Università
Ca'Foscari
Venezia

Corso di Dottorato di ricerca
in Economia

ciclo XXXIII

Tesi di Ricerca
in cotutela con Universiteit van Amsterdam

Cognitive Biases in Expectation Formation

Lab evidence on preferences for redistribution, financial forecasting, and subscription traps

SECS-P/01 Economia politica
SECS-P/03 Scienza delle finanze

Coordinatore del Dottorato
ch. prof. Giacomo Pasini

Supervisore
ch. prof. Michele Bernasconi

Supervisore cotutela
ch. prof. Jan Tuinstra

Dottorando
Frieder Neunhoeffer
Matricola 956400

Cognitive biases in expectation formation

Lab evidence on redistribution preferences, financial forecasting, and subscription traps

Cognitive biases in expectation formation

Lab evidence on redistribution preferences, financial forecasting, and subscription traps

DOCTORAL DISSERTATION IN ECONOMICS

Frieder Neunhoeffer

Joint Ph.D. program

UNIVERSITEIT VAN AMSTERDAM

Università Ca'Foscari Venezia

European joint doctorate program

Expectations and Social Influence Dynamics in Economics

ExSIDE

Thesis committee

Supervisors:	Prof. dr. M. Bernasconi	Ca' Foscari University of Venice
	Prof. dr. J. Tuinstra	University of Amsterdam
Other members:	Prof. dr. C.H. Hommes	University of Amsterdam
	Prof. dr. R. Sausgruber	Vienna University of Economics
	Prof. dr. J.H. Sonnemans	University of Amsterdam
	Dr. J.J. van der Weele	University of Amsterdam
	Prof. dr. P.P. Wakker	Erasmus University Rotterdam
	Dr. M. Weber	University of St. Gallen

University of Amsterdam
PhD thesis in *economie en bedrijfskunde*

Ca' Foscari University of Venice
SECS-P/01 Economia politica
SECS-P/03 Scienza delle finanze
Student number: 956400
Cycle: XXXIII

European joint doctorate program
Program: *Expectations and Social Influence Dynamics in Economics*
European Union Horizon 2020 research and innovation programme
Marie Skłodowska-Curie grant agreement No 721846.
Individual research project:
Testing Models of Expectations Formation in the Lab with Real World Data

This thesis has been written within the framework of the Marie Skłodowska-Curie Actions Innovative Training Network 'Expectations and Social Influence Dynamics in Economics' (ExSIDE), with the purpose of obtaining a joint doctorate degree. The thesis was prepared in the Faculty of Economics and Business at the University of Amsterdam and in the Facoltà di Economia at the Università Ca' Foscari Venezia.

Corralejo, April 1, 2021

Cognitive biases in expectation formation

Lab evidence on redistribution preferences, financial forecasting, and subscription traps

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. K.I.J. Maex
ten overstaan van een door het College voor Promoties ingestelde
commissie, in het openbaar te verdedigen in de Agnietenkapel
op donderdag 20 mei 2021, te 16.00 uur
door

Frieder Neunhoeffer

geboren te Tübingen, Duitsland

Promotiecommissie

Promotores:	Prof. dr. M. Bernasconi	Università Ca' Foscari Venezia
	Prof. dr. J. Tuinstra	Universiteit van Amsterdam
Overige leden:	Prof. dr. C.H. Hommes	Universiteit van Amsterdam
	Prof. dr. R. Sausgruber	Wirtschaftsuniversität Wien
	Prof. dr. J.H. Sonnemans	Universiteit van Amsterdam
	Dr. J.J. van der Weele	Universiteit van Amsterdam
	Prof. dr. P.P. Wakker	Erasmus Universiteit Rotterdam
	Dr. M. Weber	Universität St. Gallen

Faculteit Economie en Bedrijfskunde

Dit proefschrift is tot stand gekomen in het kader van het Marie Skłodowska-Curie Actions Innovative Training Network 'Expectations and Social Influence Dynamics in Economics' (ExSIDE), met als doel het behalen van een gezamenlijk doctoraat. Het proefschrift is voorbereid in de Faculteit Economie en Bedrijfskunde van de Universiteit van Amsterdam en in de Facoltà di Economia van de Università Ca' Foscari Venezia.

Preface

Often children (and even some adults) think that doctors and physicians are one and the same. For a long time, I, too, was under the impression that all doctors are physicians and all physicians hold a doctoral degree. Despite a strong correlation, obviously, there is no causality. Some erroneous biases are widespread in society but can be quickly dispelled by simple information provision. Others are cognitively deeper rooted. Although aware of cognitive biases, humans tend to fall for them repeatedly. Our brain is susceptible to certain pitfalls once arguably essential for survival but certainly disadvantageous in today's world.

The feeling of satiety, for example, sets in only with some delay after eating. In the past, abundant food was a rare blessing to people, which could have helped to build up body fat reservoirs for less fortunate days. Today, however, many people struggle with obesity and would be glad if the feeling of satiety would set in earlier. Although we are almost helpless against some cognitive biases in the heat of the moment, there are safeguards that can protect us from our impulsive actions. For instance, we might keep the key to the pantry in the basement. The extra time bought from having to go there first during an unbridled ravenous craving for potato chips provides us with the chance to reflect and eventually choose a healthier alternative.

Even though behavioral sciences today deal with more complex implications of these biases, the underlying core - the human mind - remains the same. This dissertation examines preferences for income redistribution, speculative behavior in stock markets, and subscription traps for said cognitive biases and presents possible remedies to cope with them.

Vorwort (German)

Kinder (und sogar noch mancher Erwachsener) sind häufig der Auffassung, dass Doktoren und Ärzte ein und dasselbe sind. Auch mir war lange nicht bewusst, dass nicht alle Doktoren zwangsläufig Ärzte sein und nicht alle Ärzte einen Doktortitel führen müssen. Trotz starker Korrelation, besteht hier natürlich kein kausaler Zusammenhang. Manche Vorurteile sind weit verbreitet, lassen sich aber mit Hilfe weniger Information schnell auflösen. Andere Voreingenommenheiten und verzerrte Wahrnehmungen, im Englischen *biases* genannt, sind kognitiv tiefer verwurzelt. Obwohl wir uns ihrer bewusst sind, gehen wir ihnen doch immer wieder auf den Leim. Das menschliche Gehirn ist anfällig für gewisse Fallstricke, die in grauer Vorzeit mitunter überlebenswichtig waren, heutzutage aber eher von Nachteil sind. Um ein fiktives Beispiel zu bemühen: das Sättigungsgefühl nach dem Essen tritt erst mit gewisser Verzögerung ein. Früher war es vermutlich existenzsichernd sich den Bauch vollzuschlagen, wenn genug zu essen da war, um mit einem gefüllten Körperfettspeicher für weniger gesegnte Tage vorzusorgen. In unserer heutigen Konsumgesellschaft haben die meisten jedoch eher mit Über- als Untergewicht zu kämpfen und würden sich glücklich schätzen, wenn das Sättigungsgefühl früher einträte. Auch wenn wir in der Hitze des Gefechts manchen *cognitive biases* fast schon hilflos ausgeliefert zu sein scheinen, gibt es Sicherheitsvorkehrungen, die uns vor unserem impulsiven Handeln schützen können. Zum Beispiel könnte man den Speisekammerschlüssel im Keller aufbewahren um sich während eines Anfalls ungezügelten Heißhungers durch den "verlängerten Weg" zur schnellen Junkfood-Lösung die nötige Zeit zu kaufen sich einer gesünderen Alternative zu besinnen.

Auch wenn sich die Verhaltensforschung im Bereich der Wirtschaftspsychologie mit komplexeren Auswirkungen jener cognitive biases beschäftigt, bleibt die Grundlage – das menschliche Gehirn – dieselbe. Die vorliegende Dissertation untersucht Präferenzen für Einkommensumverteilung, Spekulationsverhalten an Aktienmärkten, sowie Abo-Fallen auf besagte cognitive biases und gibt entsprechende Handlungsempfehlungen für Praktiker.

Acknowledgements

For a long time, I too was of the opinion that all doctors are physicians. And since I never had any serious ambitions to become a physician,[1] as a child I had probably rather dreamed of becoming an astronaut than a doctor one day.

First and foremost, I owe this doctorate to my parents, Uta and Wolfgang Neunhoeffer, who not only supported me financially during my studies, but together with my sisters Eva, Freya, and Johanna always encouraged me to pursue my own interests. Upon completing my master degree, I, therefore, decided to enroll in a Joint Ph.D. program in economics at the two universities of Venice and Amsterdam within the framework of the European Marie Curie Program ExSIDE.

During a first trip to Venice in June 2017, primarily dedicated to the purpose of getting a first overview of the housing market, I was kindly hosted by my future Italian colleague Andrea Modena and introduced to the Venetian summer life. Unfortunately, by the time I actually moved to Venice in early September, summer was already over. I learned it the hard way. Due to the precarious lack of housing on the Venetian island, I was forced to spend in a campsite tent what was probably the rainiest September since modern weather records.

That I nevertheless quickly found my way around Ca' Foscari University was primarily due to my Italian doctoral advisor Michele Bernasconi. With his thoughtful and polite manner he welcomed me in his office at anytime of the day. The experienced economist gave me, a young scientist, every freedom to develop my own research ideas. Yet, at the same time, he pointed the direction à la Adam Smith's "invisible hand", whenever I was losing my bearings.

In addition, I wish to thank my Italian colleagues and partly later flatmates, who not only shared professional advice and support (e.g., as participants in my pilot experiments), but also, when work allowed, provided a taste of the infamous "dolce vita": Enrico Longo (faithful conversation partner about Italian "Calcio"; with all the fan gear of his favorite club Atalanta Bergamo, he definitely filled the Ph.D. room

[1] At the latest I buried them after my civil service year at the University Hospital in Tübingen

with color and tone), Caterina Pavese (who introduced me to the Florentine accent and later to the subtle nuances of Italian pronunciation), Francesca Larosa (probably the most stylish Italian who made a lasting impression on my mother during a visit, and who invited us Ph.D. fellows to her Ligurian home for an unforgettable weekend trip), Elena Bassoli (who always took care of me and the other Ph.D. students like a big sister, not least with homemade Tagliatelle pasta by her "Nonna"), Irene Simonetti, Anna Macchioni, and Arianna Traini (who always spoke Italian to me, even if mine was unbearable in the beginning, and always showed me a good time over an espresso at the campus bar in their charmingly distinctive kinds).

Especially the evenings in our shared apartment at Rio Marin, experiencing first-hand what it means to be among Italians in the kitchen, will always remain with me. I would also like to mention our colleague and "chef" Giulio Cinquanta, who never ran out of giving me a lesson in "cucina italiana". For example, that the side salad and the main course are neither served on the same plate nor eaten at the same time. Or that pizza is consumed in the evening and cappuccino only until 11 am. I was able to enjoy many of these priceless experiences, incomprehensible to foreigners, together with Florian Mayer, who became a steady guest in our apartment. With the promises of the Venetian summer and over a glass of "Valpolicella" we managed to survive the gray winter months in the canal town.

Although we first met each other as fellow doctoral students, Evans Erhenhi and Francesco Furini soon became good friends off campus, too. Among other things, I was able to share my passion for "spoken word" and a few adventures on land and water with them. Furthermore, Marco Betan deserves to be mentioned in here, not only but in particular for his active involvement in Venice's sports club culture. Among other experiences, I owe him the honor of leading Spinea's American football team "Cocai Terra Ferma" onto the field as the probably first Ph.D. quarterback in club history.

My second Ph.D. year took me to the "Venice of the North". Even though the canals are called "Grachten" there, the obvious parallels between the canal towns might have helped me during the transition to Amsterdam. At the Universiteit van Amsterdam, I had the great fortune to be advised by Jan Tuinstra, a collected and mindful scholar. His intuitions saved me from many costly mistakes in designing experiments. It has always been a pleasure to meet in his office discussing promising research ideas and not being judged for contemplating about rather abstract thoughts. A special memory will remain the joint collaboration with our co-author Mikhail Anufriev, marked by sharp wit and humor, to whom I not only owe my Matlab skills but also valuable lessons of precise and sophisticated research con-

duct.

The fact that I felt right at home in the CeNDEF working group was certainly due to an atmosphere full of support and openness, owed to CeNDEF founding father Cars Hommes. He always stood up for us Ph.D. students and, despite his busy schedule, kept an open ear for me. I am equally grateful to Cees Diks, who whenever necessary took the time to answer my questions about the CeNDEF server or to even find a solution together. Also Stefanie Huber's energetic spirit and openness were extremely valuable for an early-stage researcher like me. Moreover, chatting about our shared passion for surfing made up for a healthy break from work.

Whenever I faced IT issues, my Dutch Ph.D. colleagues Johan de Jong and Myrna Hennequin were also there to help. They showed plenty of patience in explaining the experimental software oTree to me and selflessly shared their software codes, which they surely wrote with a lot of effort considering the quality. As an ice cream lover I will not forget the trip to Johan's family ice cream parlor, providing doubtless the best ice cream in North Holland. I would also like to thank my paranymph Myrna for the unforgettable road trip between the conferences in Ottawa and Vancouver and for her constant willingness to discuss my at times daring research hypotheses.

My office life in Amsterdam was further enriched by my colleagues Alex Grimaud with his incomparable Econ jokes and Eva Levelt with her apt use of impressive vocabulary. I am also grateful to the CREED research group for interesting lectures and stimulating discussions, and last but not least for the table football games, which among other benefits provided ergonomic distraction to my regular office work body position. Special mention deserves my paranymph Margarita Leib, who introduced me to CREED in her amiably direct manner. Literally, her office was always open for refreshing and even profound chats together with her office mates, the bright-minded Ivan Soraperra and the quick-witted basketball expert and fellow econ-rapper Nils Köbis, who were never shy to share a valuable fun fact on work-related topics and beyond.

