answers, and alternating between giving and subtracting points — the study explores how individuals react to gains and losses in a classroom setting. The findings suggest that studying with a loss-frame improves performance more than just earning points. For tests of heterogeneity, we find no evidence of differential gender effect. We test for differences in student-quality, again, finding no evidence for heterogeneous treatment effects. Finally, we test whether improvement in scoring is a novelty-effect, by comparing students who were losing points throughout the semester to those who only lost points at the final test. The results do not support the idea of novelty effect, suggesting that loss-framing is not diminishing over time - at least for the duration of one semester.

Chapter 2 examines reference point-based decision-making from the perspective of fairness. Its hypotheses suggest that fairness perception plays a significant role in comparison-based decision-making, especially when harmful behavior may also be present in the decision-making situation. The study shows that in social decision-making, unfair differences can lead to so-called malicious envy; however, if the decision-maker is able to rationalize the reason for the discrepancy, the harmful behavior may cease. Different types of new information yield different results: in Hungary, people are more likely to accept someone having a higher salary if they put in more effort, as opposed to higher qualifications or more experience being the reason behind wage differences.

In Chapter 3, I examine the dynamics of economic preferences and their groupings over time using data from a Hungarian survey conducted in two waves during 2020. Members of one cluster, characterised by higher cooperation, competitiveness, more patience and lower time-inconsistency and an internal locus of control, were consistently younger, more educated, and had a more balanced gender distribution. Members of this cluster also reported higher net incomes. An alternative clustering approach using only the first wave data and projecting onto the second wave confirmed these findings, highlighting a correlation between higher socio-economic status and certain economic preferences. Limitations include the survey-based measurement of preferences and the data used for the clustering analysis being a pooled cross-sectional. Despite these limitations, the findings provide insights into the stability and socio-economic correlates of clusters of economic preferences.

In Chapter 4, using Hungarian monthly survey data between 2000 and 2009, we show that the relationship between expectations (both at the macroeconomic and household levels) and socioeconomic status (SES), as represented by income rank and education level, is non-linear. In many instances, there is no significant difference in expectations between the two lower quintiles. However, individuals in the upper (fourth and top) quintiles exhibit significantly more positive expectations than those in the lower quintiles. There is also a clear difference in expectations between the fourth and the top quintiles. In terms of education level, individuals with a high-school degree have significantly more positive expectations compared to their peers without one. Significant differences in economic expectations are also observed between high-school graduates and individuals with a university diploma, particularly regarding inflation, savings expectations, and the assessment of the household's future financial situation. Disparities in household-level expectations based on SES are more pronounced than those in macroeconomic expectations. Past experiences and household-level optimism seem to be key factors influencing macroeconomic expectations. Furthermore, we document that both macroeconomic and household-level expectations predict the intention for significant expenditures, even after controlling for SES variables.

Bibliography

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. M. (2023). When should you adjust standard errors for clustering? *The Quarterly Journal of Economics*, *138*(1), 1–35.

Agarwal, S., & Mazumder, B. (2013). Cognitive abilities and household financial decision making. *American Economic Journal: Applied Economics*, *5*(1), 193–207.

Akerlof, G. A. (1978). The market for “lemons”: Quality uncertainty and the market mechanism. In *Uncertainty in economics* (pp. 235–251). Elsevier.

Akerlof, G. A. (1980). A theory of social custom, of which unemployment may be one consequence. *The Quarterly Journal of Economics*, *94*(4), 749–775.

Akerlof, G. A. (1982). Labor contracts as partial gift exchange. *The Quarterly Journal of Economics*, *97*(4), 543–569.

Alesina, A., & Giuliano, P. (2011). Preferences for redistribution. In *Handbook of social economics* (pp. 93–131, Vol. 1). Elsevier.

Alsharawy, A., Ball, S., Smith, A., & Spoon, R. (2021). Fear of covid-19 changes economic preferences: Evidence from a repeated cross-sectional mturk survey. *Journal of the Economic Science Association*, *7*(2), 103–119.

Angelico, C., & Di Giacomo, F. (2019). Heterogeneity in inflation expectations and personal experience. *Available at SSRN 3369121*.

Apostolova-Mihaylova, M., Cooper, W., Hoyt, G., & Marshall, E. C. (2015). Heterogeneous gender effects under loss aversion in the economics classroom: A field experiment. *Southern Economic Journal*, *81*(4), 980–994.

Armantier, O., Bruine de Bruin, W., Topa, G., Van Der Klaauw, W., & Zafar, B. (2015). Inflation expectations and behavior: Do survey respondents act on their beliefs? *International Economic Review*, *56*(2), 505–536.

Arnold, E., Dräger, L., & Fritsche, U. (2014). *Evaluating the link between consumers’ savings portfolio decisions, their inflation expectations and economic news* (tech. rep.). DEP (Socioeconomics) Discussion Papers-Macroeconomics and Finance Series.

Arora, M., & Kumari, S. (2015). Risk taking in financial decisions as a function of age, gender: Mediating role of loss aversion and regret. *International Journal of Applied Psychology*, *5*(4), 83–89.

Arrow, K. J. (2012). *Social choice and individual values* (Vol. 12). Yale University Press.

Bachmann, R., Berg, T. O., & Sims, E. R. (2015). Inflation expectations and readiness to spend: Cross-sectional evidence. *American Economic Journal: Economic Policy*, *7*(1), 1–35.

Bamieh, O., & Ziegler, L. (2024). Can wage transparency alleviate gender sorting in the labor market? *Economic Policy*, eiae025.

Bedeian, A. G. (1995). Workplace envy. *Organizational Dynamics*, *23*(4), 49–56.

Bennedsen, M., Larsen, B., & Wei, J. (2023). Gender wage transparency and the gender pay gap: A survey. *Journal of Economic Surveys*, *37*(5), 1743–1777.

Bies-Hernandez, N. J. (2012). The effects of framing grades on student learning and preferences. *Teaching of Psychology*, *39*(3), 176–180.

Binder, C. (2020). Coronavirus fears and macroeconomic expectations. *Review of Economics and Statistics*, *102*(4), 721–730.

Binder, C. C., & Brunet, G. (2022). Inflation expectations and consumption: Evidence from 1951. *Economic Inquiry*, *60*(2), 954–974.

Binmore, K., Shaked, A., & Sutton, J. (1985). Testing noncooperative bargaining theory: A preliminary study. *American Economic Review*, *75*(5), 1178–1180.

Blanchflower, D. G., & MacCoille, C. (2009). *The formation of inflation expectations: An empirical analysis for the uk* (tech. rep.). National Bureau of Economic Research.

Böheim, R., & Gust, S. (2021). The austrian pay transparency law and the gender wage gap. *CESifo Working Paper*.

Böhm, R., Fleiß, J., & Rybnicek, R. (2021). On the stability of social preferences in inter-group conflict: A lab-in-the-field panel study. *Journal of Conflict Resolution*, *65*(6), 1215–1248.

Bokern, P., Linde, J., Riedl, A., & Werner, P. (2021). *The effect of the covid-19 crisis on economic and social preferences*. Netspar, Network for Studies on Pensions, Aging; Retirement.

Bolton, G. E., & Ockenfels, A. (2000). Erc: A theory of equity, reciprocity, and competition. *American Economic Review*, *91*(1), 166–193.

Borgen, N. T., Haupt, A., & Wiborg, Ø. N. (2024). A new framework for estimation of unconditional quantile treatment effects: The residualized quantile regression (rqr) model. *OSF*.

Botsch, M. J., Malmendier, U. M., et al. (2020). *The long shadows of the great inflation: Evidence from residential mortgages* (tech. rep.). CEPR Discussion Papers.

Bouchouicha, R., Deer, L., Eid, A. G., McGee, P., Schoch, D., Stojic, H., Ygosse-Battisti, J., & Vieider, F. M. (2019). Gender effects for loss aversion: Yes, no, maybe? *Journal of Risk and Uncertainty*, *59*, 171–184.

Breitkopf, L., Chowdhury, S., Shambhavi, P., Schildberg-Hörisch, H., & Sutter, M. (2024). Do economic preferences of children predict behavior? *MPI Collective Goods Discussion Paper*, (2024/9).

Brown, S., & Taylor, K. (2006). Financial expectations, consumption and saving: A microeconomic analysis. *Fiscal Studies*, *27*(3), 313–338.

Bruhin, A., Fehr, E., & Schunk, D. (2019). The many faces of human sociality: Uncovering the distribution and stability of social preferences. *Journal of the European Economic Association*, *17*(4), 1025–1069.

Bruine de Bruin, W., Vanderklaauw, W., Downs, J. S., Fischhoff, B., Topa, G., & Armantier, O. (2010). Expectations of inflation: The role of demographic variables, expectation formation, and financial literacy. *Journal of Consumer Affairs*, *44*(2), 381–402.

Bryan, M. F., Venkatu, G., et al. (2001). The demographics of inflation opinion surveys.

Burke, M. A., & Ozdagli, A. (2014). Household inflation expectations and consumer spending: Evidence from panel data. *The Review of Economics and Statistics*, 1–45.

Burks, S. V., Carpenter, J. P., Goette, L., & Rustichini, A. (2009). Cognitive skills affect economic preferences, strategic behavior, and job attachment. *Proceedings of the National Academy of Sciences*, *106*(19), 7745–7750.

Carlsson, F., Johansson-Stenman, O., & Nam, P. K. (2014). Social preferences are stable over long periods of time. *Journal of Public Economics*, *117*, 104–114.

Carroll, C. D., & Dunn, W. E. (1997). Unemployment expectations, jumping (s, s) triggers, and household balance sheets. *NBER Macroeconomics Annual*, *12*, 165–217.

Chapman, J., Dean, M., Ortoleva, P., Snowberg, E., & Camerer, C. (2023). Econographics. *Journal of Political Economy Microeconomics*, *1*(1), 115–161.

Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2014). Nbclust: An r package for determining the relevant number of clusters in a data set. *Journal of Statistical Software*, *61*, 1–36.

Chowdhury, S., Sutter, M., & Zimmermann, K. F. (2022). Economic preferences across generations and family clusters: A large-scale experiment in a developing country. *Journal of Political Economy*, *130*(9), 2361–2410.

Chuang, Y., & Schechter, L. (2015). Stability of experimental and survey measures of risk, time, and social preferences: A review and some new results. *Journal of Development Economics*, *117*, 151–170.

Cohen-Charash, Y. (2009). Episodic envy. *Journal of Applied Social Psychology*, *39*(9), 2128–2173.

Cohen-Charash, Y., & Mueller, J. S. (2007). Does perceived unfairness exacerbate or mitigate interpersonal counterproductive work behaviors related to envy? *Journal of Applied Psychology*, *92*(3), 666.

Coibion, O., Georgarakos, D., Gorodnichenko, Y., & Van Rooij, M. (2019). *How does consumption respond to news about inflation? field evidence from a randomized control trial* (tech. rep.). National Bureau of Economic Research.

Coibion, O., Gorodnichenko, Y., & Weber, M. (2020). *Does policy communication during covid work?* (Tech. rep.). National Bureau of Economic Research.

Coibion, O., Gorodnichenko, Y., & Weber, M. (2022). Monetary policy communications and their effects on household inflation expectations. *Journal of Political Economy*, *130*(6), 000–000.

Cole, N. S. (1997). *The ets gender study: How females and males perform in educational settings* (tech. rep.). Educational Testing Service, Princeton, NJ.

Cullen, Z., & Perez-Truglia, R. (2022). How much does your boss make? the effects of salary comparisons. *Journal of Political Economy*, *130*(3), 766–822.

Czibor, E., Onderstal, S., Sloof, R., & van Praag, C. M. (2020). Does relative grading help male students? evidence from a field experiment in the classroom. *Economics of Education Review*, *75*, 101953.

D'Acunto, F., Hoang, D., Paloviita, M., & Weber, M. (2019). *Iq, expectations, and choice* (tech. rep.). National Bureau of Economic Research.

D'Acunto, F., Hoang, D., Paloviita, M., & Weber, M. (2021). Iq, expectations, and choice. *Chicago Booth Research Paper*, (19-20).

D'Acunto, F., Hoang, D., & Weber, M. (2016). *The effect of unconventional fiscal policy on consumption expenditure* (tech. rep.). National Bureau of Economic Research.

D'Acunto, F., Hoang, D., & Weber, M. (2018). Unconventional fiscal policy. *AEA Papers and Proceedings*, *108*, 519–23.

D'Acunto, F., Hoang, D., & Weber, M. (2022). Managing households' expectations with unconventional policies. *The Review of Financial Studies*, *35*(4), 1597–1642.

D'Acunto, F., Malmendier, U., Ospina, J., & Weber, M. (2019). *Exposure to daily price changes and inflation expectations* (tech. rep.). National Bureau of Economic Research.

D'Acunto, F., Malmendier, U., Ospina, J., & Weber, M. (2021). Exposure to grocery prices and inflation expectations. *Journal of Political Economy*, *129*(5), 1615–1639.

D'Acunto, F., Malmendier, U., & Weber, M. (2021). Gender roles produce divergent economic expectations. *Proceedings of the National Academy of Sciences*, *118*(21), e2008534118.

D'Acunto, F., Malmendier, U., & Weber, M. (2022). *What do the data tell us about inflation expectations?* (Tech. rep.). National Bureau of Economic Research.

Das, S., Kuhnen, C. M., & Nagel, S. (2020). Socioeconomic status and macroeconomic expectations. *The Review of Financial Studies*, *33*(1), 395–432.

Dasgupta, U., Gangadharan, L., Maitra, P., & Mani, S. (2017). Searching for preference stability in a state dependent world. *Journal of Economic Psychology*, *62*, 17–32.