Off campus, I am grateful to Felix Chilunga for being a fun flatmate and loyal partner during our apartment renovation, and to his wonderful girlfriend Jessica Michgelsen for delicious vegetarian dishes and for introducing us to the Dutch culture. I am also grateful to Marvin Barron, Grega Grmovsek, and Joris Amin who accepted me into the "US" basketball team.

I would also like to take this opportunity to thank other people who have contributed in various ways to the success of my dissertation. First and foremost, I am deeply indebted to Jamela Mohamed, who took on an incredible amount to support me when I was on the verge of madness with my PhD.

Daniele Checchi, Antonio Filippin (for sharing software codes), Luca Corazzini, Valeria Maggian, Tiziana Medda, Anita Kopányi-Peuker (for helpful discussions), Paolo Pellizzari, Sebastiano della Lena (for sacrificing valuable teaching time for my classroom experiments), Ailko van der Veen (for assistance with the CREED lab), Fernando Garcia (for good collaboration as local ExSIDE colleague in the Marie Curie program), Giacomo Pasini and Pietro Dindo (as local Ph.D. Coordinators), Ulrike Haake, Herbert Dawid, and Diana Grieswald (for all the work around the successful and lasting ExSIDE network), Barbara Iacampo, Lisa Negrello, Robert Helmink, Wilma de Kruijf and all the other hardworking staff members in the Venice and Amsterdam administrations.

Last but not least I would like to thank my Ph.D. committee for their precious time and extremely valuable comments (Rupert Sausgruber, Matthias Weber, Joep Sonnemans, Peter Wakker, Joël van der Weele, Cars Hommes, Jan Tuinstra, and Michele Bernasconi).

If I have not explicitly mentioned all the dear people who have accompanied me during my Ph.D. over the last years, it does not mean that I have forgotten you. Please feel particularly appreciated. Without you I would not have been able to complete the present dissertation.

Corralejo, April 1, 2021

Danksagung (German)

Auch ich war lange der Auffassung, dass alle Doktoren Ärzte sind und da ich spätestens nach meinem Zivildienst am Uniklinikum Tübingen keine größeren Ambitionen mehr hegte Mediziner zu werden, hätte ich mir schon als Kind wohl eher träumen lassen Astronaut zu werden als eines Tages Dr. Neunhoeffer.

In erster Linie habe ich diese Doktorwürde meinen Eltern, Uta und Wolfgang Neunhoeffer, zu verdanken, die mir nicht nur finanziell den Rücken während meines Studiums freihielten, sondern mich zusammen mit meinen Schwestern Eva, Freya, und Johanna stets darin bestärkten meinen eigenen Interessen nachzugehen. So entschied ich mich nach meinem Studium im Rahmen des europäischen Marie-Curie Programms ExSIDE an den beiden Universitäten Venedig und Amsterdam in Volkswirtschaftslehre zu promovieren.

Während einer ersten Venedig-Erkundungsreise im Juni 2017, die primär dem Zweck diente sich einen Überblick vom Wohnungsmarkt zu verschaffen, wurde ich von meinem zukünftigen italienischen Kollegen Andrea Modena freundlicherweise in seiner Altstadt-Wohnung beherbergt und ins venezianische Sommerleben eingeführt.

Als ich im September tatsächlich nach Venedig zog war der Sommer leider schon vorbei. Das sollte ich schnell am eigenen Leib zu spüren bekommen. Aufgrund des prekären Wohnungsmangels auf der Insel, war ich gezwungen den wahrscheinlich regnerischsten September seit modernen Wetteraufzeichnungen im Ein-Mann-Zelt auf dem Campingplatz zu verbringen.

Dass ich mich trotzdem zügig an der Ca' Foscari Universität zurechtfand, habe ich zum großen Teil meinem italienischen Doktorvater Michele Bernasconi zu verdanken. In seiner bedachten und höflichen Art war ich stets in seinem Büro mit Campus-Blick willkommen. Der erfahrene Ökonom gab mir jungem Wissenschaftler jegliche Freiheit zur Entfaltung meiner eigenen Forschungsinteressen, aber zugleich à la Adam Smith's „unsichtbarer Hand" die Richtung vor, falls ich mal die Orientierung zu verlieren glaubte.

Zusätzlich habe ich die schnelle Eingewöhnung auf dem Campus meinen italienischen Kollegen und teils späteren Mitbewohnern zu verdanken, die mir nicht nur fachlich mit Rat und Tat zur Seite standen (unter anderem als Versuchsteilnehmer bei meinen Pilotexperimenten) sondern wenn es die Arbeit zuließ mir auch hin und wieder das „Dolce Vita" vorlebten: Enrico Longo (treuer Gesprächspartner über die wichtigste Nebensache der Welt: „Calcio"; mit all seinen Fanartikeln vom Herzensverein Atalanta Bergamo füllte er den PhD-Raum definitiv mit Farbe und Ton), Caterina Pavese (die mir den florentinischen Akzent und später die feinen Nuancen der italienische Aussprache näherbrachte), Francesca Larosa (die wahrscheinlich stilechteste Italienerin, die meine Mutter nachhaltig beeindruckte und uns Doktoranden in ihre ligurische Heimat zu einem unvergessenen Wochenende einlud), Elena Bassoli (die sich immer wie eine große Schwester um uns PhD-Kollegen kümmerte, nicht zuletzt mit von ihrer „Nonna" selbstgemachten Tagliatelle), Irene Simonetti, Anna Macchioni, und Arianna Traini (die immer Italienisch mit mir sprachen, auch wenn meines anfangs bestimmt unerträglich war, und mir auf ihre unverwechselbaren Arten beim Espresso an der Campusbar stets ein Lächeln auf die Lippen zauberten).

Besonders die gemeinsamen Abende in unserer WG am Rio Marin bei denen ich hautnah miterleben durfte was es bedeutet unter Italienern in der Küche zu stehen, werden für mich immer in Erinnerung bleiben. Zu erwähnen ist an dieser Stelle auch unser Kollege und „Chefkoch" Giulio Cinquanta der für mich immer ein, zwei Lehrstunden der italienischen Küche bereithielt. Beispielsweise, dass Beilagensalat und Hauptgericht nicht auf ein- und demselben Teller serviert und auf gar keinen Fall durcheinander gegessen werden dürfen. Oder, dass Pizza abends und Cappuccino nur bis 11 Uhr morgens konsumiert wird. Viele dieser für Ausländer unbegreiflichen Erlebnisse durfte ich zusammen mit Florian Mayer erleben, der in unserer Wohngemeinschaft ein uns ausging. Beim ein oder anderen Glas „Valpolicella" haben wir uns „Exil-Tübinger" gegenseitig Mut gemacht und uns mit der Aussicht auf den venezianischen Sommer über die grauen Wintermonate gerettet.

Obwohl ursprünglich als Kollegen kennengelernt, wurden Evans Erhenhi und Francesco Furini auch abseits des Campus schnell zu guten Freunden. Unter anderem konnte ich mit ihnen meine Leidenschaft für Sprechgesang teilen und das ein oder andere Abenteuer zu Wasser und Land bestreiten. An dieser Stelle gebührt auch Marco Betan größter Respekt für sein aktives Engagement rund um Venedigs Sportvereinkultur. Neben anderen Erlebnissen habe ich ihm die Ehre zu verdanken Spinea's American Football Mannschaft „Cocai Terra Ferma" als wahrscheinlich erster Doktoranden-Quarterback aufs Feld geführt zu haben.

Mein zweites PhD- Jahr führte mich ins „Venedig des Nordens“. Auch wenn die Kanäle hier Grachten heißen fiel mir der Wechsel nach Amsterdam denkbar einfach. An der Universiteit van Amsterdam hatte ich das große Glück mit Jan Tuinstra einen unaufgeregten und scharfsinnigen Doktorvater an meiner Seite zu wissen. Seine Intuitionen haben mich vor so manchem kostspieligen Fehler beim Design von Experimenten bewahrt und es war immer ein Genuß mit ihm über Forschungsideen zu diskutieren. In besonderer Erinnerung wird das gemeinsame Projekt mit unserem dritten Co-Autor Mikhail Anufriev bleiben. Der von spitzfindigem Verstand und Humor gezeichneten Zusammenarbeit verdanke ich nicht nur meine Matlab-Kenntnisse sondern gemeinhin wertvolle Lehrstunden präziser und ausgeklügelter Grundlagenforschung.

Dass ich mich in der Arbeitsgruppe CeNDEF direkt zuhause gefühlt habe lag ganz bestimmt an der von Unterstützung und Offenheit geprägten Atmosphäre, die dem CeNDEF-Gründungsvater Cars Hommes zu verdanken ist. Er setzte sich für uns Doktoranden ein und hatte trotz seines vollen Terminkalenders stets ein offenes Ohr für mich. Genauso dankbar bin ich Cees Diks, der sich wann immer nötig die Zeit nahm mir Fragen zum CeNDEF Server ausführlich zu beantworten oder zusammen eine Lösung zu finden. Auch Stefanie Huber war mit ihrer motivierenden, offenen Art eine wertvolle Ratgeberin für einen jungen Wissenschaftler wie mich. Nebenbei war es immer eine willkommene Pause über unsere gemeinsame Passion - das Surfen - zu fachsimpeln.

Bei Problemen mit der IT waren auch meine holländischen Doktorandenkollegen Johan de Jong und Myrna Hennequin zur Stelle. Sie bewiesen reichlich Geduld damit mir die Experimental-Software oTree zu erklären und teilten selbstlos ihre unter sicherlich großem nervlichem Aufwand geschriebenen Software-Codes. Als Eisliebhaber werde ich natürlich auch den Ausflug zu Johan's Familieneisdiele mit dem fraglos besten Eis Nordhollands nicht vergessen. Bei meiner Paranymphin Myrna möchte ich mich zudem für den erlebnisreichen Roadtrip zwischen den Konferenzen in Ottawa und Vancouver bedanken und genauso für ihre stete Bereitschaft auch bisweilen waghalsige Forschungshypothesen mit mir zu diskutieren. Meinen Büroalltag in Amsterdam haben auch meine Kollegen Alex Grimaud mit unvergleichlichen Econ-Witzen und Eva Levelt mit ihrem beeindruckenden Wortschatz bereichert. Dankbar bin ich auch der CREED-Forschungsgruppe für die interessanten Fach-Vorträge und stimulierenden Gespräche, und nicht zuletzt für den Tischkicker der beim gemeinen Schreibtischtäter für ergonomische Abwechslung sorgte. Besonders zu erwähnen ist meine zweite Paranymphin Margarita Leib, die mich mit ihrer liebenswürdig direkten Art mit den CREEDern bekannt machte. Ihr Büro stand im wahrsten Sinne des Wortes immer offen für ein erfrischendes und mitunter auch tiefsinniges Pläuschchen. Auch ihre Bürokollege Nils Köbis war nicht nur als schlagfertiger Basketballexperte ein äußerst unterhaltsamer Gesprächspartner.

Abseits vom Campus bin ich Felix Chilunga als angenehmen Mitbewohner und verlässlichem Partner bei unserer großen Wohnungsrenovierung dankbar sowie seiner Freundin Jessica Michgelsen für leckere vegetarische Gerichte und das Näherbringen der holländischen Kultur. Zudem Marvin Barron, Grega Grmovsek und Joris Amin die mich ins US Basketball-Team aufnahmen.

Auch möchte ich die Gelegenheit nutzen mich bei weiteren Personen zu bedanken die in verschiedenster Form zum Gelingen meiner Doktorarbeit beigetragen haben. An erster Stelle bin ich Jamela Mohamed in tiefem Dank verbunden, die unglaublich viel auf sich nahm um mir den Rücken zu stärken wenn mich die Promotion mal wieder an den Rand des Wahnsinns trieb.

Daniele Checchi, Antonio Filippin (für das bereitwillige Teilen von Software-Codes), Luca Corazzini, Valeria Maggian, Tiziana Medda, Anita Kopányi-Peuker (für hilfreiche Diskussionen), Paolo Pellizzari, Sebastiano della Lena (für das Opfern von kostbarer Unterrichtszeit für meine Klassenzimmer-Experimente), Fernando Garcia (für die gute Zusammenarbeit als lokaler ExSIDE-Kollege im Marie-Curie-Programm), Giacomo Pasini und Pietro Dindo (als lokale Ph.D. Koordinatoren), Ulrike Haake, Herbert Dawid, und Diana Grieswald (für die viele Arbeit rund um das erfolgreiche ExSIDE Netzwerk), Barbara Iacampo, Lisa Negrello, Robert Helmink, Wilma de Kruijf und all die anderen fleißigen Mitarbeiter in der Uni-Verwaltung in Venedig und Amsterdam.

Nicht zuletzt möchte ich mich bei meinem Ph.D. Komitee für ihre kostbare Zeit und äußerst wertvollen Ratschläge bedanken (Rupert Sausgruber, Matthias Weber, Joep Sonnemans, Peter Wakker, Joël van der Weele, Cars Hommes, Jan Tuinstra und Michele Bernasconi).

Falls ich treue Weggefährten, die mich in den letzten Jahren während meiner Promotion begleitet haben nicht explizit erwähnt habe, heißt das mitnichten, dass ich Euch vergessen habe. Fühlt Euch im Besonderen gedrückt. Ohne Euch hätte ich diese Dissertation nicht vollenden können.

Corralejo, den 1. April 2021

Contents

Chapter 1

Introduction

Poor people prefer less income equality. Less decision time improves financial forecasting. Short subscriptions are valued higher than equally expensive extensions. Is less more? What seems illogical at first can be encountered in real-life situations. Human behavior is pervaded by irrational and biased decision-making. The famous Müller-Lyer illusion (1889, Figure 1.1) refers to the above question at the most basic level: the red line seems *less* long than than the green one, although the opposite is true. Analogously to this deception, stemming from outward vs. inward-pointing arrowheads, behavioral paradoxes usually arouse from changing reference points. Whether the effect is called framing, anchoring, or decoy, to name only a few, cognitive biases in decision-making share a common root, that is, context dependency.

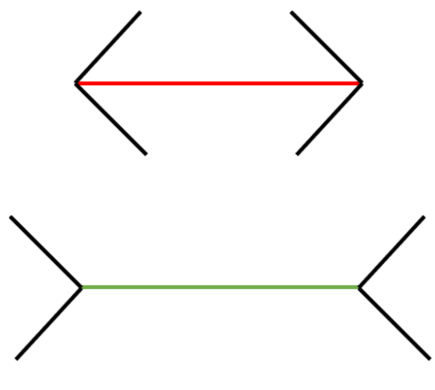


Figure 1.1: Müller-Lyer illusion (1889).

The present dissertation explores distinct cognitive biases in expectation formation through three experimental studies (Chapters 2-4). Whereas the term *expectation formation* usually applies to financial markets (Chapter 3), here it refers more broadly to people's beliefs about the future and their respective decisions. Cognitive biases in expectation formation challenge the rational expectation hypothesis (Lucas Jr, 1972; Muth, 1961), which still prevails as the dominant paradigm in mainstream economics. While empirical data suffer from confounding influences that present a risk in identifying cause-consequence relationships of biased behavior, experiments allow investigating a treatment effect in isolation. The hybrid case - empirical data collected in the form of a natural experiment - sets the gold standard in experimental economics. Yet, occasions for "running" natural experiments are rare, often costly, and difficult to implement. Thus, the controlled laboratory environment, originating from the fraternal field of psychology, experiences growing popularity as data generating process among economists (List et al., 2011). The fusion of both disciplines - *behavioral economics* - is inseparably tight to the field of experimental economics. Their compound success story speaks for itself and exemplifies the estrangement from classical theories in economics toward a more contemporary interpretation of economic problems.[2] Reflecting its root in psychology, the introduction of micro-founded models to macro theories have made the field also increasingly relevant for behavioral macroeconomists.

Lab experiments serve a variety of purposes. They can complement theoretical work by falsifying models, create novel hypotheses as a byproduct or blank tests, or, if properly designed, function as a testbed for policy intervention. Particularly its latter application faces frequent critique questioning the external validity outside a lab's ideal-world conditions. The art of designing experiments lies essentially in the ability to transform complex real-world problems into simulated lab economies to make informed statements on questions otherwise impossible to address.

Chapter 2 meets this challenge by bringing existing inequality levels in South Africa (high) and Switzerland (low) to the lab.[3] We study herein how people's preferences for redistribution change with the level of income inequality, income mobility, uncertainty of initial income positions, and source of income (random or based on real-effort), and find that uncertainty and overconfidence undermine demand for redistribution. The effect magnifies with larger income disparity (South Africa). It

[2]In recent years, the Nobel price in Economics has been frequently awarded to scholars in the field of behavioral and experimental economics, e.g, Banerjee, Duflo and Kremer in 2019, Thaler in 2016, Roth in 2012, Kahneman and Smith in 2002.

[3]Chapter 2 is based on joint work with Michele Bernasconi. I want to thank Valeria Maggian, Luca Corazzini, Tiziana Medda, and Andrea Albarea for stimulating discussions and assistance during the experiment, and Rupert Sausgruber and Matthias Weber for valuable comments. Special thanks go to Antonio Filippin and Daniele Checchi for sharing experimental programs. The work benefited from useful comments of participants at the 2018 SIEP Conf. in Padova, the 2018 Conf. on Decision Sciences in Konstanz, and the 2019 SPUDM Conf. in Amsterdam.

further induces a *reverse* prospect-of-upward-mobility effect: since wealth ambitions of rich aspirants are better preserved under low than under high mobility, demand for redistribution grows in mobility. These results combined propose an *inequality trap*: today's inequality favors income overestimation, winding up less demand for redistribution with less mobility, which propels advanced inequality tomorrow.[4]

Chapter 3 investigates the effect of time pressure in so-called Learning-to-Forecast experiments, which have been found to replicate price volatility of demand-driven asset markets quite accurately.[5] Yet, the scope of prior studies rarely exceeded 50 decision periods or limited decision time substantially, and thereby neglected two central features of financial markets: long runtime and time pressure. This study investigates whether common "bubble and crash" dynamics persist in the long run (150 periods) and how decision time (6 vs. 25 seconds per decision) influences market volatility. For the treatment with low time pressure, we observe a tendency of prices converging to the asset's fundamental value in the long run. Parallel to this change in dynamics, we identify a switch from trend-extrapolating strategies to forecasting strategies that are more adaptive. In contrast, increasing time pressure limits trend-chasing behavior and coordination right from the beginning. Consequently, we find less price volatility and faster convergence to the fundamental value.

Chapter 4 explores a novel menu effect in the context of subscriptions that violates the transitivity principle of rational choice theory.[6] Providers typically capitalize on arranging offers such that a longer but costlier option is chosen over the cheaper but shorter alternative. We find that sizing the shorter subscription down to single-use raises its attraction. This suggests that the presence of a single-use option prompts rational evaluation based on a realistic estimate to use the subscription again. Instead, when both alternatives represent time spans, an irrational mind

[4]See science slam on the topic by Neunhoeffer (2018): 'Warum wählen arme Menschen Millionäre?', English dubbing: 'Why do poor people vote for millionaires?'.

[5]Chapter 3 is based on joint work with Mikhail Anufriev and Jan Tuinstra. I thank Anita Kopányi-Peuker and Cars Hommes for stimulating discussions, and Rupert Sausgruber and Matthias Weber for valuable comments. Special thanks go to Myrna Hennequin and Johan de Jong for sharing experimental programs and to Dávid Kopányi for providing codes to estimate individual rules. The work benefited from useful comments by participants of the Barcelona GSE Summer Forum 2019, the 25th CEF Conference in Ottawa, the 2019 ESA World Meeting in Vancouver, the 2020 SEET Workshop in Naples, and a brown bag seminar at CREED (UvA).

[6]Chapter 4 benefited from useful comments by participants of the SABE Conference 2020 and the 3rd Reading Experimental and Behavioural Economics Workshop. The author is particularly grateful to Michele Bernasconi, Jan Tuinstra, Peter Wakker, Rupert Sausgruber, and Matthias Weber for valuable discussions and comments. The dissertation received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 721846, "Expectations and Social Influence Dynamics in Economics" (ExSIDE) and the ORA project "BEAM" (NWO 464-15-143) which is partly financed by the Netherlands Organization for Scientific Research.

may assign them to the same category - referred to as *pigeonholing* - with the consequence that other comparative criteria come to the fore. Two-dimensional models, present in most behavioral theories, fail to explain this type of preference reversal. Inspired by the intuition of transaction utility (Thaler, 1985) and the availability heuristic (Tversky and Kahneman, 1973), we propose a generalization of salience theory (Bordalo et al., 2012) to capture the effect of pigeonholing.

Chapters 2-4 can be read independently, each providing a separate introduction and conclusion.

Chapter 2

The inequality trap - how high stakes fuel overestimation and equity aversion

2.1 Introduction

Democracy enables the people to choose public policies via the ballot. It is therefore remarkable that, although income inequality has grown worldwide since the 1980s, demand for redistribution has declined continuously throughout the same period (Atkinson, 2015; Stiglitz, 2016).[7]

Motivated by the public relevance of socio-economic issues resulting from spreading inequality, there is resurrecting interest in research for the familiar paradox in western democracies that "the poor do not expropriate the rich" (Roemer, 1998). Why does a substantial share of poor individuals show little support for redistributive policies and even vote for regressive tax schemes although this widens the rich-poor income gap and exacerbates the hardships they experience through inequality?

Growing wealth and income gaps corrode societies, yet like themselves, consequences are multilayered.[8]It is therefore that we contribute to the literature by exploiting the controlled lab environment, which grants isolated analysis of economic drivers that favor inequality. By testing fresh leads from empirical data experimen-

[7]In Western Europe, one of the world regions with the lowest structural change, the income share ratio between the top 10% richest people and the bottom 50% increased by 26% over the period 1980-2016. In the US, shifts were more massive, and this ratio even doubled. On the global level, the top 1% richest individuals captured twice as much of the world's income growth over the period as the bottom 50% (Alvaredo et al., 2018). Despite these large inequality shifts, top marginal income tax rates reduced drastically (on average from 70% to 42 % in the major world economies), with redistribution achieved by the tax-benefit systems also falling (Atkinson, 2015; Piketty, 2014).

tally, we follow several studies that accrued more recently to the field.

Durante et al. (2014) adopt a real-world framing, described in more detail below, which highlights different features of the macroeconomy. In our view, their study sets a benchmark to measure preferences for redistribution in the lab. They find that a mix of classical motives concurs to determine support for redistributive policies. These include maximization of expected earnings, demand for self-insurance against future income shocks, and social concern for inequality and efficiency.

The present study focuses on related issues beyond the scope covered by Durante et al. (2014), including the role of pre-tax inequality on people's demand for redistribution and its interaction with social mobility. Pre-tax inequality can impact redistributive demand in many ways. Besides distortionary costs of taxation, further reasons are fairness of income differences and, from here, the legitimacy of the redistribution process, and the effect that pre-tax inequality can exert on people's expectations about their prospects in life. Regarding the latter, much work has documented a human tendency of holding overoptimistic beliefs (Moore and Healy, 2008). This overconfidence differs from the rational prospect of upward mobility (POUM hypothesis) that people with an income below average (the poor) can have to move up the income ladder when actual mobility is sufficiently high (Benabou and Ok, 2001). Though different from rational expectations, overconfidence can interact with true mobility.

By inducing distinct degrees of inequality, our experiment replicates real-world conditions in South Africa, home of the highest national pre-tax inequality, and conditions in Switzerland, representing the worldwide lowest pre-tax inequality. Whereas in Phase 1 of the experiment, subjects know the income distribution but not their individual position, in Phase 2, subjects learn about their income position, which is either randomly assigned or based on an effort task (baseline design modeled after Durante et al., 2014). By varying income determination and adding a dynamical framework (income mobility) to the treatment set we employ a 2x2x2x2 design altogether. This consists of two within-subject variables (uncertainty vs. certainty of income class, low vs. high income mobility) and two between-subject variables (low vs. high pre-tax inequality, random vs. effort-based income assignment).

In line with Durante et al. (2014), we find that uncertainty about initial income

[8]Economic inequality causes problems for well-being, social cohesion, health, access to education, public order, mortality, etc. (Wilkinson and Pickett, 2017). At the time of writing this paper, the explosion of healthcare needs and the fear of an upcoming recession have made the inequitable response to the COVID-19 pandemic evident (Ahmed et al., 2020). Moreover, the large media attention around the social unrest across major US cities in the aftermath of the death of George Floyd has raised awareness for the negative implications of social inequality and sparked a public debate.

positions arouses overconfidence. What is new, overconfidence intensifies with the level of pre-tax inequality. This induces further a reverse POUM effect. Since uncertainty of initial income positions preserves wealth ambitions of rich aspirants better under low than under high mobility, demand for redistribution is lower under the former than the latter condition. Lifting uncertainty of initial incomes is a game-changer. While demand for redistribution increases, the distributional conflict between rich and poor emerges more polarized, particularly so in the high inequality treatment. Combined results suggest that reducing pre-tax inequality and raising awareness of the own economic position can represent two measures to invert what could be called an inequality trap: inequity promotes income overestimation, which, given low mobility, depresses demand for redistribution, with the consequence of increasing inequality.

The remainder of the paper is structured as follows. The next section reviews related literature on preferences for redistribution. In Section 2.3, we describe the data set, experimental procedure, and design. Section 2.4 contains a descriptive and econometric analysis of the results. The last section concludes with a discussion of the findings, policy implications, and directions for future research. Additional material is in the appendix.

2.2 Literature review

An extensive literature has investigated people's redistributive demand. Below we review the main insights relevant to our experiment. We start with the paradigm of rational preferences and how it deals with income uncertainty.

2.2.1 Rational preferences: redistribution and income uncertainty

The workhorse to analyze redistributive policies, the median voter theorem (Downs, 1957; Meltzer and Richard, 1981), assumes that people know with certainty their income position. Under the hypothesis of the classic *homo oeconomicus*, it predicts that the below-mean income majority (poor) support redistribution, whereas the above-mean minority (rich) oppose it. The actual amount of redistribution resulting from voting decisions depends on various conditions. Nevertheless, it is widely agreed that the theorem predicts more redistribution than generally observed (Alesina and Giuliano, 2011). Experiments can control the level of democracy (Fang et al., 2016; Großer and Reuben, 2013; Ryvkin and Semykina, 2017; Tyran and Sausgruber, 2006), finding that redistribution increases with democratic quality. Sausgruber et al. (2019)

state that even the Meltzer-Richard disincentive effect of redistributive taxation reduces when redistribution is determined democratically.

Among the frequently debated ideas for why tax rates deviate from the median voter theorem, we focus on the role of uncertainty of the relative income in the following.[9] Today's voting decision affects an individual's earnings tomorrow given persistent tax schemes. In this case, individual expectations about future income positions come to the fore. The rational expectation hypothesis considers two effects.