De Oliveira, A. C., Eckel, C., & Croson, R. T. (2012). The stability of social preferences in a low-income neighborhood. *Southern Economic Journal*, *79*(1), 15–45.

Devos, K. (2013). *Factors influencing individual taxpayer compliance behaviour*. Springer Science & Business Media.

Diaz-Bonilla, E., Thomas, M., Robinson, S., & Cattaneo, A. (2000). Food security and trade negotiations in the world trade organization: A cluster analysis of country groups. *TMD Discussion Paper 59*.

Disdier, A.-C., & Van Tongeren, F. (2010). Non-tariff measures in agri-food trade: What do the data tell us? evidence from a cluster analysis on oecd imports. *Applied Economic Perspectives and Policy*, *32*(3), 436–455.

Dominitz, J., & Manski, C. F. (2004). How should we measure consumer confidence? *Journal of Economic Perspectives*, *18*(2), 51–66.

Dräger, L. (2015). Inflation perceptions and expectations in sweden–are media reports the missing link? *Oxford Bulletin of Economics and Statistics*, *77*(5), 681–700.

Dräger, L., & Nghiem, G. (2021). Are consumers' spending decisions in line with a euler equation? *Review of Economics and Statistics*, *103*(3), 580–596.

Epper, T., Senn, J., & Fehr, E. (2024). The missing type: Where are the inequality averse (students)? *CESifo Working Paper*.

Ertl, A., Horn, D., & Kiss, H. J. (2024). *Economic preferences across generations and family clusters: A comment* (tech. rep.). I4R Discussion Paper Series.

Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis*. John Wiley & Sons.

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018). Global evidence on economic preferences. *The quarterly journal of economics*, *133*(4), 1645–1692.

Fallucchi, F., Nosenzo, D., & Reuben, E. (2020). Measuring preferences for competition with experimentally-validated survey questions. *Journal of Economic Behavior & Organization*, *178*, 402–423.

Faulk, L. H., Settlage, D. M., Wollscheid, J. R., et al. (2019). Influencing positive student behaviour using the endowment effect. *e-Journal of Business Education and Scholarship of Teaching*, *13*(1), 20–29.

Fehr, E., & Charness, G. (2023). Social preferences: Fundamental characteristics and economic consequences. *Working paper series/Department of Economics*, (432).

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, *114*(3), 817–868.

Fehr, E., & Schmidt, K. M. (2000). Fairness, incentives, and contractual choices. *European Economic Review*, *44*(4-6), 1057–1068.

Festinger, L. (1954). A theory of social comparison processes. *Human relations*, *7*(2), 117–140.

Friehe, T., & Schildberg-Hörisch, H. (2018). Predicting norm enforcement: The individual and joint predictive power of economic preferences, personality, and self-control. *European Journal of Law and Economics*, *45*, 127–146.

Gáspár, A., & Reizer, B. (2020). *Average wages at exceptional times: Wage trends in hungary during the first eighteen months of the coronavirus pandemic*. The Hungarian Labor Market 2020; Centre for Economic; Regional Studies, Institute of Economics.

GKI. (2022). *Market and business cycle analyses*. Retrieved October 13, 2022, from https://www.gki.hu/language/en/surveys/

Grinblatt, M., Ikäheimo, S., Keloharju, M., & Knüpfer, S. (2016). Iq and mutual fund choice. *Management Science*, *62*(4), 924–944.

Grinblatt, M., Keloharju, M., & Linnainmaa, J. (2011). Iq and stock market participation. *The Journal of Finance*, *66*(6), 2121–2164.

Grinblatt, M., Keloharju, M., & Linnainmaa, J. T. (2012). Iq, trading behavior, and performance. *Journal of Financial Economics*, *104*(2), 339–362.

Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization*, *3*(4), 367–388.

Gutierrez, C., Obloj, T., & Zenger, T. (2022). Pay transparency and productivity. *Bocconi University Management Research Paper, Kelley School of Business Research Paper*, (2022-4341804).

Hackman, D. A., & Farah, M. J. (2009). Socioeconomic status and the developing brain. *Trends in Cognitive Sciences*, *13*(2), 65–73.

Hanspal, T., Weber, A., & Wohlfart, J. (2021). Exposure to the covid-19 stock market crash and its effect on household expectations. *The Review of Economics and Statistics*, *103*(5), 994–1010.

Hardardottir, H. (2017). Long term stability of time preferences and the role of the macroeconomic situation. *Journal of Economic Psychology*, *60*, 21–36.

Harrison, G. W., Hofmeyr, A., Kincaid, H., Monroe, B., Ross, D., Schneider, M., & Swarthout, J. T. (2022). Subjective beliefs and economic preferences during the covid-19 pandemic. *Experimental Economics*, *25*(3), 795–823.

Harsanyi, J. C. (1955). Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of Political Economy*, *63*(4), 309–321.

Harsanyi, J. C. (1975a). Can the maximin principle serve as a basis for morality? a critique of john rawls's theory. *American Political Science Review*, *69*(2), 594–606.

Harsanyi, J. C. (1975b). Nonlinear social welfare functions. *Theory and Decision*, *6*(3), 311–332.

Hassan, T. R., Khalid, W., & Habib, A. (2014). Overconfidence and loss aversion in investment decisions: A study of the impact of gender and age in pakistani perspective. *Research Journal of Finance and Accounting*, *5*(11), 148–157.

Herrnstein, R. J., & Murray, C. (2010). *The bell curve: Intelligence and class structure in american life*. Simon; Schuster.

Hoffman, P. J., Festinger, L., & Lawrence, D. H. (1954). Tendencies toward group comparability in competitive bargaining. *Human Relations*, *7*(2), 141–159.

Hundley, G., & Kim, J. (1997). National culture and the factors affecting perceptions of pay fairness in korea and the united states. *The International Journal of Organizational Analysis*, *5*(4), 325–341.

Ichiue, H., & Nishiguchi, S. (2015). Inflation expectations and consumer spending at the zero bound: Micro evidence. *Economic Inquiry*, *53*(2), 1086–1107.

Jonung, L. (1981). Perceived and expected rates of inflation in sweden. *The American Economic Review*, *71*(5), 961–968.

Jöreskog, K. G., & Moustaki, I. (2001). Factor analysis of ordinal variables: A comparison of three approaches. *Multivariate Behavioral Research, 36*(3), 347–387.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1986a). Fairness and the assumptions of economics. *Journal of Business*, S285–S300.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1986b). Fairness as a constraint on profit seeking: Entitlements in the market. *American Economic Review*, *76*(4), 728–741.

Kahneman, D., Krueger, A. B., Schkade, D. A., Schwarz, N., & Stone, A. A. (2004). A survey method for characterizing daily life experience: The day reconstruction method. *Science*, *306*(5702), 1776–1780.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, Vol.48., No.2*, 263–291.

Kampen, J., & Swyngedouw, M. (2000). The ordinal controversy revisited. *Quality and quantity*, *34*(1), 87–102.

Kasik, L. (2015). *Személyközi problémák és megoldásuk / interpersonal problems and their solutions (in hungarian)*. Gondolat Kiadó.

Keynes, J. M. (1937). The general theory of employment. *The Quarterly Journal of Economics*, *51*(2), 209–223.

Khan, A. K., Peretti, J.-M., & Quratulain, S. (2009). Envy and counterproductive work behaviors: Is more fairness always preferred. *20th AGRH Conference (9th-11th Septembre 2009, Toulouse, France) of the French-speaking HR/OB Research Scholarly Association*.

Khayouti, S., Kiss, H. J., & Horn, D. (2021). Versengeni és együttműködni? egy reprezentatív felmérés tanulságai. *Közgazdasági Szemle*, *68*(9), 966–986.

Kim, J.-O. (1975). Multivariate analysis of ordinal variables. *American Journal of Sociology*, *81*(2), 261–298.

Kim, J., & Park, Y. (2018). Malicious envy and benign envy in organization: The relationships with perceived fairness, self-esteem, and counterproductive work behavior. *Korean Journal of Industrial and Organizational Psychology*, *31*(1), 103–121.

Kim & Leung, K. (2007). Forming and reacting to overall fairness: A cross-cultural comparison. *Organizational Behavior and Human Decision Processes*, *104*(1), 83–95.

Kim, Sung, L.-K., & Park, T.-Y. (2020). Pay differentials, work effort, and social undermining: Role of envy. *Academy of Management Proceedings*, *2020*(1), 22052.

Kőszegi, B., & Rabin, M. (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics*, *121*(4), 1133–1165.

Krupka, E. L., & Stephens Jr, M. (2013). The stability of measured time preferences. *Journal of Economic Behavior & Organization*, *85*, 11–19.

Kuchler, T., & Zafar, B. (2019). Personal experiences and expectations about aggregate outcomes. *The Journal of Finance*, *74*(5), 2491–2542.

Kuhnen, C. M., & Miu, A. C. (2017). Socioeconomic status and learning from financial information. *Journal of Financial Economics*, *124*(2), 349–372.

Lades, L. K., Laffan, K., & Weber, T. O. (2021). Do economic preferences predict pro-environmental behaviour? *Ecological Economics*, *183*, 106977.

Larson, K., Russ, S. A., Nelson, B. B., Olson, L. M., & Halfon, N. (2015). Cognitive ability at kindergarten entry and socioeconomic status. *Pediatrics*, *135*(2), e440–e448.

Levitt, S. D., List, J. A., Neckermann, S., & Sadoff, S. (2016). The behavioralist goes to school: Leveraging behavioral economics to improve educational performance. *American Economic Journal: Economic Policy*, *8*(4), 183–219.

Li, K. K., & Huang, B. (2020). How does the covid-19 outbreak affect macroeconomic expectations? *Available at SSRN 3567937*.

Lin, C.-A., & Bates, T. C. (2021). Who supports redistribution? replicating and refining effects of compassion, malicious envy, and self-interest. *Evolution and Human Behavior*, *42*(2), 140–147.

Loewenstein, G. (2000). Emotions in economic theory and economic behavior. *American Economic Review*, *90*(2), 426–432.

Lombardelli, C. (2003). Public expectations of uk inflation. *Bank of England Quarterly Bulletin*.

Lotz, S., Schlösser, T., Cain, D. M., & Fetchenhauer, D. (2013). The (in) stability of social preferences: Using justice sensitivity to predict when altruism collapses. *Journal of Economic Behavior & Organization*, *93*, 141–148.

Lusardi, A., & Mitchell, O. S. (2011). Financial literacy around the world: An overview. *Journal of Pension Economics & Finance*, *10*(4), 497–508.

Major, B., & Forcey, B. (1985). Social comparisons and pay evaluations: Preferences for same-sex and same-job wage comparisons. *Journal of Experimental Social Psychology*, *21*(4), 393–405.

Malmendier, U., & Nagel, S. (2011). Depression babies: Do macroeconomic experiences affect risk taking? *The Quarterly Journal of Economics*, *126*(1), 373–416.

Malmendier, U., & Nagel, S. (2016). Learning from inflation experiences. *The Quarterly Journal of Economics*, *131*(1), 53–87.

Marescaux, E., De Winne, S., & Rofcanin, Y. (2021). Co-worker reactions to i-deals through the lens of social comparison: The role of fairness and emotions. *Human Relations*, *74*(3), 329–353.

McEvoy, D. M. (2016). Loss aversion and student achievement. *Economics Bulletin*, *36*(3), 1762–1770.

McFadden, D. (2001). Economic choices. *American economic review*, *91*(3), 351–378.

Meier, S., & Sprenger, C. D. (2015). Temporal stability of time preferences. *Review of Economics and Statistics*, *97*(2), 273–286.

Mrkva, K., Johnson, E. J., Gächter, S., & Herrmann, A. (2020). Moderating loss aversion: Loss aversion has moderators, but reports of its death are greatly exaggerated. *Journal of Consumer Psychology*, *30*(3), 407–428.

Neisser, U., Boodoo, G., Bouchard Jr, T. J., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J., et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, *51*(2), 77.

Novemsky, N., & Kahneman, D. (2005). The boundaries of loss aversion. *Journal of Marketing Research*, *42*(2), 119–128.

Okun, A. M. (1981). Prices and quantities: A macroeconomic analysis (brookings institution, washington, dc).

Parrott, W. G., & Smith, R. H. (1993). Distinguishing the experiences of envy and jealousy. *Journal of personality and social psychology*, *64*(6), 906.

Pearlin, L. I., & Schooler, C. (1978). The structure of coping. *Journal of Health and Social Behavior*, 2–21.

Piatek, R., & Pinger, P. (2010). Maintaining (locus of) control? assessing the impact of locus of control on education decisions and wages. *ZEW - Centre for European Economic Research Discussion Paper No. 10-093, Available at SSRN: https://ssrn.com/abstract=1 or http://dx.doi.org/10.2139/ssrn.1739746*.

Premik, F., & Stanisławska, E. (2017). The impact of inflation expectations on polish consumers' spending and saving. *Eastern European Economics*, *55*(1), 3–28.

Rawls, J. (1971). *A theory of justice*. Belknap Press.

Reid, M., Siklos, P., & Du Plessis, S. (2021). What drives household inflation expectations in south africa? demographics and anchoring under inflation targeting. *Economic Systems*, *45*(3), 100878.

Reynaud, A., & Couture, S. (2012). Stability of risk preference measures: Results from a field experiment on french farmers. *Theory and Decision*, *73*, 203–221.

Ripley, B., Venables, B., Bates, D. M., Hornik, K., Gebhardt, A., Firth, D., & Ripley, M. B. (2013). Package 'mass'. *Cran R*, *538*, 113–120.