Risk aversion

Risk mitigation implies that also the rich welcome redistribution as a form of self-insurance against future income shocks and that its demand grows in uncertainty (Sinn, 1996). Examining the effect of insurance motives in real-world data can be difficult due to several confounding factors, including multiple equilibria when other insurance opportunities are available from the private sector (Benabou, 2000). Yet, various experiments have confirmed the impact of income uncertainty to generate demand for redistribution under different set-ups (e.g., Cowell and Schokkaert, 2001; Durante et al., 2014; Höchtl et al., 2012; Schildberg-Hörisch, 2010).

Prospects of mobility

Preferences for redistribution can also depend on agents' expectations to ascend or descend the income ladder. Benabou and Ok (2001) introduced the prospect-of-upward-mobility (POUM) hypothesis. Assuming a concave transition function, limited risk aversion, and some duration of the implemented tax scheme, the theorem proves that a rational median income earner votes for limited redistribution today as she expects an above-mean income tomorrow. Various scholars have studied the impact of mobility prospects with survey data (e.g., Alesina and La Ferrara, 2005; Bernasconi, 2006; Cojocaru, 2014; Laméris et al., 2020). While studies often find that preferences for redistribution decrease with greater mobility, the relationship between perceived and actual mobility is often weak (e.g., Alesina et al., 2018; Cheng and Wen, 2019; Swan et al., 2017). Studies, therefore, are generally inconclusive for the POUM hypothesis.

Few experiments have tested the POUM in the lab. Checchi and Filippin (2004) find decreasing average tax rates as income mobility and the length of the imple-

[9] Several critics note that the median voter model is too naive for reality as democratic institutions work less than perfect for various reasons (Harms and Zink, 2003), including the impact that money can have to alter the democratic game through multiple channels (Bartels, 2018).

mented tax scheme rise. Agranov and Palfrey (2020) develop a dynamic model of redistributive taxation, which casts the POUM in a Meltzer-Richard set-up. They find in a corresponding experiment that preferred tax rates fall with increasing mobility. Instead, the level of mobility affects inequality not significantly. Both studies, though not fully comparable, bear theory-driven experiments validating the intuition of the POUM hypothesis in view of a dynamic trade-off between actual levels of inequality and mobility. Yet, neither experiment considers a possible impact that personal beliefs and perceptions can have on this trade-off. By incorporating income uncertainty in a dynamic framework, the present study addresses the POUM question in the context of varying inequality.

2.2.2 Social preferences, merit, luck

In addition to self-interest, economists have long been aware that people hold preferences for what they perceive as socially just. Initially, they have considered social preferences mainly from a normative perspective, reflecting an impartial position (Frohlich and Oppenheimer, 1993; Harsanyi, 1955; Rawls, 1971). From a more practical view, personal interest and social concern are likely to determine jointly attitudes towards economic divisions, including income distributions (Bolton and Ockenfels, 2000; Charness and Rabin, 2002; Fehr and Schmidt, 1999).

Many models of social preferences underlie the close connection between the notions of risk aversion and inequality aversion (Atkinson, 1970). At least for the latter it strictly implies, the greater is inequality, the higher the demand for redistribution ceteris paribus. Moreover, according to the same argument and given redistribution is free of costs, the level of pre-tax inequality ought not to affect preferences for post-tax inequality. The caveat ‘ceteris paribus’, however, can be very important from a positive perspective. Much work has affirmed how people’s culture, history, or traits can affect beliefs about justice and fairness, thus redistributive policy decisions (Alesina and Angeletos, 2005; Benabou and Tirole, 2006; Piketty, 1995).[10]

One conjecture that has attracted attention based on evidence from social surveys argues that people are willing to accept more inequality when income is obtained by effort/merit instead of pure luck (Cojocaru, 2014; Fong, 2001; Luttmer and Singhal,

[10] Roemer (1998) has shown that sociological traits (e.g., race, religion, ethnicity) can affect demand for redistributions also when they represent politically relevant secondary divides used by low-tax parties to distract a fraction of the poor from voting for high redistribution (for recent evidence see Corneo and Neher, 2015). While experiments can control for secondary channels, e.g., by using homogeneous samples like university students, results of experiments can still be influenced by cultural and sociological factors when they affect moral values globally.

2011). Experiments have found some support for this hypothesis, but the evidence is not overwhelming (Balafoutas et al., 2013; Gee et al., 2017; Jiménez-Jiménez et al., 2018; Kesternich et al., 2018; Krawczyk, 2010; Ku and Salmon, 2013; Lefgren et al., 2016). Accordingly, Durante et al. (2014) find more demand for redistribution when luck decides over income positions rather than effort. The difference only occurs before individuals learn their income position but disappears when preferences are stated and the own income position is known. They relate the effect to people's overconfidence about their ability in the effort task.

Cappelen et al. (2007, 2013) confirm that people hold different views on fairness, including egalitarian fairness (no form of income difference is justifiable), libertarian fairness (any form is perceived acceptable), and meritocratic fairness (only inequality arising from merit, not luck, is acceptable). Luck alone can also be viewed as a form of fairness under certain conditions, for example, when goods are indivisible (Broome, 1984; Diamond et al., 1967; Sen, 1970).[11] It can be further regarded as necessary to prevent society from drifting towards too rigid structures that bear negative socio-economic externalities (Arrow et al., 2000; Young, 1958).

2.2.3 Behavioural approaches

Explaining limited redistribution by misperceiving wealth differences has caught growing attention in the debate on inequality. Misperception can affect both how people evaluate their current situation and the way they look into the future.

Misperception of social inequality and mobility

Studies focusing on current income, resp. wealth, find large discrepancies between underestimated and existing inequality (Ashok et al., 2015; Cruces et al., 2013; Hauser and Norton, 2017; Karadja et al., 2017; Norton and Ariely, 2011). They observe general preference for even less inequality than the already underestimated levels.

Few studies have analyzed the relationships between people's expectations of future income positions based on perceived versus actual mobility (Davidai and Gilovich, 2015; Kraus and Tan, 2015). The evidence is mixed. In line with the "American Dream", Alesina et al. (2018) finds that Americans' beliefs about mobility are more optimistic than Europeans' and that while the former are also over-optimistic regarding the level of actual mobility, the latter tend to be over-pessimistic. Other

[11] A recent experiment by Charité et al. (2015) finds that even placed as an external observer, people refrain from redistributing from rich to poor to respect others initial endowments and prevent them from experiencing loss-aversion.

studies find contrasting evidence and argue that Americans' beliefs of social mobility are instead pessimistic (Cheng and Wen, 2019; Swan et al., 2017).[12]

Overconfidence

A different form of misperception that can affect expectations is overconfidence. It is a broad term used in behavioral sciences to describe a tendency of individuals assuming they are better at doing something than they actually are (Moore and Healy, 2008). Various explanations can underlie overconfidence, including cognitive errors, self-motivated beliefs, self-esteem, or social context (Köszegi, 2006; Logg et al., 2018; Schwardmann and Van der Weele, 2019). Forms of overconfidence specifically relevant in economics occur not only when people overestimate their absolute abilities but also when they believe their abilities are better than others', referred to as overplacement or 'better-than-average' bias[13] (DellaVigna, 2009).

Overconfidence can also be related to other forms of optimistic beliefs, such as people's tendency to over-estimate preferred outcomes (Brunnermeier and Parker, 2005; Heger and Papageorge, 2018). It can respond to social stereotyping and exhibit gender effects with men displaying some higher overconfidence, particularly prominent in competitive tasks (Bordalo et al., 2019; Buser et al., 2020; Charness et al., 2018).

We are interested in two possible implications of overconfidence in the context of redistributive preferences. First, we argue that greater pre-tax inequality can boost overplacement of the income position; in a world where incomes are close to each other little reason for overestimation exists. The larger a society's median-mean income gap, the greater is the motivation for being overconfident.

Second, we suggest that overconfidence can weaken or even reverse the ramifications of the POUM hypothesis. The argument rests on the mean marking the theoretical tipping point in any income distribution between beneficiaries (below-mean or "poor") and benefactors (above-mean or "rich") of direct redistribution. While the

[12] Most existing studies focus on inter- rather than intra-generational mobility. Cheng and Wen (2019) also remind that a difficulty to reconcile different results can arise from the multi-faceted nature of income mobility, with transition probability matrices that cannot be easily consigned to a unique measure (Fields and Ok, 1999; Jäntti and Jenkins, 2015).

[13] According to Logg et al. (2018), the better-than-average bias ought to apply to cases where the majority of people claim they are better than the median, while overplacement refers to situations in which a person thinks to obtain a higher ranking on a task test than she does. Note that parallel to the diversity of definitions, the literature maintains a variety of approaches to estimate overconfidence. E.g., Benoît and Dubra (2011) demonstrate that in some cases, beliefs and behaviors consistent with overconfidence can result from a population of rational Bayesians with incomplete information regarding their abilities. Yet, recent evidence by Benoît et al. (2015) show that true overconfidence is also robust to such Bayesian critique (see also Cheung and Johnstone, 2017).

poor ought to demand less redistribution with growing mobility, the rich hold rational preferences for more redistribution, considering the enhanced probability to move below average in tomorrow's income distribution. Which of the two opposing forces dominates depends on the shape of the distribution.

Assuming limited risk aversion, concavity of the transition function, and a minimum length of the redistributive rule, Checchi and Filippin (2004) experimentally confirm a negative relationship between mobility and stated tax rates for the standard case of right-skewed distributions as predicted by the POUM hypothesis.

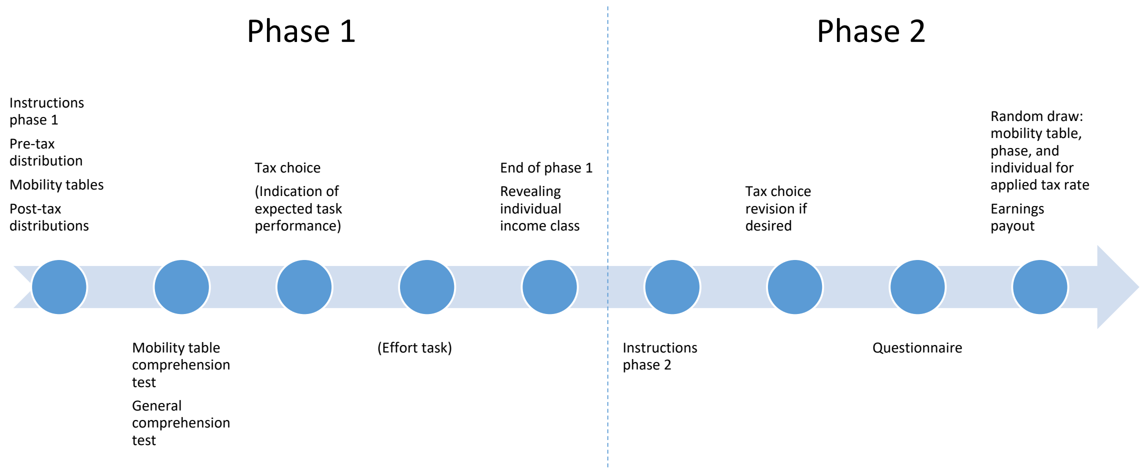
The distribution of expected incomes, and thus the aggregated losses of the rich and gains of the poor are invariant of the distribution skewness but solely determined by the concavity of the transition function. Yet, the gains of the poor concentrate on fewer people with decreasing skewness. Given growing mobility, a lower share of beneficiaries today means less demand for redistribution tomorrow.

In contrast to the standard POUM framework assuming certainty about current income positions, we invite the reader to think of a state in which there is uncertainty about people's current positions. Under the premise of overconfidence, uncertainty should inflate the fraction of (perceived) benefactors reversing the negative relationship between mobility and demand for redistribution.

Given people believe to be more able than others and deserve being rewarded for this, they will prefer economic positions invariably awarded on merit. Put differently, they will prefer a rather rigid society in determining social ranking as too much mobility could harm their economic ambitions. One could conjecture that overconfident individuals dislike redistribution under low mobility, while they may be more neutral or even supportive of some redistribution under high mobility. We refer to this hypothesis as the reverse POUM effect.

2.3 Experiment

We run a laboratory experiment to study people's preference for redistribution conducted at the CERME lab of the Ca' Foscari University in April 2018 using z-Tree (Fischbacher, 2007). A total of 160 college students (mainly from the economics faculty) were recruited and assigned randomly via ORSEE (Greiner, 2015) to one of eight sessions with 20 participants each. The sample averages 21 years and is fairly gender-balanced (55% males).



Steps in brackets are only relevant for the effort condition.

Figure 2.1: Experimental flow.

2.3.1 Procedure

The procedure of our experiment is similar to Durante et al. (2014), in which subjects are asked to state their preferences for redistribution under different conditions, both within- and between-subject. Figure 2.1 illustrates the flow of a typical session which lasted 75 minutes and was divided into Phase 1 and Phase 2. Each of the two phases asked participants to choose two tax rates for redistributive transfers (hence, four tax rates in total) that could affect their own payoff and the payoffs of the other participants.[14]

Phase 1

At the beginning of each session, the experimenters hand out instructions for Phase 1 and read them aloud to subjects. The instructions mention a division into two phases and that only one of the phases, randomly chosen, counts for payment. However, the content of Phase 2 is not further specified until its very start.[15]

Pre-tax income distribution

The instructions of Phase 1 explain participants that they and the other 19 participants form a society with a pre-tax income distribution consisting of five income classes X_i, $i = 1,2,...,5$ and four subjects per class. They further learn that the income distribution represents one of the world's top 40 national economies scaled down to an average income $\overline{X} = 10$ €. The actual country's income distribution varies among

[14] Subjects received an additional show-up fee of 5 €.