Robitzsch, A. (2020). Why ordinal variables can (almost) always be treated as continuous variables: Clarifying assumptions of robust continuous and ordinal factor analysis estimation methods. *Frontiers in Education*, *5*, 589965.

Roth, C., & Wohlfart, J. (2020). How do expectations about the macroeconomy affect personal expectations and behavior? *Review of Economics and Statistics*, *102*(4), 731–748.

Roy, S., & Lewis, C. (2023). Examining grade loss aversion–the importance of reference points. *Education Economics*, 1–13.

Salamanca, N. (2018). The dynamic properties of economic preferences. *Melbourne Institute Working Paper No. 04/18, Available at SSRN: https://ssrn.com/abstract=3164693 or http://dx.doi.org/10.2139/ssrn.3164693*.

Schildberg-Hörisch, H. (2018). Are risk preferences stable? *Journal of Economic Perspectives*, *32*(2), 135–154.

Schmidt, U., & Traub, S. (2002). An experimental test of loss aversion. *Journal of Risk and Uncertainty*, *25*, 233–249.

Shachat, J., Walker, M. J., & Wei, L. (2020). The impact of the covid-19 pandemic on economic behaviours and preferences: Experimental evidence from wuhan. *ESI working papers*.

Simon, H. A. (1957). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 99–118.

Simon, H. A. (1966). *Theories of decision-making in economics and behavioural sciences*. Springer.

Smith, A. (2010). *The theory of moral sentiments*. Penguin.

Smith, Shrader, R., White, D. R., Wooten, J., Dogbey, J., Nath, S., O'Hara, M., Xu, N., & Rosenman, R. (2019). Improving student performance through loss aversion. *Scholarship of Teaching and Learning in Psychology*, *5*(4), 278.

Smith, V. L., & Wilson, B. J. (2019). *Humanomics: Moral sentiments and the wealth of nations for the twenty-first century*. Cambridge University Press.

Sugden, R. (1985). Why be consistent? a critical analysis of consistency requirements in choice theory. *Economica*, *52*(206), 167–183.

Sutter, M., Kocher, M. G., Glätzle-Rützler, D., & Trautmann, S. T. (2013). Impatience and uncertainty: Experimental decisions predict adolescents' field behavior. *American Economic Review*, *103*(1), 510–531.

Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics*, *106*(4), 1039–1061.

Vaaramo, M., Huikari, S., Ala-Mursula, L., Miettunen, J., & Korhonen, M. (2024). Temperament traits and economic preferences predict occupational choice beyond human capital variables. *Labour Economics*, 102576.

Vahalík, B., & Staníčková, M. (2016). Key factors of foreign trade competitiveness: Comparison of the eu and brics by factor and cluster analysis. *Society and Economy*, *38*(3), 295–317.

Van Rooij, M. C., Lusardi, A., & Alessie, R. J. (2012). Financial literacy, retirement planning and household wealth. *The Economic Journal*, *122*(560), 449–478.

Varian, H. R. (1973). Equity, envy, and efficiency. *Journal of Economic Theory*.

Varian, H. R. (2014). *Intermediate microeconomics with calculus: A modern approach*. WW Norton & company.

Vellekoop, N., & Wiederholt, M. (2019). Inflation expectations and choices of households.

Von Neumann, J., & Morgenstern, O. (2007). Theory of games and economic behavior: 60th anniversary commemorative edition. In *Theory of games and economic behavior*. Princeton University Press.

Yoshida, Y., & Ito, S. (2022). Average sensitivity of euclidean k-clustering. *Advances in Neural Information Processing Systems*, *35*, 32487–32498.

Zikmund-Fisher, B. J., & Parker, A. M. (1999). Demand for rent-to-own contracts: A behavioral economic explanation. *Journal of Economic Behavior & Organization*, *38*(2), 199–216.

Appendices

4.1 Appendix- Chapter 1

1 Effect of Loss aversion on learning outcomes - Regular points

In this section, we show the regression results of Table 1.3. For the Best 3 tests, the treatment effect is 2.35 points, which accounts for 7.1 percent of the average score (the average was 32.9 points out of 48.

Similarly, in the final test, the treatment effect for losing points was 3.1 points, while the average score was 25.5 out of 40; that is 12.1 percent of the average. We argue that these effects are great, considering that implementing these treatments is cheap, in the sense that no alteration of the materials and teaching methodology is required.

Table 4.7: Regression results for tests written throughout the semester and the Final Test in regular points

	Test 1	Test 2	Test 3	Test 4	Best 3 Tests	Final Test	
Losing Points	-0.270	1.066***	0.700	1.212***	2.358**	3.164**	3.128***
	(0.556)	(0.124)	(0.419)	(0.197)	(0.927)	(1.142)	(0.633)
Female	0.878**	0.924**	1.182*	0.717	2.857**	0.967	-0.390
	(0.292)	(0.301)	(0.621)	(0.577)	(1.221)	(1.189)	(0.706)
Best 3 Tests							0.465***
							(0.0945)
Constant	10.18***	10.68***	10.33***	9.557***	33.43***	23.31***	6.700*
	(0.561)	(0.480)	(1.019)	(0.878)	(1.726)	(2.553)	(3.380)
Observations	321	321	321	321	321	321	321
R-squared	0.154	0.172	0.124	0.094	0.162	0.160	0.351
Residual Std. Error	3.693	3.855	4.220	5.026	9.611	9.395	8.269

Standard errors in parentheses, clustered by seminar group;* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Control variables include: mother's university education, knowing how to take derivatives,working or not, and teacher, practice session and campus fixed effects. Note that *Losing points (!)* correspond to students losing points during the semester tests and during the final test, respectively.

2 Treatment Effects without control variables

In this section, as a robustness check, we run the two important regressions (comparing losings points vs. gaining points and comparing hybrid and loss treatments to analyze the novelty-effect) without using control variables. Below we show that not including control variables increases the estimated effect significantly in the former – that is, the estimated difference between losing and gaining points is at least 11.86 percent.

As for the novelty effect, not controlling for the other variables decreases the difference between losing points throughout the semester and losing points just for the final test to the point that even if the difference is statistically significant, economically the significance is marginal.

Table 4.8: Comparing treatments who were losing points vs. gaining points – without control variables

	Test 1	Test 2	Test 3	Test 4	Best 3 Tests	Final Test	
Losing Points	2.722	10.71**	12.30**	13.14***	9.487**	15.62**	11.86**
	(5.947)	(4.433)	(4.348)	(3.427)	(3.291)	(5.814)	(3.898)
Best 3 Tests							0.547***
							(0.124)
Constant	62.02***	54.17***	59.41***	53.91***	64.87***	53.41***	18.33*
	(3.074)	(3.210)	(2.743)	(1.818)	(2.364)	(5.017)	(8.570)
Observations	321	321	321	321	321	321	321
R-squared	0.003	0.041	0.048	0.040	0.047	0.087	0.300
Residual Std. Error	24.586	25.441	26.980	31.721	20.944	> 24.016	21.058

Standard errors in parentheses, clustered by practice-group;* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Note that *Losing points (!)* correspond to students losing points during the semester tests and during the final test, respectively.

Table 4.9: Testing the Novelty Effect – Comparing Hybrid and Loss treatments without control variables

	Test 1	Test 2	Test 3	Test 4	Best 3 Tests	Final Test	
Loss	-0.00347	6.321	14.06***	14.44**	8.478	2.255	-0.271
	(5.773)	(5.989)	(3.876)	(4.600)	(4.631)	(2.794)	(2.388)
Best 3 Tests							0.298*
							(0.152)
Constant	64.74***	58.57***	57.65***	52.61***	65.88***	67.68***	48.05***
	(3.768)	(5.134)	(2.768)	(3.375)	(4.294)	(2.303)	(11.26)
Observations	212	212	212	212	212	212	212
R-squared	0.000	0.016	0.064	0.051	0.043	0.003	0.075
Residual Std. Error	24.028	24.722	26.376	30.724	19.759	21.873	21.116

Standard errors in parentheses, clustered by practice-group;* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Note that *Losing points (!)* correspond to students losing points during the semester tests and during the final test, respectively.

3 Regressions excluding dropouts

As mentioned in the main text, we run robustness checks with filtering out students scoring zero points at the end of the semester - thus, they are likely dropouts. In our sample, 30 students scored zero points in the Final Test - only five of those were losing points during the semester[15]. Running a Linear Probability Model on dropouts shows that losing points is a significant predictor of not being a dropout, decreasing the probability of scoring zero between 5.4 - 8.7 percentage, as per Table 4.10

Table 4.10: Linear Probability Models for dropouts measured by scoring zero in the final test

	Dropout = Scoring zero on Final Test			
Losing Points	-0.0872**	-0.0545**		
	(0.0389)	(0.0229)		
Loss Treatment			-0.132**	-0.0638
			(0.0529)	(0.0356)
Hybrid Treatment			-0.101	-0.0243
			(0.0665)	(0.0567)
Female	-0.0429	-0.0269	-0.0390	-0.0271
	(0.0290)	(0.0263)	(0.0265)	(0.0264)
Constant	0.151***	0.0815	0.174***	0.0912
	(0.0337)	(0.0591)	(0.0421)	(0.0728)
Controls	NO	YES	NO	YES
Observations	321	321	321	321
R-squared	0.028	0.089	0.041	0.090
Residual Std. Error	0.288	0.284	0.286	0.284

Standard errors in parentheses, clustered by practice-group;* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Control variables include: female, mother's university education, knowing how to take derivatives, working or not, and teacher, practice session and campus fixed effects. Note that *Losing points (!)* correspond to students losing points during the semester tests and during the final test, respectively.

We do this robustness check to make sure that the effect is not rooted in the difference between dropout rates. We can expect that the treatment is going to decrease, as part of the mechanism of loss aversion is students do not want to ”lose” their 40 points (40 percent of the grade!). Table 4.11 shows the estimates for OLS regressions without dropouts. Compared to the results of Table 1.3, the effect becomes not statistically significant for the

[15] With 19 of them being in the control group, 5 of them in the ”loss” group and 6 in the ”hybrid” group

tests written during the semester; for the Final Test, however, the effect is still significant, although - as expected - decreased to 5 -5.5 percent.

Table 4.11: Analysis of practice group tests and Final Test after filtering out dropouts

	Test 1	Test 2	Test 3	Test 4	Best 3 Tests	Final Test	
Losing Points	-5.074	3.596***	1.876	4.504**	1.651	5.286***	5.685***
	(3.391)	(0.948)	(2.273)	(1.879)	(1.985)	(0.882)	(0.814)
Female	5.485**	5.778**	7.388*	4.481	5.953**	2.418	-0.974
	(1.826)	(1.882)	(3.879)	(3.608)	(2.543)	(2.973)	(1.765)
Best 3 Tests							0.558***
							(0.113)
Constant	66.29***	69.75***	69.09***	64.39***	73.49***	62.57***	49.46***
	(2.988)	(2.522)	(5.120)	(6.524)	(2.635)	(2.833)	(7.124)
Observations	291	291	291	291	291	291	291
R-squared	0.156	0.148	0.103	0.075	0.138	0.132	0.167
Residual Std. Error	21.412	23.223	24.082	28.700	17.343	14.509	14.237

Standard errors in parentheses, clustered by seminar group;* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Control variables include: mother's university education, knowing how to take derivatives, working or not, and teacher, practice-session and campus fixed effects. Note that *Losing points (!)* correspond to students losing points during the semester tests and during the final test, respectively.

Similarly, we wanted to check for the robustness of heterogeneous gender effects. There were a bit more female students who scored zero points of the Final Test (18 vs. 12 males), although the Linear Probability Model suggests that gender is not a significant predictor. Table 4.12 contains the same regression specification as Table 1.6 but without the dropouts. The coefficients of the interaction term decreased compared to Table 1.6, and in the case of the Final Test, they became positive (but statistically not significant from zero). Same as in our analysis of the main text, we find no evidence for statistically significant heterogeneous treatment effects on gender.

Table 4.12: Testing gender-heterogeneity of loss aversion without outliers

	Test 1	Test 2	Test 3	Test 4	Best 3 Tests	Final Test	
Losing Points	-2.647	7.531**	2.501	4.613	3.977	5.234*	5.553**
	(4.239)	(2.916)	(5.028)	(3.887)	(3.167)	(2.702)	(2.389)
Female	7.168***	8.482***	5.732	1.990	6.472**	0.416	-0.484
	(1.849)	(1.977)	(4.614)	(4.902)	(2.634)	(4.263)	(3.856)
Female x Losing Points	-4.353*	-7.055	-1.120	-0.196	-4.171	0.101	0.255
	(2.217)	(4.184)	(6.788)	(4.829)	(3.449)	(4.322)	(3.840)
Best 3 Tests							0.169*
							(0.0807)
Constant	65.41***	68.32***	68.86***	64.35***	72.65***	62.60***	49.53***
	(3.328)	(2.656)	(5.822)	(7.388)	(3.042)	(3.196)	(6.584)
Observations	291	291	291	291	291	291	291
R-squared	0.159	0.153	0.103	0.075	0.142	0.132	0.167
Residual Std. Error	21.423	23.199	24.124	28.752	17.344	14.535	14.263

Standard errors in parentheses, clustered by seminar group;* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Control variables include: mother's university education, knowing how to take derivatives,working or not, and teacher, practice session and campus fixed effects.

4 Regressions with all available test scores

As mentioned in the main text, we asked students to provide some background information on themselves, with the addition to consent using their personal data in the study. However, we wanted to make sure that sampling bias does not affect our results. First, we ran a simplified regression on the final test results for all students (n = 461 after filtering out those who scored 0 point throughout the semester) and our sample. The results are shown in Table 4.13; running a simplified regression on all students slightly increases the estimated coefficients, but doe not change the results. Important to note, however, that the coefficient of the Hybrid treatment (i.e., the ones who were only losing at the Final Test) becomes lower by 0.7 points, which is 1.75 percent of the total score. This suggests some selection bias in this treatment, as we might lose some high-performers in the Hybrid treatment.

This could explain the differences observed between Tables 1.3 and 1.4. However, if we want to test the novelty effect, this would cause us the report a bigger difference between the Hybrid and Loss treatments in the Final Test. Thus, still getting an insignificant difference between the two treatments for the Final Test strengthens our results (and explains why adding students' performance during the semester mitigates the difference between the two groups). Additionally, we ran two sample t-tests for all test scores to compare all students with our sample, and we could not reject the null hypothesis of the means being equal to zero at the five percent significance level for either score.

Table 4.13: Testing selection bias: treatment effects of points scored on the Final Test

	Dependent variable:	
	Final Test	
	(All students)	(Our sample)
Loss	3.873***	3.523***
	(0.681)	(0.853)
Hybrid	4.203***	3.527***
	(1.112)	(1.351)
Best 3 Tests	0.475***	0.473***
	(0.080)	(0.105)
Constant	7.495***	7.424**
	(2.879)	(3.570)
Observations	461	321
R^2	0.318	0.341
Adjusted R^2	0.303	0.320
Residual Std. Error	8.971	8.279
F Statistic	20.972***	16.050***

Standard errors in parentheses, clustered by practice group;* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Clustered standard errors in parantheses. Control variables include practice-group and teacher fixed-effects.

5 Quantile Regression Results

Table 4.14: Quantile Regression results

	Final Test			
Quantiles	All data		No dropouts	
Q.15	-1.2379	-1.2380	0.0000	0.0000
	(10.2371)	(13.5021)	(3.2949)	(3.3716)
Q.20	1.5831	1.5844	3.7642	3.1072
	(3.5871)	(6.7790)	(2.8745)	(3.1848)
Q.25	3.8654	3.8592	2.5829	2.0589
	(2.4706)	(4.9799)	(2.6772)	(2.6854)
Q.30	2.2435	2.2517	3.8920	3.2424
	(2.6824)	(3.0857)	(3.1582)	(2.9406)
Q.35	5.0339	5.0315*	6.4887*	6.0927*
	(2.6816)	(2.3677)	(3.2349)	(3.0696)
Q.40	6.8023*	6.8000**	7.8725*	6.6615*
	(3.0441)	(2.3582)	(3.5001)	(3.1306)
Q.45	7.5018**	7.5000***	7.8341*	7.5000**
	(2.4650)	(2.0379)	(3.5864)	(2.6586)
Q.50	8.3130**	8.3293***	8.7981*	8.5747***
	(2.8518)	(1.8383)	(3.5478)	(2.4107)
Q.55	8.6288**	8.6312***	8.6056*	8.3936**
	(3.0375)	(2.4105)	(3.5670)	(3.2208)
Q.60	8.6586*	9.2720**	9.0534*	8.7034*
	(3.4754)	(2.9814)	(4.1619)	(3.8503)
Q.65	8.9695*	8.9773*	8.8361	7.4335
	(3.8149)	(3.6292)	(5.0287)	(4.5293)
Q.70	8.9915	8.9887**	8.6820	8.6575
	(4.8632)	(3.1377)	(5.1514)	(4.7752)
Q.75	6.8805	6.8983	7.0985	6.9959
	(4.2460)	(3.6388)	(5.3086)	(5.0052)
Q.80	8.0764	8.0773	5.4413	5.5476
	(4.6313)	(4.2720)	(4.7630)	(4.7841)
Q.85	6.3737	6.3798	6.0024	6.0303
	(4.6294)	(4.2661)	(3.9296)	(4.4277)
Q.90	7.3439*	7.3564*	6.5990*	6.7619
	(3.1392)	(2.9294)	(3.1783)	(3.9954)
Control Variables	YES	YES	YES	YES
3 Best Tests	YES	NO	YES	NO
N	321	321	291	291

6 Final Questionnaire and Teacher Evaluation

After the final test, we asked students to fill out our second questionnaire mentioned in section 1.2, where they provided us with some feedback to the course, about the treatments and their perceived fairness, and the hours put in the course. Unfortunately, less than a third of our sample sent back the questionnaire. With missing values filtered out, our sample size was reduced to 101, so the interpretations of these results are limited. Most of the questionnaires sent back are from individuals from the loss treatment (n= 50), while the hybrid and gain groups consisting of 27 and 24 respondents, respectively. Nonetheless, we think that the answers are useful to analyse on the mechanisms behind the effect identified.

First, we asked students how motivated they were to study each week on a 1-7 likert-scale. We created a dummy variable for being motivated if their answer was four or above. One thing is to be motivated, and another is to actually put in the effort. In our next question, we asked students to approximate how many minutes did they spend studying each week. Finally, we asked them whether they knew about the other treatments, and whether they discussed it among their peers, to account for contamination effect.

Table 4.15 shows the regression results for the final test scores. Due to sample size issues, we elected to not include the dummies for teachers, only the time-slot for the practice groups[16]. While the effect sizes of the treatment are different compared to regression (1), the increase in effect is not drastic. However, the significance dropped - which can be caused by the selection of the students actually choosing to fill out such a questionnaire. The key result, however, is that the effect size increased once we take into account motivation, knowledge on other groups, and minutes studied in each week (with the note that the last one being self-reported is less reliable).

At the end of the semester, students also have the opportunity to evaluate their teachers and give feedback in several aspects in a formal way through a platform called MyView[17]. On the last course occasion, a questionnaire can be filled out with questions ranging from the materials available, the teachers' merits, etc.. Important to note that here, we classified "hybrid" as a "gain" treatment, as up to the deadline, these students did not encounter losing points.

The results of this feedback are available to us, aggregated on the practice group levels, with average score of the practice group leader (ranging from 1 to 5), and with the percent of students filling out the questionnaire. Our hypothesis is that if students felt that losing

[16] We ran the regressions with several specifications, and it did not change the outcome significantly.

[17] For more information on MyView, visit the official announcement: https://www.uni-corvinus.hu/post/hir/launching-myview-the-new-student-feedback-system/?lang=en

Table 4.15: Regression Results for final test using data from the final questionnaire

	Dependent variable:			
	Final Test			
	(1)	(2)	(3)	(4)
Loss	2.219^{*}	2.701^{**}	2.714^{**}	2.783^{**}
	(1.202)	(1.375)	(1.360)	(1.352)
Hybrid	3.185^{*}	3.597^{**}	3.624^{*}	3.745^{**}
	(1.651)	(1.832)	(1.861)	(1.739)
Female	-0.316	-0.266	-0.276	-0.255
	(0.840)	(0.858)	(0.822)	(0.916)
Best 3 Tests	0.289^{**}	0.292^{**}	0.286^{**}	0.295^{**}
	(0.116)	(0.120)	(0.124)	(0.118)
Knew about other groups		-1.191	-1.214	-1.231
		(1.263)	(1.286)	(1.228)
Motivated			0.310	
			(1.491)	
Minutes studied per week				-0.008^{*}
				(0.005)
Constant	15.594^{***}	16.013^{***}	15.996^{***}	16.715^{***}
	(4.088)	(4.346)	(4.373)	(4.552)
Observations	101	101	101	101
R^2	0.222	0.231	0.231	0.250
Adjusted R^2	0.164	0.164	0.155	0.176
Residual Std. Error	4.909	4.910	4.935	4.874
F Statistic	3.797^{***}	3.445^{***}	3.039^{***}	3.370^{***}

Clustered Standard errors in parentheses;* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Control variables include practice-group and teacher fixed-effects.

points was more unfair towards them, they would fill out this questionnaire, and would criticize the teacher, thus, a) more students would fill it out and b) would give a lower score). However, running t-tests on both measures between loss and gain groups suggests that we cannot reject the null hypothesis of the average scores being equal in the two treatments. That is: we find no evidence that students were more discontent with the course in the loss treatment. We ran the same test on the amount of feedback given, and found similar results.

7 Power Calculations and Non-inferiority Tests

Power analysis is used to determine the sample size required to detect an effect of a given size with a specified level of confidence. In this section, we derive the relationship between the minimum detectable effect size (MDE) and sample size when comparing two independent group means. Specifically, we test the MDE on the Final Test scores, where the average score was 63.72% and the standard deviation was 25.097. Our observations with the least populous treatment was 84 (Hybrid treatment).

Definitions and Assumptions

We define the following parameters:

- **Significance level (Type I error):** $(\alpha = 0.05)$
- **Power (1 - Type II error):** $(1-\beta = 0.80)$
- **Standard deviation of the pooled sample:** $(\sigma = 25.097)$
- **Sample size per group:** (n)
- **Minimum detectable effect size:** (Δ)

The test statistic for comparing two independent means is given by:

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sigma_{\bar{X}_1 - \bar{X}_2}}$$

where the standard error of the difference in means is:

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sigma\sqrt{\frac{2}{n}}$$

Power Calculation – Minimum Detectable Effect Size

The power of the test is determined by the ability to detect a true difference of (Δ). This means that we require:

$$Z_{1-\alpha/2} + Z_{1-\beta} = \frac{\Delta}{\sigma\sqrt{2/n}}$$

where: $(Z_{1-\alpha/2})$ is the critical value corresponding to $(\alpha = 0.05)$ (1.96 for a two-tailed test) and $(Z_{1-\beta})$ is the critical value corresponding to $(1-\beta = 0.80)$ (0.84)

Substituting these values:

$$1.96 + 0.84 = \frac{\Delta}{0.45\sqrt{2/n}}$$

which simplifies to:

$$2.80 = \frac{\Delta\sqrt{n}}{0.45\sqrt{2}}$$

Solving for (n):

$$n = \frac{(2.80 \times 0.45\sqrt{2})^2}{\Delta^2}$$

Rearranging the above equation to express the minimum detectable effect size Δ in terms of (n):

$$\Delta = \frac{2.80 \times 0.45\sqrt{2}}{\sqrt{n}}$$

This equation shows that the minimum detectable effect size is inversely proportional to the square root of the sample size. That is, increasing the sample size decreases the minimum detectable effect size, allowing for the detection of smaller differences between group means. The following plot illustrates the relationship between sample size (n) and minimum detectable effect size Δ:

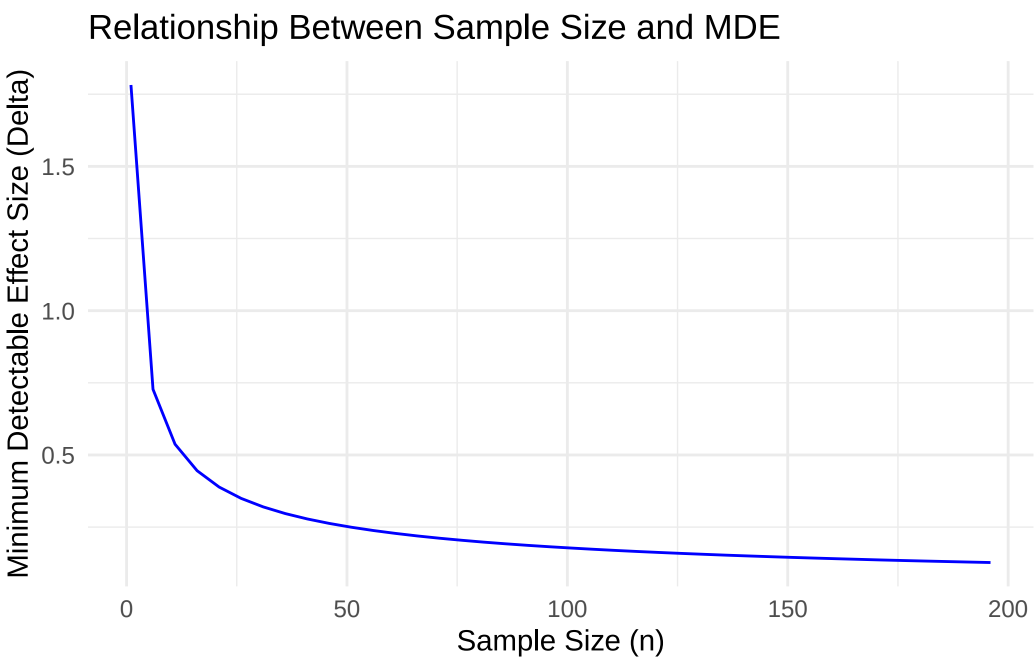

Figure 4.5: Power Calculation: Relationship Between Sample Size and MInimal Detectable Effect Size

Power Calculation and Non-inferiority tests

One of our main findings in the paper is a null result, that is: we find no differences between the Hybrid and Loss treatments when looking at the Final Test scores. However, this result might come from the fact that we have a small sample available to us in order to find a statistically significant difference.

Indeed, running a one-sided power-calculation for $\Delta = 5.48$ with $\sigma = 25.097$ suggests that we have low power (0.4). However, as we argue in the analysis, due to the treatment also taking effect during the semester, thus we have to control for tests scores during the semester. That decreases the treatment effect size to 2.128 percent. Running the power calculation with $\Delta = 2.128$ with $\sigma = 25.097$ and $\beta = 0.8$, we would need n = 1720 in order to detect the effect. And even if it is still significant, the effect size is big enough that

As an alternative, I run a non-inferiority test, where I test whether the estimated effect ($\hat{\beta}$) is not worse than a predefined margin (Δ) compared to zero effect, where the test statistic is given by:

$$T = \frac{\hat{\beta} - \Delta}{SE}$$

Testing for $\hat{\beta} = 5.484$, $SE = 4.939$, $\Delta = 5$, we get $T = 0.098$ ($p = 0.461$), leading us to reject the null hypothesis of non-inferiority. Therefore, we cannot conclude that the Hybrid treatment is non-inferior to the Loss treatment.