[15] The instructions for the two phases are in Appendix 2.A.

two between-subject treatments. Note, whereas the instructions inform subjects of the pre-tax income distributions, their own income assignment remains unknown to them until the end of Phase 1. They only know the assignment mechanism of income positions, which could either be random or based on an ability task in two further between-subject treatments.

Mobility process

The instructions further inform participants that the pre-tax income distribution lasts for two periods t=0 and t=1. The payoffs received by each subject at the end of the experiment will be determined by a tax rate chosen at random among all tax rates indicated by participants and applied to the final period (t=1) pre-tax income distribution. Subjects learn that the income levels remain constant between periods, but subjects' initial income assignments can change between periods according to a specific mobility process, which translates incomes $X_{i,0}$ in the initial period t=0 into incomes $X_{i,1}$ in t=1 (see Section 2.3.2). A mobility table or transition matrix M specifies the probability $p_{i,j}$ to move from a certain income class i in t=0 to an income class j in the final period t=1.

$$M = \begin{bmatrix} p_{11} & \cdots & p_{15} \\ \vdots & \ddots & \vdots \\ p_{51} & \cdots & p_{55} \end{bmatrix} = (p_{i,j}) \in \mathbb{R}_{+}^{5\times 5} \qquad \sum_{i=1}^{5} p_{i,j} = 1 \quad \forall\ j = 1, 2, ..., 5$$

Income distributions in t=0 and t=1 are identical as no economic growth is included. Phase 1 presents subjects with two mobility tables and asks them to indicate their preferred tax rates, one for each mobility process. If the computer selects Phase 1 for payment, one of the two mobility tables will be randomly picked and used to determine the second period pre-tax distribution to which the tax rate chosen is then applied.

Taxation

The last piece of information given in Phase 1 concerns the tax and transfers system used to determine the post-tax income distribution. The experiment applies a standard formula to tax incomes, based on full and equal redistribution of collected tax revenues among all income classes:

$$Y_j = X_j - \tau X_j + \frac{1}{20}\tau \sum_{k=1}^{20} X_k = (1-\tau)X_j + \tau\overline{X} \tag{2.1}$$

where X_j is a subject's pre-tax income class after transitioning, τ the applied tax rate, and Y_j her post-tax income counting for payment. Albeit the instructions list the formula only in the appendix, they include post-tax distributions (Table 2.2) generated by applying τ, ranging from 0% to 100% in increments of 10%, to the pre-tax distribution.

Before indicating their preferred tax rates for the two mobility tables, subjects have to pass two comprehension tests to ensure their understanding of the experiment. One test is directed at the mobility treatment. The other guarantees participants' comprehension of the overall procedure. Subjects can not advance to the main experiment until both tests are correctly answered. Upon successful completion, they express their preferred tax rates, not knowing their initial pre-tax incomes. As indicated, the latter are disclosed not before the submission of tax choices at the end of Phase 1 (see Fig. 2.1).

Phase 2

After uncovering initial pre-tax incomes, the computer program informs subjects about the start of Phase 2. Ultimately knowing their initial income position they are asked again for their preferred tax rate in association with the two mobility tables. Whether experimental earnings are based on subjects' choices in Phase 1 or in Phase 2 is randomly determined at the end of the experiment, together with the mobility table and the corresponding tax rate of one participant. Before receiving payments in cash, subjects complete a questionnaire, including demographic information.

2.3.2 Treatments

We use the procedure outlined in Figure 2.1 to implement four experimental variations, two between-subject and two within-subject.

Phase 1/income uncertainty vs. Phase 2/certainty (within-subject)

The first source of variation distinguishes whether subjects are uninformed or informed of their initial incomes. In this chapter, we will refer interchangeably between Phase 1/Phase 2 and uncertainty/certainty of initial positions. Some authors (e.g., Durante et al., 2014) tend to assimilate the condition in Phase 1 with choice under the *veil of ignorance*. Since the latter term directs to a construct used in political philosophy to induce a position of impartiality in judging distributions from a purely normative perspective, we refer to the uncertainty condition of initial incomes rather not by the veil of ignorance.

Low (Switzerland)/high (South Africa) inequality, between-subject

The second source of variation concerns the level of pre-tax inequality. To study the effect of pre-tax inequality on preferences for redistribution we opt for a between-subject design avoiding any carry-over effects. Comparing national pre-tax inequality on a global level identified South Africa (highest inequality index, World Bank, 2014) and Switzerland (lowest, Bundesamtfür Statistik, 2017) as the two extremes. Table 2.1 replicates the real pre-tax distributions in quintiles scaled down to an average income of 10 €. One-half of the subjects are exposed to the Switzerland pre-tax distribution (CH, low inequality) and the other half to the South African (ZA, high inequality). Countries are not specified to exclude any confounding effects.

Table 2.1: Pre-tax income distributions

Income quintile	Switzerland (Low inequality)	South Africa (High inequality)
1	4.80 €	1.25 €
2	7.15 €	2.35 €
3	8.90 €	4.00 €
4	11.25 €	7.95 €
5	17.95 €	34.45 €
Income average ($\overline{X}$)	10.00 €	10.00 €

Table 2.2: Post-tax distributions in € shown to subjects.

Switzerland

applied tax rate τ	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
1^{st} quintile	4.80	5.32	5.84	6.36	6.88	7.40	7.92	8.44	8.96	9.48	10.00
2^{nd} quintile	7.15	7.44	7.72	8.01	8.29	8.58	8.86	9.15	9.43	9.72	10.00
3^{rd} quintile	8.90	9.01	9.12	9.23	9.34	9.45	9.56	9.67	9.78	9.89	10.00
4^{th} quintile	11.25	11.13	11.00	10.88	10.75	10.63	10.50	10.38	10.25	10.13	10.00
5^{th} quintile	17.95	17.16	16.36	15.57	14.77	13.98	13.18	12.39	11.59	10.80	10.00
Income ratio $5^{th}/1^{st}$	3.74	3.23	2.80	2.48	2.55	1.89	1.66	1.47	1.29	1.14	1

South Africa

applied tax rate τ	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
1^{st} quintile	1.25	2.13	3.00	3.88	4.75	5.63	6.50	7.38	8.25	9.13	10.00
2^{nd} quintile	2.35	3.12	3.88	4.65	5.41	6.18	6.94	7.71	8.47	9.24	10.00
3^{rd} quintile	4.00	4.60	5.20	5.80	6.40	7.00	7.60	8.20	8.80	9.40	10.00
4^{th} quintile	7.95	8.16	8.36	8.57	8.77	8.98	9.18	9.39	9.59	9.80	10.00
5^{th} quintile	34.45	32.01	29.56	27.12	24.67	22.23	19.78	17.34	14.89	12.45	10.00
Income ratio $5^{th}/1^{st}$	27.56	15.03	9.85	6.99	5.19	3.95	3.04	2.35	1.80	1.36	1

Note: Countries' names and income ratios are not shown to participants.

Random vs. effort based income assignment (between-subject)

The second between-subject variable adds alternative allocation of pre-tax income. While in one-half of the sessions, the computer assigns pre-tax incomes randomly to subjects, in the other half, they are based on relative performance in a real effort game (Gill and Prowse, 2012). Within two minutes, subjects have to place as many sliders as possible in the center of a bar. In addition, subjects indicate their expected income quintile beforehand according to their expected performance. Although Gill and Prowse (2019) argue for the superiority of the slider task in comparison with other real-effort tasks, we intend to minimize task-specific effects on subjects' performance expectations by placing it after the first tax choice (Fig. 2.1).[16]

Low vs. high income mobility (within-subject)

One purpose of our experiment is to analyze the impact of actual mobility on preferences for redistribution. Accordingly, subjects state their preferred tax rate in both phases of the experiment for two 5x5 income transition matrices (Table 2.3). In the low (high) mobility matrix, much (little) weight lays on the diagonal resulting in a low (high) likelihood to move upward and downward in the income distribution. In fact, the low mobility matrix does not entail the POUM hypothesis - the expected income after the final period for subjects with a median income in the initial period (i.e., subjects in the third quintile) lingers below average - while the high mobility matrix does.[17] Instructions emphasize that the pre-tax distributions remain constant across the mobility process comprising four subjects per quintile at all times. The choice of a (5x5) mobility table is partly due to making the matrix functioning intelligible to subjects and partly because of maintaining a sufficient degree of real-world diversity of social classes. Feedback from a pilot experiment in which we tested subjects' comprehension of mobility tables confirmed our considerations in favor of 5x5 order matrices.

[16]Alternative real-effort tasks used to allocate income positions in related experiments include manual ability tests, quizzes of general knowledge, or mathematical or linguistic (reading, spelling, etc.) competence games. Results do not seem to vary greatly. Durante et al. (2014) do not find significant performance differences between a Tetris game or a general knowledge quiz.

[17]The low mobility matrix is inspired by existing levels of income mobility in South Africa (Finn and Leibbrandt, 2013). The high mobility table replicates Scandinavian countries (Jantti et al., 2006).

Table 2.3: Income mobility tables

Low mobility

from\to	1^{st} quintile	2^{nd} quintile	3^{rd} quintile	4^{th} quintile	5^{th} quintile
1^{st} quintile	75%	25%	0%	0%	0%
2^{nd} quintile	25%	50%	25%	0%	0%
3^{rd} quintile	0%	25%	50%	25%	0%
4^{th} quintile	0%	0%	25%	50%	25%
5^{th} quintile	0%	0%	0%	25%	75%

High mobility

from\to	1^{st} quintile	2^{nd} quintile	3^{rd} quintile	4^{th} quintile	5^{th} quintile
1^{st} quintile	50%	25%	0%	25%	0%
2^{nd} quintile	25%	25%	25%	25%	0%
3^{rd} quintile	25%	0%	25%	25%	25%
4^{th} quintile	0%	25%	25%	25%	25%
5^{th} quintile	0%	25%	25%	0%	50%

2.3.3 Theoretical predictions

Two separate sessions with 20 participants were devoted to each of the 2x2x2x2 treatments, for a total of 40 observations per condition. Note that two of the treatment variables are within-subject. Table 2.4 summarises all treatments in regard to three classical criteria being the most directly linked to the approaches of rational and social preferences reviewed above. We discuss other possible effects that do not enter the most standard concepts at the section finale.

Profit maximization

The first criterion represents expected payoff maximization (Table 2.4, column 2). It is equivalent to risk-neutral behavior. Accordingly, the optimal tax rate $\tau_{opt,h}$ for any subject h is given by the comparison between the society's average income equal to $\overline{X} = 10$ in all treatments, and the subject's expected income before tax and transfers, denoted by $E_h[X]$. In Phase 1 treatments with random income assignments and all Phase 2 treatments, $E_h[X]$ can be computed for all individuals using objective probabilities. In the random Phase 1 treatments holds $E_h[X]=10$ for all subjects h, implying that any subject can indifferently choose her preferred tax rate τ_h in the interval $[0,1]$, leading to an average tax rate of 0.5. In Phase 2, after subjects have been assigned to an income class, $\tau_{opt,h}$ depends on the respective mobility matrix. Under low mobility, it is $E_h[X] < 10$ and hence $\tau_{opt,h} = 1$, for all subjects in the three lower quintiles, that is 60% of subjects; and $E_h[X] > 10$ hence $\tau_{opt,h} = 0$, for all subjects in the two top quintiles, that is 40% of subjects. Opposite frequencies hold for the high mobility matrix, that is 40% of subjects with $\tau_{opt,h} = 1$ and 60% with $\tau_{opt,h} = 0$. Remark that since the predictions identify the poorer and richer than average in terms of expected future incomes (short *prospective poor* and *prospective rich*), they incorporate the POUM hypothesis directly. In the effort treatments of Phase 1 comparison between $\overline{X}$ and $E_h[X]$ depends on subject h's expected performance in the effort task.

Risk aversion

Column 3 in Table 2.4 provides the predictions from risk aversion. Risk aversion implies that subjects with $E_h[X] < 10$ vote for full redistribution, i.e., $\tau_h = 1$; while subjects with $E_h[X] > 10$ will prefer τ_h greater than 0 by a precise amount depending on the individual risk attitude. Risk aversion can also affect the predictions of the POUM hypothesis when people with $E_h[X] > 10$ prefer redistribution because they fear the chance to move downward, which is more pronounced with high mobility than with low mobility. Moreover, due to risk aversion, the demand for redistribution must be higher when experiencing high inequality (ZA) than low inequality (CH).

Other-regarding preferences

The predictions for social preferences in column 4 of Table 2.4 are based on the model of Charness and Rabin (2002). Durante et al. (2014) also use this model with some modifications (see Appendix 2.C). It expresses people's utility in the form of a linear combination of selfish and social motivations. A person's expected payoff represents the selfish motivation in the original model, while the other-regarding component, relevant in our experiment, is captured by a Rawlsian concern for the person with the lowest post-tax income.[18] The relative weighting of the two motivations points the direction for subjects' preferred tax rates. The predictions are straightforward. Subjects with $E_h[X] < \overline{X} = 10$ always vote for full redistribution ($\tau_h = 1$); subjects with $E_h[X] > 10$ choose either $\tau_h = 0$ or 1 depending on their degree of inequality aversion relative to selfish motivation. In Appendix 2.C, we estimate an extension of the model employed by Durante et al. (2014), which allows distinguishing the two attitudes.

[18] The original model accounts further for efficiency in the social component. Here it is disregarded since the redistribution scheme applied does not include efficiency costs.

Table 2.4: Treatments and predictions.

Treatments				Predictions: Maximization of expected earnings[1]	Predictions: Risk aversion[2] (demand of self-insurance)	Predictions: Social preferences[3]
Phase 1	Random	ZA	Low	Any τ in $[0,1]$	$\tau = 1$	$\tau = 1$
			High			
		CH	Low			
			High			
	Effort	ZA	Low	Predictions are the same as in the treatments of Phase 2, with $E_h[X]$, which depends on subject h's expected performance in effort task.	Predictions are the same as in the treatments of Phase 2, with $E_h[X]$, which depends on subject h's expected performance in effort task.	Predictions are the same as in the treatments of Phase 2, with $E_h[X]$, which depends on subject h's expected performance in effort task.
			High			
		CH	Low			
			High			
Phase 2	Random	ZA	Low	$\tau = 1$ for subjects with $E_h[X] < 10$ (prospective poor); $\tau = 0$ for subjects with $E_h[X] > 10$ (prospective rich).	$\tau = 1$ for subjects with $E_h[X] < 10$; $\tau > 0$ for subjects with $E_h[X] > 10$, size of τ depends on the degree of risk aversion. For the same degree of risk aversion and same income quintile, τ in South Africa expected to be higher than in Switzerland.	$\tau = 1$ for subjects with $E_h[X] < 10$ and for subjects with $E_h[X] > 10$ with high degree of inequality aversion; $\tau = 0$ for subjects with $E_h[X] > 10$ and low degree of inequality aversion. For same degree of inequality aversion and same income quintile, $\tau = 1$ more likely in ZA than in CH.
			High			
		CH	Low			
			High			
	Effort	ZA	Low			
			High			
		CH	Low			
			High			

Notes: The first and the fourth column refer to the within-subject treatments, while the second and the third column to the between-subject treatments. Thus, the 4×4 combinations listed in the fourth column represent the 16 sub-treatments of the experiment.

1) The probability π_{hj} of subject h being in income quintile j after the mobility process (t=1), determining her expected income $E_h[X] = \sum_{j=1}^{5} \pi_{hj} X_j$ is given by: i) $\pi_{hj} = \frac{1}{5}\sum_{i=1}^{5} p_{ij}$ in the random treatments of Phase 1, where p_{ij} is the transition probability of moving from quintile i (t=0) to quintile j (t=1) in the mobility tables (low/high); ii) $\pi_{hj} = \sum_{i=1}^{5} \pi_h^i p_{ij}$ in the real-effort treatments of Phase 1, where π_h^i is the subjective probability of subject h to end-up in income quintile i (t=0) in the real-effort task; and iii) $p_{hj} = p_{ij}$ in all treatments of Phase 2, for any subject h who at the end of Phase 1 is in quintile i.

2) Predictions under risk aversion follow from Jensen's inequality stating that $E_h\left[u_h(Y_h)\right] \leq u_h(E_h\left[Y_h\right])$ for any increasing and concave utility function u_h, where $E_h\left[Y_h\right] = E_h\left[(1-\tau)X_h + \tau\overline{X}\right]$ is subject h's expected post-tax income, so that $u_h\left(E_h\left[Y_h\right]\right) < u_h(\overline{X})$ whenever $E_h\left[X_h\right] < \overline{X}$.

3) Predictions for social preferences are based on the following utility function adapted from Charness and Rabin (2002) and Durante et al. (2014, see Appendix 2.C for a thourough discussion): $V_h = (1-\lambda)E_h[Y_h] + \lambda Y^{min}$, where $E_h[Y_h] = \sum_{j=1}^{5} \pi_{hj} X_j (1-\tau) + \tau\overline{X}$ is subject h's post-tax expected income, $Y^{min} = X^{min}(1-\tau) + \tau\overline{X}$ is the society's minimum post-tax income, and $\lambda \in (0,1)$ is the relative weight of personal versus inequality concern.

Further behavioural drivers

Other issues reviewed in Section 2.2 that are not captured by the analyzed standard theories above can alter the predictions in Table 2.4. First of all, the mechanism of income assignment can pivot preferences towards higher or lower demand for redistribution. Subjects considering the random assignment less fair than the one based on effort are inclined to support more redistribution under the first than under the second condition; the opposite applies for individuals who find some fairness in the random assignment. In the effort treatments of Phase 1, the method of income determination can interact with subjects' expectations about their income position. Mainly, overestimating income positions due to overconfidence can depress the demand for redistribution in the effort treatments of Phase 1, but not in the random treatments. As there are more reasons to overestimate when incomes are further away than when they are closer to each other, one question is whether overconfidence grows in pre-tax inequality. A second question concerns the possibly defusing effect of income mobility on overconfidence, resulting in lower demand for redistribution under low than high mobility (the reverse POUM effect).

2.4 Analysis and results

This section presents descriptive statistics and an econometric analysis of the tax choices to conclude with the main findings of the study.[19]

2.4.1 Descriptive analysis

Figure 2.2 reports the distributions of tax rates chosen by participants. The top panels show the distributions in treatments of Phase 1, and the bottom panels these of Phase 2. The vertical red bars mark average tax rates.

Phase 1

Tax rates in most treatments of Phase 1 are not statistically different from 0.5.[20] Two exceptions represent the low mobility treatments with effort, where distributions skew slightly to the right, and average tax rates are just above 0.4 both for

[19] Appendix 2.B reports a series of non-parametric tests in support of the descriptive analysis.

[20] A series of one-sample Kolmogorov-Smirnov tests for discrete distributions (reported in App Table B2) accept the null-hypothesis of uniform distributions in all but three treatments of Phase 1, i.e., the Switzerland treatment with random assignment and high mobility ($p = 0.070$), and the South Africa treatments with effort (low mobility and high mobility: $p = 0.030$ and $p = 0.021$, respectively).

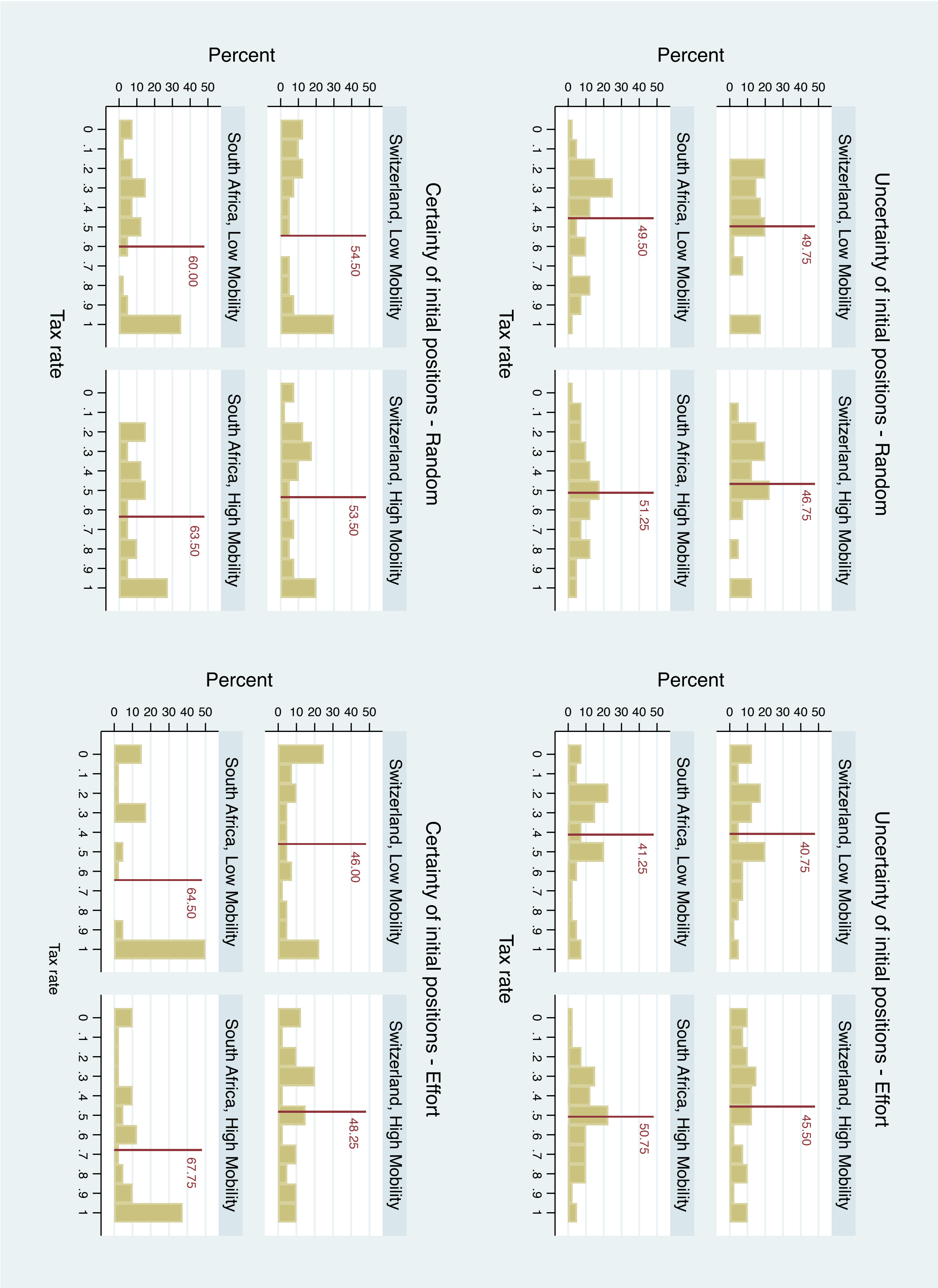


Figure 2.2: Tax choice histograms.

the Switzerland and South Africa condition. Uniform distributions are consistent with the prediction by expected payoff maximization: subjects choose with equal probability any τ between $[0, 1]$. Yet, the uniformity is inconsistent with inequality aversion and risk aversion (compare Table 2.4). Tax rates even below 0.5 in the effort treatments under low mobility but not under high mobility are suggestive of overconfidence. On the other hand, it advocates an effect which we have termed as the reverse POUM hypothesis, namely that under high mobility, people's overconfidence is offset by a higher chance of moving downward in the mobility process. Below we investigate this conjecture in depth.

Phase 2

After lifting uncertainty of initial positions in Phase 2, we can easily detect an increase in tax rates through tobit regressions (Table 2.6). Average rates between Phase 1 and Phase 2 rise by 15 percentage points ($p < .001$, Model 1). The South Africa treatments drive the effect (+24%, $p < .001$, Model 2), while the +5% increase for Switzerland in Model (3) bears no significance.

The distributions in Phase 2 take on different shapes as compared to those in Phase 1. In line with the theoretical predictions in Table 2.4, weights move toward the ends of the tax scale, i.e., fat tails. There are, however, notable deviations from the strict predictions of corner solutions. First and foremost, we observe very few choices of $\tau = 0$ across all treatments (overall 11% in Phase 2). These frequencies are lower than predicted by the principle of expected profit maximization. Whereas the three criteria specified in Table 2.4 project full redistribution ($\tau = 1$) for at least 60% (40%) of subjects in the low (high) mobility treatments of Phase 2, the data reports instead overall shares of 34%, respectively 24%. Only in the low mobility treatments of South Africa, the proportions obtain the majoritarian share of 50%.

The frequency of tax choices $\tau = 1$ in Phase 2 is for corresponding conditions always greater in the low than in the high mobility treatment. While this is consistent with the rationale behind the POUM, the hypothesis itself is even stronger as it requires that under low mobility, more people demand τ=1 than τ=0, but the opposite under high mobility. The histograms of Phase 2 clearly contract this.

Post-tax inequality

Table 2.5 reports the level of post-tax inequality obtained by applying the majoritarian rule (median tax rates) in the treatments. Two results are worth noticing. The first confirms that also post-tax inequality based on median tax rates amounts between corresponding treatments always higher in Phase 1 than in Phase 2. This

Table 2.5: Redistribution and post-tax inequality under majority rule.

Switzerland (CH)	U_R_LM	U_R_HM	U_E_LM	U_E_HM	C_R_LM	C_R_HM	C_E_LM	C_E_HM
Median tax rate	0.40	0.40	0.40	0.40	0.50	0.50	0.40	0.50
Post-tax income ratio $5^{th}/1^{st}$	2.55	2.55	2.55	2.55	1.89	1.89	2.55	1.89
South Africa (ZA)	U_R_LM	U_R_HM	U_E_LM	U_E_HM	C_R_LM	C_R_HM	C_E_LM	C_E_HM
Median tax rate	0.40	0.50	0.40	0.50	0.50	0.60	0.90	0.80
Post-tax income ratio $5^{th}/1^{st}$	5.19	3.95	5.19	3.95	3.95	3.04	1.36	1.80

Acronyms are as follows: the first digit is U/C for uncertainty (Phase 1)/certainty (Phase 2) of initial positions; the second is R/E for random/effort income assignment; the third is LM/HM for low/high mobility. E.g., U_R_LM stands for the treatment with uncertain income position (Phase 1), random income assignment, and low mobility.

rejects the prediction of more demand for redistribution with less income certainty. The second result follows from comparing CH to ZA treatments. Since there are virtually no differences in tax choice distributions between corresponding treatments in Phase 1,[21] taxation corrects only partially for varying pre-tax inequality in the two countries. In all Phase 1 treatments, post-tax inequality, measured by the ratio between the highest and lowest income quintile, is twice as large in South Africa than in Switzerland. The same holds for the random treatments of Phase 2. Only in the effort treatments of Phase 2, we find comparable levels of post-tax inequality between the two countries.

We come back to the evidence on post-tax inequality after having investigated subjects' choices using regression analysis. We start with Phase 1.