Another null result is that we find no evidence of gender heterogeneity, contrary to Apostolova-Mihaylova et al. (2015). To test for non-inferiority, we chose the lowest effect size found in their study, which is $\Delta = -6.7 as per Table 5 of the original study.$

For parameters, we chose specification (5) in Table 6 for Best 3 Tests, as it has the highest estimated effect with lowest standard errors. With the parameters $\hat{\beta} = -5.9$, $SE = 4.901$, $\Delta = -6.7$, we get $T = -0.163$ $(p = 0.435)$. thus cannot conclude that the gender effect is non-inferior to the predefined margin.

Additionally, we argue that in our regression estimates, the coefficients for $\widehat{\beta_{Female}}$ and $\widehat{\beta_{FemalexLoss}}$ work in opposite directions to the point where even if we have enough statistical power to show that effect sizes themselves are statistically significant, it would not be a big effect (the greatest estimated difference being 2.5% (with the average being around 60%).

8 Introduction to the students during the first lecture[18]

Dear Students,

I am Antal Ertl, a second-year PhD student at Corvinus University of Budapest, and I would like to ask for your help with our research.

This semester, the students of the course are participating in a teaching methodology research project, where we are examining the effectiveness of various teaching and assessment methods. Related to this, a point system specific to each practice group has been uploaded, which can be found in the syllabus.

The study material and the evaluation process, which serve as the basis for assessment, are the same for all practice groups. We emphasize that all practice groups will receive the same material, the same number of questions from the same question bank, and each content element will be worth the same number of points.

Both the Student Government and the Ethics Committee of Corvinus University have been informed about the research. No student will face any disadvantage during the evaluation process.

We kindly ask you to contribute to our research by filling out our questionnaire link. Only the researchers involved in the study will have access to the content of the questionnaire, which will be handled anonymously and will not affect your final grade for the course.

If you have any questions during the semester, please follow the instructions below: For professional topics, please contact your practice group leader directly.

For any questions regarding the grading of assignments in Moodle, please send an email to bce.phd.kutatas@gmail.com. In your message, please make sure to include your practice group number!

Thank you for your help!

[18] We held an information session during the first lecture to the students participating in the course. 5 minutes were also allocated from the lecture to students to complete our questionnaire, and give their consent to use their data in our research. The following text was read to them during the lecture in Hungarian; additionally, we made the text available to them on the Moodle study-site.

9 First Questionnaire

Introduction

Dear Student! This questionnaire is part of a pedagogical research within the framework of the Macroeconomics course, during which we examine the effectiveness of various teaching and assessment methods. The following questionnaire contains questions related to your university studies, mainly concerning the Macroeconomics course. In addition, we are interested in some personal information to better interpret your responses. Completing the questionnaire takes approximately 5 minutes, and you can stop at any time and return to it if you save the link. Participation is voluntary and anonymous. The data will only be analyzed in aggregate form. Neither the course instructors nor any third party can access your answers. Please help us with your responses! If you have any questions about the research, please contact us at `bce.phd.kutatas@gmail.com`.

Demographics

Q1. Gender:

- Female
- Male
- Other, prefer not to answer

Q2. Year of birth: _____

Q3. What type of settlement did you grow up in (where you spent most of your time before university)?

- Capital city
- County seat, city with county rights
- City
- Town, large village
- Foreign settlement

Q4. What is your mother's highest level of education?

- Up to 8 years of primary school

- Vocational school, trade school (without high school diploma)
- High school giving a diploma (grammar school, vocational high school)
- College, university
- Don’t know, didn’t know her, or she is no longer living

Q5. What is your father’s highest level of education?

- Up to 8 years of primary school
- Vocational school, trade school (without high school diploma)
- High school giving a diploma (grammar school, vocational high school)
- College, university
- Don’t know, didn’t know him, or he is no longer living

Q6. Do you currently live in a dormitory?

- Yes
- No

Q7. Are you currently receiving any social scholarship/support?

- Yes
- No

Q8. Are you a member of any student organization or special college?

- Yes
- No

Information about University Studies

Q9. What is your major? _____

Q10. In what form of financing are you currently studying in this major?

- Corvinus scholarship
- Tuition fee, self-financing
- Public service scholarship

Q11. What year are you in this major?

- First year
- Second year
- Third year

Q12. How many passive semesters have you had so far in this major? _____

Q13. How many courses have you taken in total this semester? _____

Q14. How many credits do these courses total? _____

Q15. How many of these courses are mandatory? _____

Q16. How many time slots do you have classes in on average per week according to the schedule? (One time slot refers to a 90-minute class.) _____

Q17. Are you studying another major simultaneously with your current one?

- No
- Yes, but I have a passive semester in that major
- Yes, and I have an active semester in that major as well

Q18. Are you working or planning to work this semester alongside university?

- No
- Yes, part-time
- Yes, full-time

Questions Regarding the Macroeconomics Course

Q19. What category does the Macroeconomics course fall under for your studies?

- Mandatory course
- Mandatory elective course
- Freely elective course

Q20. How many times have you taken the Macroeconomics course?

- First time
- Second time
- Third or more times

Q21. If this is your second or more time, what is the reason for retaking it? _____

Q22. How interested are you in the Macroeconomics course? Please rate on a scale from 1 to 5, where 1 means not interested at all, and 5 means very interested.

Q23. What grade do you hope to achieve in the Macroeconomics course?

- 1
- 2
- 3
- 4
- 5

General Preferences

Q24. How willing are you to take risks or avoid risks? Please answer on a scale from 0 to 10, where 0 means ”not willing to take risks at all” and 10 means ”very willing to take risks.”

Q25. Suppose you receive 10,000 HUF today. What is the minimum amount you would ask for a week later to refuse this 10,000 HUF today?

10 Final Questionaire

Introduction

Dear Student! This questionnaire is part of a pedagogical research conducted within the framework of the Macroeconomics course, during which we examine the effectiveness of various teaching and assessment methods. The following questionnaire is related to assessments and study experiences, so we can gain a more accurate picture of the experiment and your experiences. Completing the questionnaire takes about 3 minutes, and you can stop at any time and return to it if you save the link. Participation is voluntary and anonymous. The data will only be analyzed in aggregate form. Neither the course instructors nor any third party can access your answers. Please help us with your responses! If you have any questions regarding the research, please contact us at `bce.phd.kutatas@gmail.com`.

Personal Information

Q1. Please provide your Neptun code (in uppercase): This is necessary so that we can match your previously provided data with your responses to this questionnaire. ____________________

Experiment-Related Questions

Q2. During the experiment, some groups were graded differently. How much did you communicate about this with participants from other groups?

- I was aware of it and talked to others about it.
- I was not aware of it.

Q3. On average, how many minutes per week did you spend preparing for the classes (including homework and tests)? ________________________

Q4. On a scale from 1 to 7, where 7 means fully motivated and 1 means not at all, how motivated were you to study week by week?

Final Remarks

Q.5 If there is anything important that we haven't covered or if you have any comments regarding the questionnaire, please write it here: ________________________ ________________________