2.4.2 Tax choices when initial incomes are unknown - Phase 1

Table 2.6 reports the estimates of various tobit regressions. The dependent variable in all models is the tax rate τ_h, chosen by subject h.[22]

Model (1) tests for the general effect of uncertain initial income positions. Consistent with the histograms in Figure 2.2, the Phase 1 dummy indicates a negative effect from income uncertainty on tax rates. The effect is highly significant, with a reduction in average tax rates between Phase 1 and Phase 2 by almost 15 percentage points.[23] The gender dummy on the whole experiment is statistically not significant.

Mobility effects

Model (4) regresses preferred tax rates in Phase 1 on three dummies for, in the order,

[21] A series of unmatched Mann-Whitney-Wilcoxon rank-sum tests (App Table B4) confirm the visual inspection from the histograms in Fig. 2.2, indicating significantly different tax rate distributions between CH and ZA only for the effort treatments of Phase 2, both with low and high mobility.

[22] Since there are both between-subject and within-subject treatments, standard errors in the following regressions are always adjusted for correlation within subject h's responses.

[23] Note the difference between the tobit estimate and the one based on sample averages (+11% reported in Section 2.4.1), due to the correction for censored data.

Table 2.6: Tobit regressions - Phase 1: Uncertainty of initial positions

	Phase 1 and Phase 2			Phase 1				
	(1) All	(2) ZA	(3) CH	(4) All	(5) Random	(6) Effort	(7) Effort	(8) Effort
Dependent variable	τ_h	τ_h	τ_h	τ_h	τ_h	τ_h	τ_h	τ_h
Income uncertainty	-0.148***	-0.242***	-0.053					
	(0.034)	(0.049)	(0.045)					
High inequality (ZA)	0.103**			0.014	-0.013	0.022	0.027	0.057
	(0.052)			(0.043)	(0.059)	(0.064)	(0.066)	(0.065)
Low mobility	-0.030	-0.054**	-0.006	-0.046**	-0.013	-0.080**	-0.087**	-0.081**
	(0.020)	(0.025)	(0.032)	(0.023)	(0.028)	(0.036)	(0.040)	(0.040)
Effort	-0.044	0.019	-0.111	-0.053				
	(0.053)	(0.075)	(0.078)	(0.043)				
Prospective poor from self-ass.							0.032	0.153*
							(0.065)	(0.089)
Prosp. poor from self-ass. × Female								-0.254**
								(0.119)
Female	-0.024	-0.016	-0.047	-0.030	-0.120**	0.065	0.058	0.144*
	(0.051)	(0.075)	(0.074)	(0.042)	(0.059)	(0.063)	(0.065)	(0.083)
Constant	0.614***	0.741***	0.597***	0.524***	0.564***	0.445***	0.437***	0.388***
	(0.060)	(0.077)	(0.081)	(0.046)	(0.060)	(0.057)	(0.060)	(0.065)
Observations	640	320	320	320	160	160	160	160
Log-pseudolikelihood	-442.569	-212.214	-224.090	-116.290	-45.129	-66.097	-65.940	-63.264
Pseudo R-squared	0.033	0.061	0.014	0.020	0.072	0.033	0.035	0.074

* p < .1, ** p < .05, *** p < .01.

high inequality, low mobility, and whether subjects are in the random or in the effort treatments. Alone the mobility dummy is negatively significant. Yet, its sign is inconsistent with the POUM hypothesis, predicting if agents are not too risk averse and hold rational expectations, tax rates under low mobility must be greater than under high mobility. The evidence can instead be consistent with a reverse POUM effect driven by overconfidence (see subsequent paragraphs).

Random vs. effort treatments

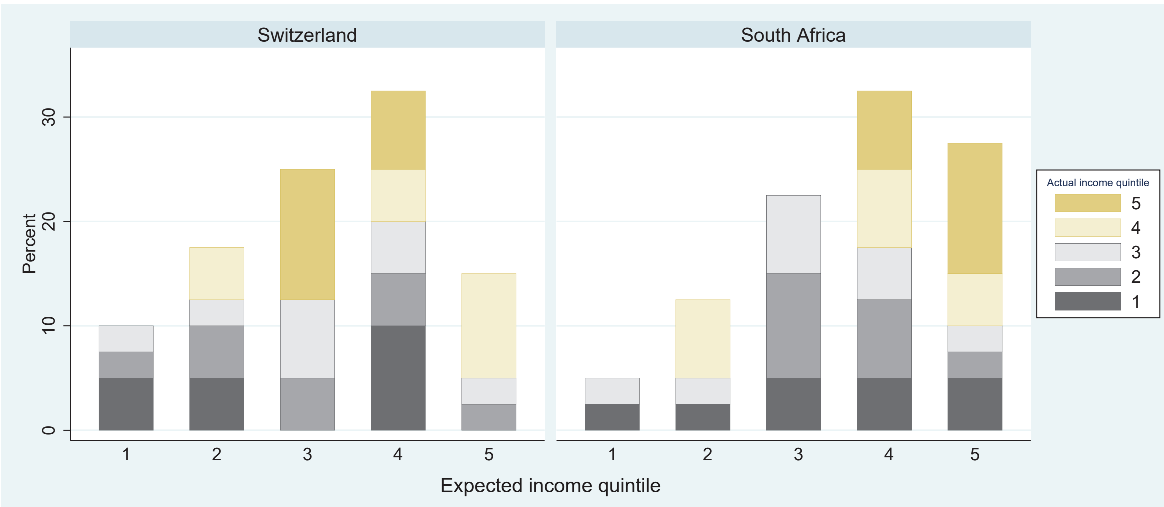
Models (5) and (6) distinguish random from effort treatments in Phase 1. In the random treatments, exclusively, the gender dummy is statistically significant, indicating a lower demand for redistribution by women than men. Related studies find often the opposite (e.g., Alesina and Giuliano, 2011). Here the effect stems from a lower proportion of women choosing full redistribution ($\tau = 1$).[24] Overall, choices in the random treatments are largely consistent with risk neutrality.[25] In line with

[24] 15 out of 160 observations in the random treatments of Phase 1 report $\tau = 1$, of which 14 were stated by males and 1 by a female. Omitting the upper corner solution, the average tax rate for $\tau \in [0, 1)$ is equal to 0.43 for males and females. We will return to the evidence that women are more reluctant than men to vote for full redistribution when discussing the results from Phase 2.

[25] Estimated average tax rates for men are .56, and .44 for women, not statistically different from 0.5 at $p = .288$ and $p = .191$, respectively.

the credited effect of overconfidence, the impact of low mobility bears significance alone in the effort treatments (Model 6). The gender dummy in Model (6) is not statistically significant, while the negative impact of low mobility cuts the estimated average tax rate from .45 down to .37 (significantly less than 0.5, p<.012). Interestingly, the dummy for South Africa is also not significant. Different from the random treatments, however, this cannot be explained by mere risk neutrality.[26] It can be further attributed to overconfidence being higher in South Africa than in Switzerland, compensating for greater pre-tax inequality, thus, larger stakes for high performers in the effort task. The latter conjecture is also confirmed by subjects' self-assessments about expected performances (Fig. 2.3). Accordingly, subjects overestimate the probability of obtaining better-than-average income quintiles while underestimating the probability of obtaining lower-than-average income quintiles. Effects are large and significant for South Africa and men; they are less pronounced

Figure 2.3: Subjects' self-assessments in the effort task.



Note: In Switzerland the overall proportions of lower-than-average (quintiles 1 and 2) and better-than-average (quintiles 4 and 5) self-assessments are .28, respectively .48 (difference-of-proportion test $d = 1.85$, $p = .064$). In South Africa the proportions in the same order are 0.18 and 0.60 ($d = 3.90$, $p < .001$). For females the proportions are 0.25 and 0.50 ($d = 1.264$, $p = .206$) in Switzerland; and 0.285 and 0.333 ($d = 0.334$, $p = .739$) in South Africa; for males the proportions are .29 and .46 ($d = 1.38$, $p = .167$) in Switzerland; and .05 and .90 ($d = 5.198$, $p < .001$) in South Africa.

[26]Note that the argument of risk neutrality postulates that a subject h's expectations about her performance in the effort task imply $E_h[X] = 10$ (estimating chances to end up in any income class equally likely). In this case, however, the dummy for low mobility ought to be insignificant either (as it is the case in the random treatments).

for Switzerland and generally not significant for females.[27]

Self-assessments must be taken with care, though. They are not incentivized and based on subjects' expectations to obtain one specific quintile rather than on the distribution of subjective beliefs to end up in any quintile. Moreover, different psychological drivers can affect stated self-assessments and actual decisions, e.g., in forms of wishful thinking and optimism (Heger and Papageorge, 2018; Schwardmann and Van der Weele, 2019).

Prospective poor vs. prospective rich

It is nevertheless worth checking how self-assessments relate to tax choices. To this end, we compute subjects' expected pre-tax income $E_h[X]$ based on the self-assessments and separate the subjects with $E_h[X] < 10$ from those with $E_h[X] > 10$. We term the first group as 'prospective poor from self-assessment' and the second as 'prospective rich from self-assessment'. Figure 2.4 shows the distribution of stated tax rates for the two groups in CH and ZA treatments. Minor differences suggest that stated self-assessments by themselves cannot explain the preferred tax rates.

Nonetheless, a dummy for 'prospective poor' bears no significance in Model (7), Table 2.6. Since fewer females exhibit overconfidence in self-assessments, Model (8) interacts dummies for female and 'prospective poor' to improve estimates. While both dummies for 'prospective poor' and female turn positive now, the interaction term becomes negative by an amount which compensates almost entirely for the two positive dummies ($0.153 + 0.144 - 0.254 = 0.043$, $p = .581$).[28]

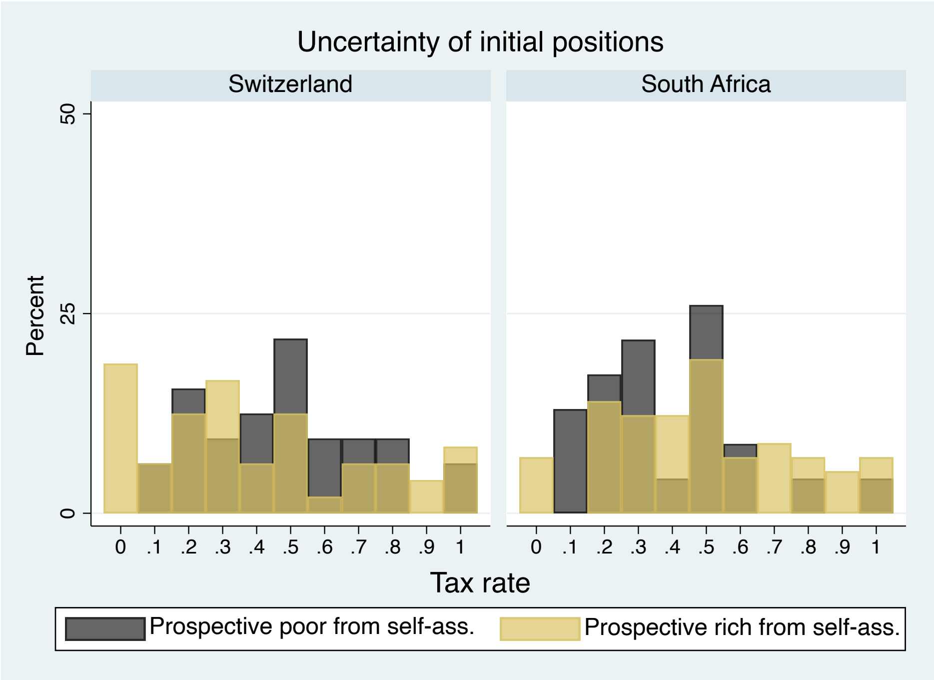
This indicates that women are less exuberant than males to inflate stated self-assessments in the experiment. Tobit regressions in Appendix Table D2 confirm gender differences in the degree of overconfidence (cf. Buser et al., 2020). A measure for revealed overconfidence in Phase 1, comparing subjects' expected income quintiles to realized income quintiles, finds a significant negative influence on the chosen tax rate among males, whereas the contrary for females.[29] Further specifying *critical overconfidence* as expecting an above mean income while realizing a below mean income, Model (8) states that a pronounced bias among men leads to significantly reduced tax preferences. Although similarly prone to overoptimism about income quintiles, women of this type indicate substantially higher tax rates. The

[27]The vast majority of subjects (76%) indicate that they are 'fairly confident' about their self-assessments, 18% are 'very confident', and 6% are 'not confident at all'.

[28]Remark that in this case, the lower tendency of women to vote for $\tau = 1$ cannot explain the negative effect of 'prospective poor' females, since in all effort treatments of Phase 1, only 11 subjects vote for full redistribution (five male and six female participants).

[29]This overconfidence measure was suggested by Rupert Sausgruber.

Figure 2.4: Prospective poor/rich self-assessed in effort treatments of Phase 1



overconfidence measures cannot explain the general tax rate jumps when moving from Phase 1 to Phase 2 (App. Table D2, Models 1-4).

The evidence supports the idea that stated self-assessments capture only part of the effect arising from overconfidence, optimism, or wishful-thinking. We sum up the evidence of Phase 1 as follows.