4.2 Appendix - Chapter 3

1 Histograms for Preferences

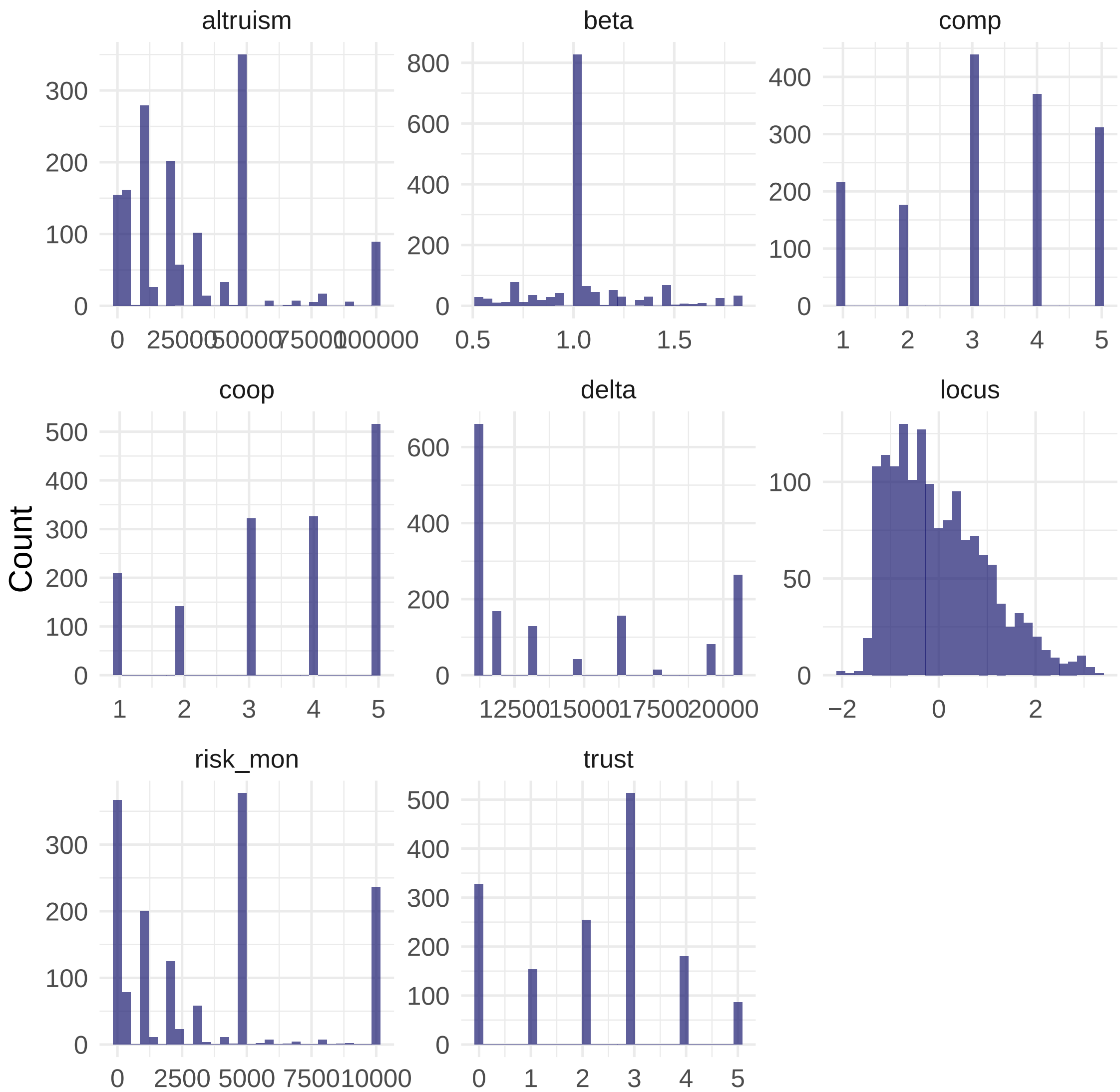

Figure 4.6: Histograms of measured economic preferences

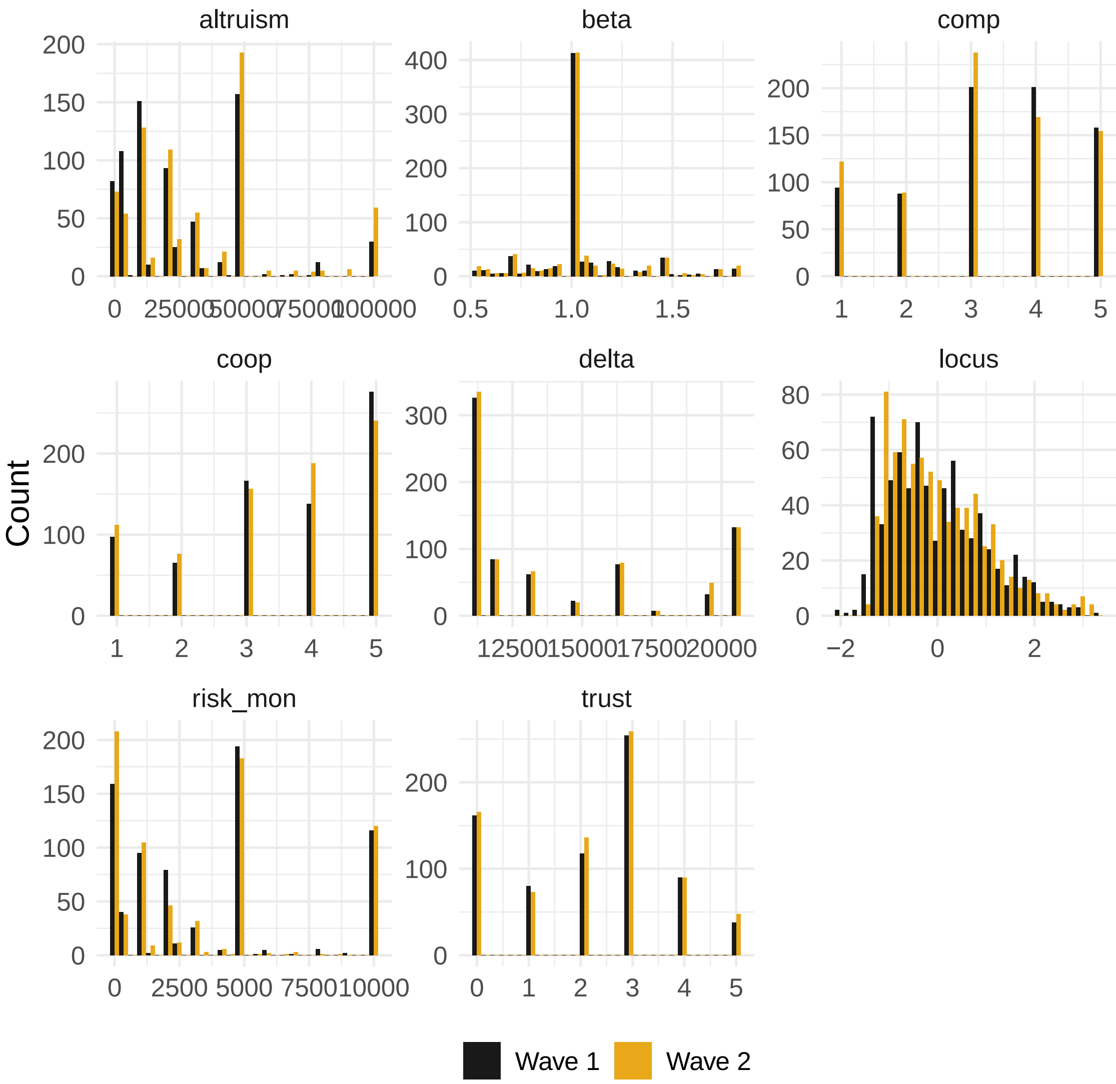

Figure 4.7: Histograms of measured economic preferences by wave

2 Filtering of variables for Wave 1 and Wave 2

Table 4.16: Comparison of removing missing values for preferences - Wave 1

Variable	Restricted, N=742	Unrestricted, N=1.025	Test_results
age			
Mean (SD)	51 (17)	52 (17)	0.0215 (Wilcoxon rank sum test)
hh_size			
Mean (SD)	2.55 (1.28)	2.45 (1.25)	0.0654 (Wilcoxon rank sum test)
Female	379 (51%)	526 (51%)	
Male	363 (49%)	499 (49%)	0.9210 (Pearson's Chi-squared test)
netwage			
Mean (SD)	181,746 (206,758)	176,715 (192,346)	0.6489 (Wilcoxon rank sum test)
Less than secondary school	270 (36%)	384 (37%)	
Secondary school	304 (41%)	412 (40%)	
University Degree	168 (23%)	229 (22%)	0.8978 (Pearson's Chi-squared test)

Note: in the restricted sample, 269, in the unrestricted sample, 445 missing values were found for netwage

Table 4.17: Comparison of removing missing values for preferences - Wave 2

Variable	Restricted, N=772	Unrestricted, N=1.013	Test_results
age			
Mean (SD)	51 (17)	52 (17)	0.0940 (Wilcoxon rank sum test)
hh_size			
Mean (SD)	2.73 (1.40)	2.68 (1.40)	0.4060 (Wilcoxon rank sum test)
Female	407 (53%)	526 (52%)	
Male	365 (47%)	487 (48%)	0.7389 (Pearson's Chi-squared test)
netwage			
Mean (SD)	191,570 (158,646)	186,159 (151,999)	0.4335 (Wilcoxon rank sum test)
Less than secondary school	250 (32%)	329 (32%)	
Secondary school	331 (43%)	434 (43%)	
University Degree	191 (25%)	250 (25%)	0.9990 (Pearson's Chi-squared test)

Note: in the restricted sample, 235, in the unrestricted sample, 378 missing values were found for netwage.

3 Results for Hierarchical Clusters

Analysing the results of hierarchical clustering , however, do not provide us with such implications. While there are differences is measured preferences, the significance varies between waves. Important to note that Cluster 2 for Wave 1 only consists of 45 observations, suggesting that this group consists of outliers. This view is strengthened by looking at the structure of the cluster: individuals are more risk-taking, more altruistic, but are less cooperative, and have a significantly lower discount factor. Similar patterns emerge in Wave 2, although with more moderate differences for risk and time-preferences (which might be an artifact of the higher number of observations lowering the averages). For the background variables - with the exception of education level in Wave 1 - no significant differences can be observed. I argue that these groupings - while stable - are less meaningful compared to the ones found with k-medoid clustering; therefore, I will proceed to analyse cluster stability based on that method.

Table 4.18: Comparison of clusters for Wave 1 and Wave 2, using Hierarchical clustering with k = 2

	Wave 1			Wave 2		
Variable	Cluster 1, N=697	Cluster 2, N=45	p-value	Cluster 1, N=547	Cluster 2, N=225	p-value
Risk	3,395	6,067	<0.001	2,831	4,541	<0.001
	(3,344)	(3,653)		(3,137)	(3,917)	
Trust	2.20	2.07	0.5	2.11	2.52	0.001
	(1.49)	(1.53)		(1.56)	(1.34)	
Cooperation	3.62	2.93	0.005	3.55	3.31	0.013
	(1.37)	(1.62)		(1.40)	(1.36)	
Altruism	24,529	42,422	<0.001	32,114	32,688	0.6
	(23,657)	(33,782)		(26,636)	(29,769)	
Competition	3.32	3.38	0.7	3.11	3.37	0.027
	(1.28)	(1.40)		(1.37)	(1.16)	
Time-inconsistency	1.01	1.57	<0.001	0.96	1.24	<0.001
	(0.19)	(0.26)		(0.17)	(0.30)	
Discount Factor	14,314	11,579	<0.001	15,230	11,804	<0.001
	(3,725)	(898)		(3,914)	(1,237)	
Locus of Control	-0.01	0.13	0.8	-0.16	0.38	<0.001
	(0.98)	(1.32)		(0.99)	(0.93)	
Age			0.2			0.6
Household Size	2.56	2.38	0.4	2.73	2.72	0.7
	(1.29)	(1.13)		(1.40)	(1.42)	
Net Wage	181,109	191,150	0.7	187,834	199,634	0.7
	(208,197)	(187,144)		(141,989)	(189,913)	
Unknown Wage	254	15		180	55	
Education Level:			0.039			0.3
Less than secondary school	256 (37%)	14 (31%)		170 (31%)	80 (36%)	
Secondary school	278 (40%)	26 (58%)		234 (43%)	97 (43%)	
University Degree	163 (23%)	5 (11%)		143 (26%)	48 (21%)	
Gender:			0.4			0.2
Female	353 (51%)	26 (58%)		297 (54%)	110 (49%)	
Male	344 (49%)	19 (42%)		250 (46%)	115 (51%)	
Health Concerns	2.91	2.93	0.5	2.82	2.84	0.5
	(0.66)	(0.69)		(0.73)	(0.65)	
Family Health Concerns	2.98	3.18	0.5	2.86	2.80	0.8
	(0.82)	(1.43)		(1.05)	(0.75)	
Financial Concerns	2.65	2.78	0.6	2.54	2.60	0.7
	(0.93)	(1.43)		(0.82)	(0.95)	

4 Results for Mixed Clustering

One of the key assumptions in the main analysis is the use of Euclidean-distance as the distance measure for the clustering algorithm. To relax this assumption, I run the analysis with the k-prototype algorithm. Similarly to the k-medoid, it is a partitioning algorithm, however, it is fit to create clusters on mixed databases by combining Euclidean distance in the case of continuous and Gower-distance for discrete variables.

Here, we find that the differences between economic preferences by clusters are more nuanced than in the main analysis. However, even though the differences are not always significant, the direction of the differences are the same for Cluster 2 in Wave 1 and Cluster 1 in Wave 2. The main differences come from Trust, Altruism and Time preferences. Also worth noting that these clusters are smaller in number compared to their counterparts, and people tend to be relatively older, lower educated, and lower earners.

Table 4.19: Comparison of clusters for Wave 1 and Wave 2, using K-prototype clustering

	Wave 1			Wave 2		
Variable	Cluster 1, N = 508	Cluster 2, N = 234	p-value	Cluster 1, N = 265	Cluster 2, N = 507	p-value
Risk	3,521 (3,425)	3,637 (3,419)	0.6	3,481 (3,397)	3,250 (3,506)	0.2
Trust	2.30 (1.45)	1.97 (1.57)	0.012	2.06 (1.54)	2.32 (1.48)	0.026
Cooperation	3.62 (1.37)	3.49 (1.45)	0.3	3.53 (1.42)	3.45 (1.38)	0.3
Altruism	22,992 (23,833)	31,308 (25,726)	<0.001	33,445 (26,323)	31,673 (28,203)	0.13
Competition	3.37 (1.24)	3.24 (1.37)	0.3	3.08 (1.42)	3.24 (1.25)	0.2
Time-inconsistency	1.12 (0.23)	0.87 (0.16)	<0.001	0.86 (0.17)	1.13 (0.24)	<0.001
Discount Factor	11,874 (1,255)	19,085 (1,942)	<0.001	18,957 (1,947)	11,761 (1,059)	<0.001
Locus of Control	-0.03 (0.96)	0.06 (1.07)	0.4	0.09 (1.07)	-0.05 (0.96)	0.2
Age	49 (16)	54 (17)	<0.001	54 (17)	49 (16)	<0.001
Household Size	2.53 (1.26)	2.60 (1.31)	0.5	2.71 (1.57)	2.73 (1.31)	0.3
Net Wage	195,225 (225,483)	153,826 (158,248)	<0.001	163,702 (107,919)	205,504 (177,192)	0.003
Unknown Wage	189	80		86	149	
Education Level:			<0.001			0.002
Less than secondary school	160 (31%)	110 (47%)		102 (38%)	148 (29%)	
Secondary school	228 (45%)	76 (32%)		116 (44%)	215 (42%)	
University Degree	120 (24%)	48 (21%)		47 (18%)	144 (28%)	
Gender:			0.6			0.062
Female	263 (52%)	116 (50%)		152 (57%)	255 (50%)	
Male	245 (48%)	118 (50%)		113 (43%)	252 (50%)	

4.3 Appendix - Chapter 4

1 Robustness: latent variable for economic expectations

An argument can be made that taking the average of the "optimism"-variables is not essentially the best way to analyze the phenomena. We chose to do this due to the fact that we could still analyze, for example, $UNEMP$ separately to $opt - macro$. Alternatively, we ran a Principal Component Analysis, and used the first factor to capture the latent "economic expectations". Then, using this factor, we ran all the main regressions of the paper. As per Table 4.20, we find that while the lower income quintiles are now significantly different from each other (an effect which only disappears once we add $HH.Prev.Year$), the effect is increasing by income-quintiles. Similarly, the effect of recession seems to be stronger with the upper quintiles, again suggesting that differences in economic expectations among income ranks get closer to each other, while for education, this interaction term is not significant. Overall, our results do not differ significantly from the results above.

2 Correlation matrix

This appendix contains the correlation table with all variables, including those (CAR $HOME$ and $HOME - exp$) for which we have a considerably lower number of observations. The associations observed in Table 4.3 still hold.

Table 4.20: Main Regressions in the paper using PCA as a measure for the latent variable ”optimism”

	Dependent variable:			
	First factor of PCA			
	(1)	(2)	(3)	(4)
Quintile 2	0.038***		0.030**	−0.003
	(0.011)		(0.012)	(0.010)
Quintile 3	0.119***		0.131***	0.059***
	(0.011)		(0.012)	(0.010)
Quintile 4	0.241***		0.254***	0.135***
	(0.011)		(0.012)	(0.010)
Quintile 5	0.423***		0.445***	0.237***
	(0.011)		(0.012)	(0.011)
Quintile (linear)		0.105***		
		(0.003)		
Rec. × Quintile 2			0.040	0.049**
			(0.028)	(0.024)
Rec. × Quintile 3			−0.079***	−0.031
			(0.028)	(0.024)
Rec. × Quintile 4			−0.083***	−0.058**
			(0.028)	(0.024)
Rec. × Quintile 5			−0.133***	−0.081***
			(0.029)	(0.025)
Has University Diploma	0.229***	0.247***	0.220***	0.141***
	(0.011)	(0.011)	(0.012)	(0.011)
Has high-school degree	0.123***	0.127***	0.119***	0.074***
	(0.011)	(0.011)	(0.012)	(0.010)
Recession	−0.238***	−0.238***	−0.192***	−0.105*
	(0.061)	(0.061)	(0.063)	(0.055)
Rec. × Univ. Diploma			0.045	0.033
			(0.029)	(0.026)
Rec. × High-school degree			0.016	−0.003
			(0.028)	(0.025)
HH.Prev.Year				0.499***
				(0.003)
Constant	−0.411***	−0.550***	−0.417***	0.124***
	(0.046)	(0.047)	(0.046)	(0.040)
Observations	75,713	75,713	75,713	75,619
R^2	0.179	0.177	0.179	0.385
Adjusted R^2	0.177	0.176	0.178	0.384
Residual Std. Error	0.907 (df = 75588)	0.908 (df = 75591)	0.907 (df = 75582)	0.785 (df = 75487)
F Statistic	132.582***	134.519***	126.904***	360.676***
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 4.21: Correlation table with all variables

	Inc	Age	ECON-macro	INF	UNEMP	ECON-hh	HH.Prev.Year	SAV	DUR	OPT-macro	OPT-hh	Diploma	CAR	HOME	DUR-worth
Income	1														
Age	-0.023	1													
ECON-macro	0.013	-0.059	1												
INF	0.050	0.0001	0.352	1											
UNEMP	0.005	0.004	0.475	0.394	1										
ECON-hh	0.024	-0.116	0.644	0.329	0.410	1									
HH.Prev.Year	0.070	-0.070	0.464	0.258	0.348	0.574	1								
SAV	0.122	-0.110	0.380	0.238	0.302	0.431	0.434	1							
DUR	0.058	-0.102	0.458	0.289	0.360	0.512	0.479	0.464	1						
OPT-macro	0.026	-0.027	0.826	0.690	0.807	0.613	0.471	0.403	0.484	1					
OPT-hh	0.087	-0.136	0.611	0.354	0.443	0.800	0.615	0.805	0.805	0.619	1				
Diploma	0.194	0.017	0.054	0.066	0.066	0.050	0.085	0.157	0.079	0.078	0.122	1			
CAR	0.113	-0.197	0.169	0.117	0.141	0.213	0.215	0.316	0.243	0.186	0.323	0.097	1		
HOME	0.082	-0.192	0.112	0.071	0.083	0.138	0.141	0.186	0.154	0.116	0.200	0.079	0.280	1	
DUR-worth	0.047	-0.066	0.169	0.096	0.130	0.171	0.176	0.223	0.215	0.174	0.253	0.072	0.127	0.079	1
HOME-exp	0.100	−0.167	0.200	0.094	0.141	0.224	0.219	0.307	0.243	0.193	0.324	0.079	0.270	0.307	0.168

3 Validity of economic expectations on economic data

While the focus of this study is differences in expectations based on SES, and their influence on purchases, we also want to highlight the relationship between expectations and actual macroeconomic data. There are antecedents of such exercises in the literature. For example, Coibion et al., 2022 show that households tend to marginally overestimate actual inflation: the average estimation was of 2.5 percent compared to the 2.3 percent actual value of the CPI index in 2018. However, they also report that when asked about the FED's inflation target, less than 20 percent answered correctly that the target is 2 percent, and more than 40 percent answered that it is over 10 percent.

To see whether expectations indeed reflect actual macroeconomic processes, we compare them to actual economic data. In Figures 4.8 and 4.9, we plot our constructed macroeconomic optimism index (OPT-macro) and the monthly unemployment data.

We observe three patterns:

- Expectations reflect the seasonality of unemployment.
- Political events play an important role in expectations. For example, macroeconomic expectations became more positive after elections (2002 and 2006), they also increased with Hungary joining the European Union (May of 2004), but they decreased with the political crisis of 2006.
- For unemployment, overall, the downward-sloping trend of optimism coincides with the upward-sloping unemployment rate observed in the period under consideration. This connection is even more visible when compared to unemployment expectations in Figure 4.9.

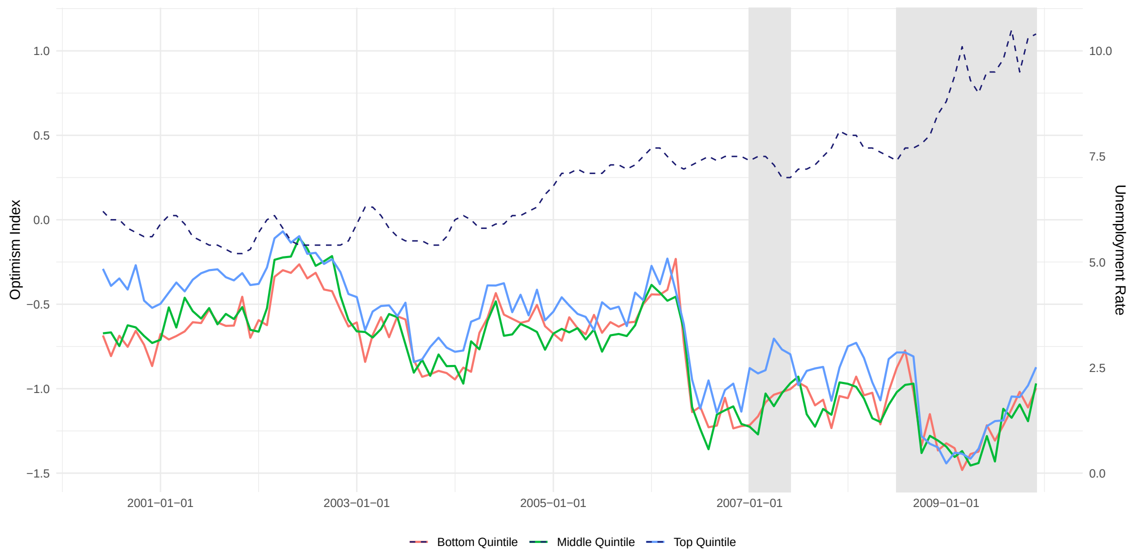

Figure 4.8: Macroeconomic Optimism Index by Income Quintiles (solid lines, left axis) and the monthly Unemployment Rate (scattered line, right axis). Shaded areas indicate recession.
Source: FRED

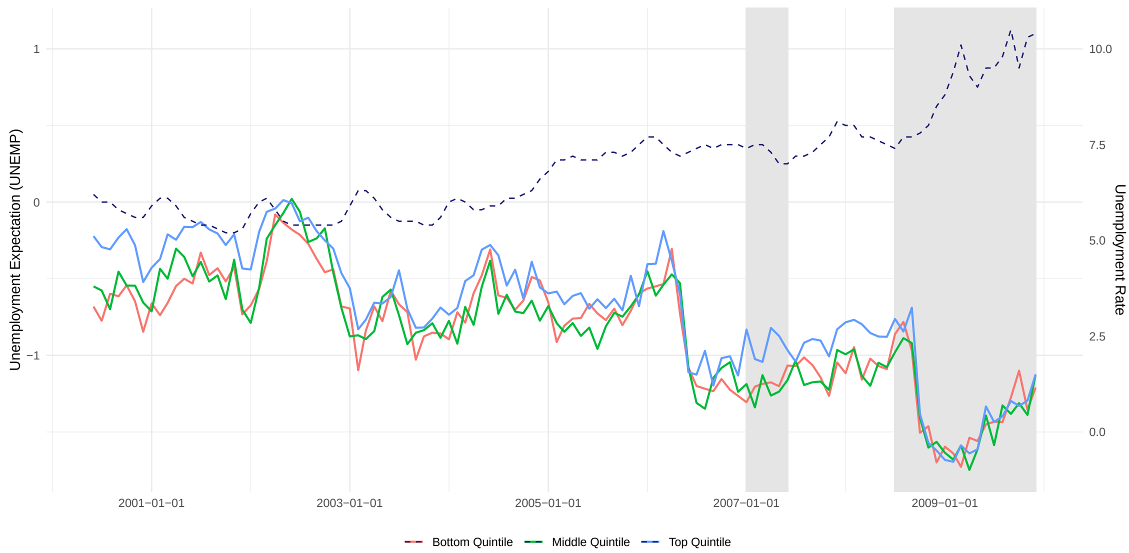

Figure 4.9: Unemployment expectations by Income Quintiles (solid lines, left axis) and the monthly Unemployment Rate (scattered line, right axis). Shaded areas indicate recession.
Source: FRED

In Figure 4.10, we present a plot comparing inflation expectations with the actual year-on-year inflation rate. During the period under investigation, the average year-on-year inflation in Hungary was approximately 6 percent, with higher inflation rates observed prior to 2002 and during the recession of 2007.

The relationship between inflation expectations and actual inflation is not as straightforward. Before 2002, there is no clear pattern indicating a decrease in inflation expectations despite a decline in actual inflation. However, increases in inflation are preceded by more pessimistic outlooks in 2003 and 2006. Additionally, the decline in inflation following the peak in 2007 is accompanied by only a modest increase in optimism.

Overall, it can be concluded that inflation expectations generally align with actual data, with some exceptions. In cases where expectations deviate from actual inflation, they are more likely to be influenced by political events rather than purely economic factors.

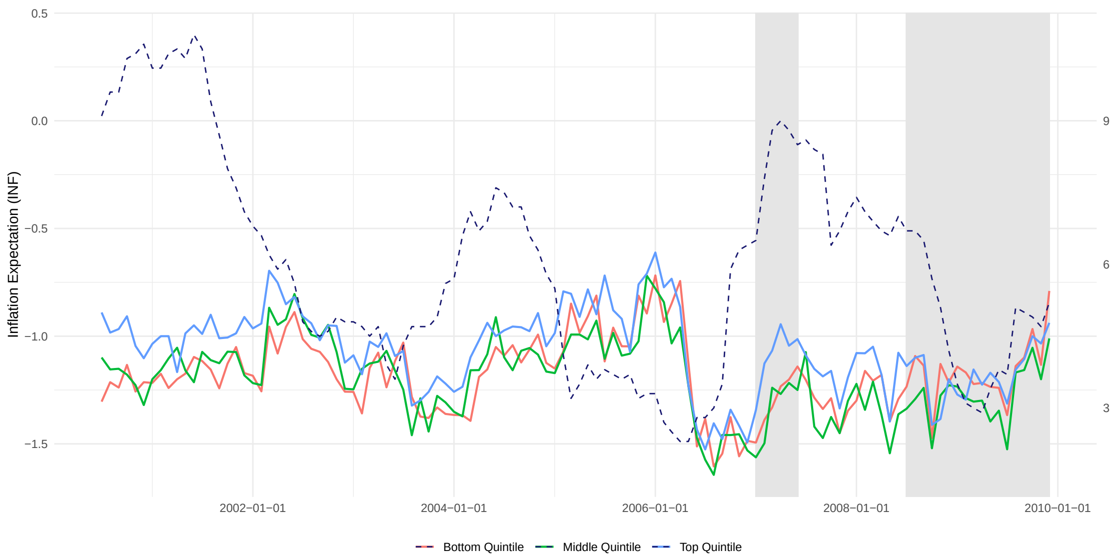

Figure 4.10: Inflation expectations by Income Quintiles (solid lines, left axis) and the year-on-year Inflation Rate (scattered line, right axis). The grey area marks quarters when the economy was in recession.
Source: FRED

4 Robustness of SES

To address the strong correlation between education and income level, we conduct additional robustness checks on the non-linearity findings presented in Table 4.4. Specifically, we perform separate regressions by including only the income and education dummy variables. The results of these regressions are presented in Tables 4.22 and 4.23.

The results from these robustness checks confirm that while the correlation between education and income is an important issue, it generally has a limited effect on the estimated coefficients. This suggests that the non-linearity observed in Table 4.4 remains robust and is not solely driven by the correlation between education and income.

Table 4.22: Robustness check: multicollinearity between macroeconomic optimism components

	Dependent variable:								
	ECON-macro			INF			UNEMP		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Quintile 2	0.036***		0.038***	−0.002		0.0001	0.001		0.003
	(0.012)		(0.012)	(0.009)		(0.009)	(0.010)		(0.010)
Quintile 3	0.093***		0.090***	0.041***		0.037***	0.049***		0.045***
	(0.012)		(0.012)	(0.009)		(0.009)	(0.010)		(0.010)
Quintile 4	0.169***		0.157***	0.100***		0.087***	0.120***		0.105***
	(0.012)		(0.012)	(0.009)		(0.009)	(0.010)		(0.010)
Quintile 5	0.254***		0.220***	0.184***		0.148***	0.232***		0.193***
	(0.012)		(0.012)	(0.009)		(0.009)	(0.010)		(0.011)
Has high-school degree		0.124***	0.101***		0.109***	0.092***		0.135***	0.113***
		(0.012)	(0.012)		(0.009)	(0.009)		(0.010)	(0.010)
Has University diploma		0.188***	0.122***		0.183***	0.133***		0.208***	0.142***
		(0.012)	(0.012)		(0.008)	(0.009)		(0.010)	(0.011)
Recession	−0.015	−0.007	−0.007	0.318***	0.326***	0.326***	−0.209***	−0.200***	−0.200***
	(0.066)	(0.066)	(0.066)	(0.048)	(0.048)	(0.048)	(0.058)	(0.058)	(0.058)
Constant	−0.911***	−0.812***	−0.929***	−1.452***	−1.400***	−1.469***	−1.292***	−1.224***	−1.312***
	(0.051)	(0.050)	(0.050)	(0.037)	(0.037)	(0.037)	(0.044)	(0.044)	(0.044)
Observations	75,713	75,713	75,713	75,713	75,713	75,713	75,713	75,713	75,713
R^2	0.167	0.164	0.169	0.063	0.061	0.066	0.158	0.155	0.161
Adjusted R^2	0.166	0.163	0.167	0.061	0.060	0.065	0.156	0.154	0.159
F Statistic	128.495***	128.046***	128.004***	42.918***	42.369***	44.837***	120.221***)	120.016***	120.876***

Standard errors in parentheses.
*/**/*** denotes significance at 1 / 5 / 10 % level.
All regressions include year-month dummies,age, gender, family status.

Table 4.23: Robustness check: multicollinearity between household-level optimism components

	Dependent variable:								
		ECON-hh			SAV			DUR	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Quintile 2	0.029***		0.030***	0.019		0.022*	0.025**		0.028**
	(0.011)		(0.011)	(0.013)		(0.013)	(0.011)		(0.011)
Quintile 3	0.071***		0.067***	0.155***		0.145***	0.096***		0.091***
	(0.011)		(0.011)	(0.013)		(0.013)	(0.011)		(0.011)
Quintile 4	0.152***		0.139***	0.352***		0.318***	0.202***		0.185***
	(0.011)		(0.011)	(0.013)		(0.013)	(0.011)		(0.011)
Quintile 5	0.267***		0.232***	0.683***		0.587***	0.375***		0.328***
	(0.011)		(0.011)	(0.013)		(0.013)	(0.011)		(0.011)
Has high-school degree		0.099***	0.074***		0.230***	0.165***		0.172***	0.137***
		(0.011)	(0.011)		(0.013)	(0.013)		(0.011)	(0.011)
Has University diploma		0.207***	0.133***		0.568***	0.373***		0.280***	0.173***
		(0.011)	(0.011)		(0.013)	(0.013)		(0.011)	(0.011)
Recession	−0.034	−0.026	−0.026	−0.171**	−0.150**	−0.152**	−0.129**	−0.117*	−0.118*
	(0.063)	(0.063)	(0.063)	(0.071)	(0.072)	(0.071)	(0.061)	(0.061)	(0.061)
Constant	−0.683***	−0.584***	−0.697***	−1.541***	−1.302***	−1.571***	−1.017***	−0.888***	−1.042***
	(0.048)	(0.047)	(0.048)	(0.054)	(0.055)	(0.054)	(0.047)	(0.046)	(0.047)
Observations	75,713	75,713	75,713	75,713	75,713	75,713	75,713	75,713	75,713
R^2	0.161	0.157	0.162	0.114	0.092	0.124	0.120	0.111	0.124
Adjusted R^2	0.159	0.155	0.161	0.112	0.091	0.123	0.118	0.109	0.122
Residual Std. Error	0.938	0.940	0.937	1.068	1.080	1.061	0.912	0.917	0.910
F Statistic	123.315***	122.180***	123.040***	81.961***	66.824***	89.645***	88.136***	81.854***	90.012***

Standard errors in parentheses.
*/**/*** denotes significance at 1 / 5 / 10 % level.
All regressions include year-month dummies, age, gender, family status.

5 Standardized coefficients

To facilitate the comparison of the coefficients, we include the standardized coefficients of the SES variables. Note that for inflation (INF) and unemployment (UNEMP), the influence of being in a recession is greater than that of the observed SES variables, but in the other instances SES variables play a more important role.

Table 4.24: Standardized coefficients of Table 4.4

	ECON-macro	INF	UNEMP	ECON-hh	SAV	DUR
Quintile 2	0.014	0.00005	0.001	0.012	0.008	0.011
Quintile 3	0.033	0.020	0.019	0.026	0.051	0.038
Quintile 4	0.058	0.047	0.045	0.055	0.113	0.077
Quintile 5	0.082	0.080	0.082	0.091	0.209	0.136
Has University Diploma	0.035	0.056	0.047	0.041	0.104	0.056
Has high-school degree	0.029	0.038	0.037	0.022	0.045	0.044
Recession	-0.002	0.165	-0.080	-0.010	-0.051	−0.046

6 Ordinal Logit Models

As an additional step to validate our results, we use an Ordinal Logit Model framework implemented in R in the ”MASS” package (Ripley et al., 2013). Throughout our regression analysis, we assumed a linear relationship along the responses (-2: will be much worse; -1: will be worse, 0: will remain the same, etc.). We can also analyze the issue by looking at the effect of factors on proportional odds of having a more favorable outlook. We include results for the OPT-macro components (that is: ECON-macro, INF, and UNEMP). P-values are obtained by comparing the t-values against the normal distribution. Overall, we see similar results to our OLS regressions, although it is worth noting that there is significant variability in the effects. For example, in the case of inflation, recession and education seem to have a higher effect than in the estimates of UNEMP and ECON-macro.

Table 4.25: Ordinal Logit Models estimate for ECON-macro

	Value	Std. Error	t-value	p-value
Quintile 2	0.025	0.022	1.132	0.258
Quintile 3	0.073	0.022	3.314	0.001
Quintile 4	0.128	0.022	5.799	0
Quintile 5	0.121	0.023	5.284	0
Has University Diploma	0.127	0.023	5.487	0
Has High-School Degree	0.127	0.022	5.643	0
Recession	0.171	0.122	1.397	0.162
HH.Prev.Year	0.867	0.008	105.967	0
Intercepts:				
-2|-1	-1.599	0.094	-17.080	0
-1|0	-0.057	0.093	-0.614	0.540
0|1	1.573	0.094	16.805	0
1|2	5.334	0.100	53.501	0

Table 4.26: Ordinal Logit Models estimate for INF

	Value	Std. Error	t-value	p-value
Quintile 2	-0.040	0.023	-1.715	0.086
Quintile 3	0.043	0.023	1.829	0.067
Quintile 4	0.138	0.023	5.887	0
Quintile 5	0.223	0.024	9.221	0
Has University Diploma	0.305	0.024	12.787	0
Has High-School Degree	0.224	0.024	9.494	0
Recession	1.042	0.131	7.929	0
HH.Prev.Year	0.534	0.008	65.596	0
Intercepts:				
-2|-1	-0.457	0.101	-4.534	0
-1|0	2.269	0.101	22.435	0
0|1	4.710	0.105	44.882	0
1|2	5.063	0.107	47.489	0

Table 4.27: Ordinal Logit Models estimate for UNEMP

	Value	Std. Error	t-value	p-value
Quintile 2	-0.023	0.022	-1.013	0.311
Quintile 3	0.031	0.022	1.404	0.160
Quintile 4	0.114	0.022	5.109	0
Quintile 5	0.211	0.023	9.189	0
Has University Diploma	0.233	0.023	10.287	0
Has High-School Degree	0.206	0.023	9.139	0
Recession	-0.350	0.125	-2.797	0.005
HH.Prev.Year	0.611	0.008	77.443	0
Intercepts:				
-2|-1	-0.773	0.095	-8.144	0
-1|0	1.199	0.095	12.629	0
0|1	3.134	0.096	32.760	0
1|2	6.146	0.107	57.492	0

7 Income level as a linear regressor

In this Appendix, we report regressions that follow Das et al., 2020 by imposing a linear structure. That is, instead of using quintile dummies, we introduce a variable called Quintile that takes the value of the corresponding quintile (1 for the bottom quintile, 2 for the second quintile and so on). In specification (1) (which is the most akin to Das et al., 2020), we find that the linear income rank has a significant and positive coefficient. However, once we include relevant factors in specifications (2) and (3), the coefficient becomes insignificant, as in our preferred regression (see Table 4.5). As for household-level optimism, the linear income variable (Quintile) remains significant in all specifications.

Table 4.28: Recreation of Table 4.5 with linear income rank specification

	Dependent variable:					
	OPT-macro			OPT-hh		
	(1)	(2)	(3)	(4)	(5)	(6)
Quintile	0.051***	0.002	−0.001	0.103***	0.070***	0.047***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Has University diploma	0.145***	0.030	0.037*	0.239***	0.145***	0.145***
	(0.024)	(0.020)	(0.020)	(0.028)	(0.023)	(0.021)
Rec. × Quintile	−0.008*	0.007**	0.007**	−0.031***	−0.026***	−0.020***
	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)
Rec. × Univ. Diploma	−0.005	−0.004	−0.006	−0.003	0.0001	−0.008
	(0.006)	(0.005)	(0.005)	(0.007)	(0.006)	(0.005)
OPT-hh		0.482***	0.420***			
		(0.003)	(0.003)			
HH.Prev.Year			0.093***			0.321***
			(0.003)			(0.002)
OPT-macro					0.646***	0.467***
					(0.003)	(0.003)
Recession	0.059	0.063*	0.078**	−0.009	−0.047	0.014
	(0.045)	(0.038)	(0.037)	(0.053)	(0.044)	(0.039)
Constant	−1.289***	−0.697***	−0.661***	−1.229***	−0.396***	−0.247***
	(0.033)	(0.028)	(0.028)	(0.039)	(0.032)	(0.029)
Observations	75,713	75,713	75,619	75,713	75,713	75,619
R^2	0.191	0.443	0.452	0.191	0.443	0.546
Adjusted R^2	0.190	0.442	0.452	0.190	0.442	0.545
Residual Std. Error	0.651	0.541	0.536	0.756 (	0.628	0.567
F Statistic	149.653***	505.892***	522.460***	146.736*** (	501.706***	757.878***

Standard errors in parentheses.
*/**/*** denotes significance at 1 / 5 / 10 % level.
All regressions include year-month dummies, age, gender, family status.

8 Income level using income deciles

For further analysis, we included a specification where instead of income quintiles, we use income decile dummies (again, by defining the income deciles by clustering for age and the month of the survey). Again, we find that the first three - and in the case of $UNEMP$, the first four - income deciles are not significantly different from each other. At the higher end, however, we can see a jump starting from Deciles 9 in many cases, and not just between 9 and 10, meaning that this increase in effect is not strictly concentrated on the highest income level.

Table 4.29: Regression results for economic expectations based on separate decile dummies

	Dependent variable:							
	ECON-macro	INF	UNEMP	OPT-macro	ECON-hh	SAV	DUR	OPT-hh
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Decile3	0.019	−0.006	−0.009	0.001	0.010	−0.010	0.013	0.004
	(0.014)	(0.010)	(0.012)	(0.009)	(0.013)	(0.015)	(0.013)	(0.011)
Decile4	0.060***	0.008	0.014	0.027***	0.050***	0.050***	0.040***	0.047***
	(0.014)	(0.010)	(0.012)	(0.009)	(0.013)	(0.015)	(0.013)	(0.011)
Decile5	0.076***	0.023**	0.030**	0.043***	0.056***	0.107***	0.072***	0.078***
	(0.014)	(0.010)	(0.012)	(0.009)	(0.013)	(0.015)	(0.013)	(0.011)
Decile6	0.106***	0.051***	0.058***	0.072***	0.078***	0.182***	0.111***	0.124***
	(0.014)	(0.010)	(0.012)	(0.009)	(0.013)	(0.015)	(0.013)	(0.011)
Decile7	0.156***	0.079***	0.095***	0.110***	0.126***	0.282***	0.164***	0.191***
	(0.014)	(0.010)	(0.012)	(0.009)	(0.013)	(0.015)	(0.013)	(0.011)
Decile8	0.162***	0.097***	0.117***	0.125***	0.157***	0.359***	0.208***	0.242***
	(0.014)	(0.010)	(0.012)	(0.009)	(0.013)	(0.015)	(0.013)	(0.011)
Decile9	0.180***	0.120***	0.168***	0.156***	0.177***	0.460***	0.260***	0.299***
	(0.014)	(0.010)	(0.013)	(0.009)	(0.014)	(0.015)	(0.013)	(0.011)
Decile10	0.267***	0.178***	0.221***	0.222***	0.292***	0.724***	0.401***	0.473***
	(0.015)	(0.011)	(0.013)	(0.010)	(0.014)	(0.016)	(0.013)	(0.011)
Has University Diploma	0.113***	0.128***	0.137***	0.126***	0.122***	0.347***	0.160***	0.210***
	(0.012)	(0.009)	(0.011)	(0.008)	(0.012)	(0.013)	(0.011)	(0.009)
Has High-School Degree	0.100***	0.092***	0.112***	0.101***	0.073***	0.162***	0.135***	0.123***
	(0.012)	(0.009)	(0.010)	(0.008)	(0.011)	(0.013)	(0.011)	(0.009)
Constant	−0.931***	−1.470***	−1.313***	−1.238***	−0.697***	−1.573***	−1.043***	−1.104***
	(0.050)	(0.037)	(0.044)	(0.033)	(0.048)	(0.054)	(0.046)	(0.038)
Observations	75,713	75,713	75,713	75,713	75,713	75,713	75,713	75,713
R^2	0.169	0.067	0.161	0.194	0.163	0.128	0.125	0.198
Adjusted R^2	0.168	0.065	0.159	0.192	0.162	0.126	0.123	0.196
Residual Std. Error (df = 75584)	0.988	0.720	0.867	0.649	0.936	1.058	0.910	0.749
F Statistic (df = 128; 75584)	120.077***	42.170***	113.207***	141.949***	115.135***	86.324***	84.227***	145.389***
Note:								*p<0.1; **p<0.05; ***p<0.01

Standard errors in parentheses.
*/**/*** denotes significance at 1 / 5 / 10 % level.
All regressions include time-year dummies, age, gender, family status.
OPT-macro = (ECON-macro + INF + UNEMP)/3
OPT-hh = (ECON-hh + SAV + DUR) / 3

9 Heterogeneity of past experience

Table 4.30: Heterogeneous effects of past experiences on SES as a channel for optimism

	Dependent variable:			
	OPT-macro		Opt-HH	
	(1)	(2)	(3)	(4)
HH. Prev.Year $\times Quintile$	-0.002		0.011^{***}	
	(0.002)		(0.002)	
HH. Prev.Year $\times Quintile2$		0.003		-0.003
		(0.007)		(0.008)
HH. Prev.Year $\times Quintile3$		0.013^{*}		0.014^{*}
		(0.007)		(0.008)
HH. Prev.Year $\times Quintile4$		0.005		0.024^{***}
		(0.007)		(0.008)
HH. Prev.Year $\times Quintile5$		-0.014^{*}		0.034^{***}
		(0.007)		(0.008)
Constant	-0.956^{***}	-0.931^{***}	-0.714^{***}	-0.627^{***}
	(0.031)	(0.031)	(0.033)	(0.032)
Observations	75,619	75,619	75,619	75,619
R^2	0.320	0.320	0.436	0.437
Adjusted R^2	0.318	0.319	0.435	0.436
Residual Std. Error	0.596	0.596	0.628	0.628
F Statistic	288.278^{***} (	275.184^{***}	473.972^{***}	453.302^{***}

Standard errors in parentheses.
*/**/*** denotes significance at 1 / 5 / 10 % level.
All regressions include year-month dummies, age, gender, income quintiles, education level, family status and recession.