Result 1. *When subjects are uncertain about initial income positions, self-interest and risk neutral behavior dictate their demand for redistribution, while risk aversion and inequality aversion are rejected as explanatory drivers. That means subjects choose similar tax rates under low inequality and high inequality in treatments with random income assignment. When initial income positions depend on real-effort, subjects' choices are further affected by general overconfidence. This biased behavior sets in stronger under high inequality (South Africa) than under low inequality (Switzerland) realizing in similar low tax rates across the two countries and causing a reverse POUM effect after all.*

2.4.3 Tax choices when initial incomes are known - Phase 2

Table 2.7 provides results of tobit regressions of Phase 2, after that subjects are informed of initial positions but before the transition process determines subjects' final pre-tax incomes. Model (1) encompasses only a dummy for gender, which is

Table 2.7: Tobit regressions - Phase 2: Certainty of initial positions.

	(1) Phase 2	(2) Phase 2	(3) Phase 2	(4) Phase 2 Random	(5) Phase 2 Effort
Dependent variable	τ_h	τ_h	τ_h	τ_h	τ_h
High inequality (South Africa)		0.232***	0.237***	0.124	0.330***
		(0.087)	(0.070)	(0.087)	(0.111)
Low mobility		-0.006	-0.133***	-0.109**	-0.158**
		(0.038)	(0.039)	(0.050)	(0.062)
Effort		-0.032	-0.038		
		(0.086)	(0.069)		
Prospective poor			0.610***	0.533***	0.697***
			(0.070)	(0.087)	(0.113)
Female	0.010	-0.022	-0.077	-0.224**	0.081
	(0.086)	(0.085)	(0.068)	(0.087)	(0.107)
Constant	0.636***	0.554***	0.337***	0.481***	0.165*
	(0.065)	(0.083)	(0.066)	(0.080)	(0.086)
Observations	320	320	320	160	160
Log-pseudolikelihood	-291.393	-285.176	-232.401	-103.544	-120.593
Pseudo R-squared	0.000	0.021	0.202	0.221	0.228

* p < .1, ** p < .05, *** p < .01.

statistically not significant. The constant (0.636) exceeds 0.5 ($p = .002$) indicating an influence of risk aversion and inequality aversion in Phase 2.[30]

Model (2) adds dummies for high inequality, low mobility, and effort-based income assignment. The dummy for inequality is highly significant, indicating that tax rates stated in Phase 2 are 23% higher in ZA than in CH. The dummies for effort and low mobility do not exert significant effects, violating the POUM hypothesis.

Prospective poor vs. prospective rich

Model (3) abandons risk neutrality. To this end, the model adds a dummy for the 'prospective poor' ($E_h[X] < 10$). Were the data fully explained by maximization of expected earnings, the coefficient on the dummy ought to be 1 while all other coefficients should be 0, including the constant. Instead, the dummy coefficient is significantly less than 1 ($p < .01$) and the constant significantly greater than 0.

The positive constant implies that less 'prospective rich' ($E_h[X] > 10$) than predicted by maximization of expected earnings vote for zero redistribution, while the

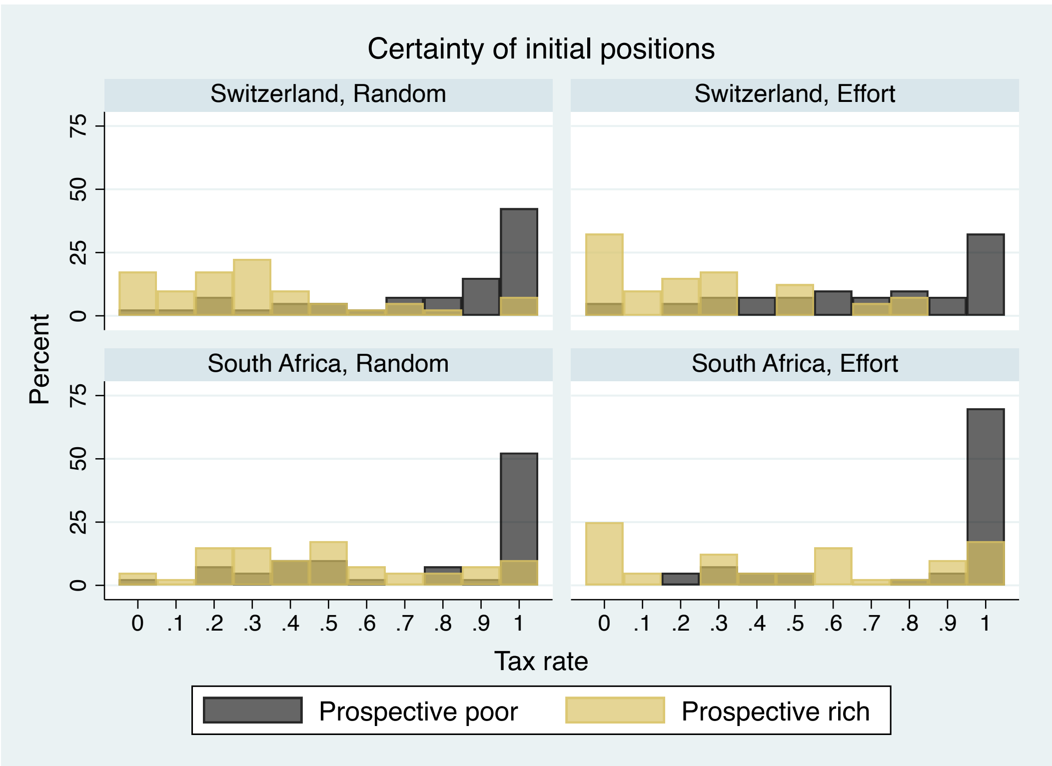
[30]Strictly speaking, either risk aversion or inequality aversion alone can explain demand for redistribution greater than predicted by the principle of expected profit maximization. Yet, estimations of the structural preferences model based on Charness and Rabin (2002) and Durante et al. (2014) in Appendix 2.C confirms that both risk aversion and inequality aversion are present in Phase 2.

dummy coefficient lower than 1 indicates that less 'prospective poor' than predicted vote for full redistribution. The first observation can be explained with risk aversion or inequality aversion, while the second violates both theories. Moreover are the dummies on high inequality (positive) and low mobility (negative) highly statistically significant, which can be, at least for the latter, explained by risk aversion.

Models (4) and (5) conduct separate regressions for the random and the effort treatments that surface notable effects. The dummy for high inequality is only significant across effort treatments; similarly, the dummy coefficient for 'prospective poor' is higher and the constant lower in the effort than in the random treatments. These results signal that the redistributive conflict between 'prospective poor' and 'prospective rich' deems more polarized in effort than in random treatments, particularly so in the high inequality context. Similarly to Phase 1, there is a negative gender effect in the random treatments due to a lower propensity of females to vote for full redistribution. The polarization of preferences between 'prospective rich' and 'prospective poor' in Phase 2 also becomes visible in Figure 2.5, showing distributions of reported tax rates for the two income groups divided between low/high inequality and random/effort assignment.

Figure 2.5 highlights the demand for full redistribution by the 'prospective poor'

Figure 2.5: Prospective poor and prospective rich in Phase 2.



and the opposition to any redistribution by the 'prospective rich'. Table 2.8 deepens the analysis of such behavior employing probit regressions separated for random and effort assignment. The dependent variable in the probit model for 'prospective poor' is 1 if subject h reports a tax rate $\tau_h = 1$, and 0 otherwise; while in the probit model for 'prospective rich', the dependent variable is 1 if subjects choose $\tau_h = 0$, and 0 otherwise.

The probit models of Table 2.8 and the histograms in Figure 2.5 illustrate that 'prospective rich' are more likely to be fully selfish in effort treatments (Model 4) than in random treatments (Model 2). The predicted probability of $\tau = 0$ at mean values is 27% in the effort compared to 6.9% in the random treatment. 'Prospective rich' become also less selfish in the random treatment with growing inequality (marginal effect of ZA in Model (2) is -13%). Less so in the effort treatment. Conversely, in the effort treatment 'prospective poor' are more ready to impose full expropriation over the rich in South Africa than in Switzerland (marginal effect of ZA in Model (3) is +36%). The positive sign of the low mobility dummies in Models (2) and (4) for 'prospective rich' in the effort and the random treatment is in line with risk aversion. The dummies for 'prospective poor' are positive but only marginally significant in the random treatment (Model (1)).[31] As anticipated, the gender dummy is significantly negative in Model (1) with large marginal effects (-36%). In contrast, gender does not play a significant role in effort treatments, which

Table 2.8: Probit regressions - prospective poor and prospective rich in Phase 2.

	(1) Phase 2 - Random Prosp. poor	(2) Phase 2 - Random Prosp. rich	(3) Phase 2 - Effort Prosp. poor	(4) Phase 2 - Effort Prosp. rich
Dependent variable	$\Pr(\tau_h = 1)$	$\Pr(\tau_h = 0)$	$\Pr(\tau_h = 1)$	$\Pr(\tau_h = 0)$
High inequality (South Africa)	0.298	-0.836*	1.054***	-0.227
	(0.385)	(0.432)	(0.395)	(0.349)
Low mobility	0.326*	1.056***	0.365	0.722***
	(0.181)	(0.395)	(0.223)	(0.272)
Female	-1.068***	-0.247	-0.155	-0.354
	(0.389)	(0.435)	(0.395)	(0.329)
Constant	0.142	-1.375***	-0.637**	-0.658**
	(0.346)	(0.417)	(0.318)	(0.292)
Observations	80	80	80	80
Log-pseudolikelihood	-47.767	-23.015	-48.841	-44.228
Pseudo R-squared	0.137	0.182	0.119	0.078

* p < .1, ** p < .05, *** p < .01.

[31] A positive sign is consistent with the rationale behind the POUM hypothesis. Although the definitions of 'prospective poor' and 'prospective rich' are already based on expected incomes, all 'prospective poor' have better chances to move up under high than low mobility.

corroborates the evidence of similar behavior between males and females when income positions are effort-based (except for stated self-assessments).

Using probit regression to compute predicted probabilities of the 'prospective poor' to demand full redistribution gives for ZA 71% in the effort compared to 53% in the random treatment. For CH, probabilities are 32% in the effort and 32% in the random treatment. While the data confirms general reluctance by the poor to expropriate the rich, it also indicates that this timidity shrinks substantially in effort treatments of high inequality. We sum up the evidence from Phase 2 as follows.

Result 2. *After Phase 2 reveals initial income positions, participants' tax choices are rather apprehended by considerations of risk aversion and inequality aversion, less by risk-neutral behavior. Several other factors influence subjects' preferences, too. First of all, demand for redistribution is overall greater under high than under low inequality, yet only when incomes are assigned through the real-effort task and not randomly. The distributional conflict between 'prospective poor' and 'prospective rich' also emerges more polarized when effort rather than luck determines income positions. Under high inequality, this competition pushes tax rates close to full redistribution. In all other treatments, preferred taxation remains well below full redistribution, including individuals who expect to be poor. This reluctance to expropriate the rich is more pronounced among female participants than male participants.*

2.5 Discussion

The final section recaps the experiment results in light of the redistribution preference models reviewed in Section 2.2. We combine the main findings to elaborate on the idea of an inequality trap and to point out questions for future research.

2.5.1 What shapes redistributive preferences?

A first observation lays open differences in redistributive preferences between the two phases of the experiment. While tax choices in Phase 1 appear to be mainly driven by selfish considerations, behavior in Phase 2 is overall better explained by a mix of motivations, including risk aversion and social concerns. The evidence on Phase 1 is surprising since standard theories of rational choice assume that demand for self-protection increases in uncertainty. Not to mention related work, which argues that missing transparency about the relative standing in the income ladder provokes pro-social behavior (see Section 2.2.2). The idea to judge behind a veil of ignorance has been further taken up by theorists to explain and model people's impersonal or social preferences (e.g., Harsanyi, 1977).

A possible explanation for the lack of social concern in Phase 1 are equal ex-ante chances to clinch any income position among all individuals. This might be perceived as a situation of equal or fair opportunities and, for this reason, a situation society does not need to correct. Given complete uncertainty about one's position, subjects may direct their cognitive attention in Phase 1 rather to 'oneself' than to 'others', leading to more egoistic behavior than when initial positions are known.

Certainly, a relevant 'ego-centric' reason for low tax choices in Phase 1 is subjects' overconfidence about the likelihood of reaching top income positions. The bias is apparent in the effort treatments; in random treatments, overconfidence cannot affect responses. Misperception and over-optimism about the own prospects in life are behavioral effects well-known in the literature and documented in experiments on preferences for redistribution (Durante et al., 2014). We add two relevant features. First, by providing room for overconfidence in the social mobility process, we find that the bias can nullify the concavity implication of the actual mobility process and induce a reverse POUM effect. Second, we observe a positive relationship between the level of inequality in a society and personal overconfidence. Whether due to optimistic expectations or genuine beliefs to be 'better' than others, the effects of overconfidence in the context of redistributive preferences can contribute to an inequality trap, discussed in more detail in the following Section 2.5.2.

When subjects are well informed about income positions, there remains some behavior opposing any common understanding of economic principles. Possibly the most puzzling one represents a sub-sample in the random treatment of Phase 2. Despite being handicapped by initial low-income positions prohibiting them with certainty to leave the below-mean income sector in the mobility process, most of these subjects still refrain from demanding full redistribution. In other words, these subjects voluntarily forgo the chance to equalize earnings regardless of no apparent reason in the random treatment why some subjects are entitled to larger payoffs than others. A large fraction (about 50% on the whole) with an expected income lower than the mean does not support full redistribution either. Classical arguments appealing to the notion of rational agents cannot explain such reluctance to expropriate the rich. It cannot even be attributed to social factors causing secondary divides between subjects benefiting from redistribution. Cultural and ideological factors, including concepts such as reference point theory, instead may provide a plausible interpretation. Subjects enter the lab unconsciously equipped with ideas, attitudes, experiences, and other stimuli. In consequence, anchors based on those experiences and impulses can affect subjects' decisions. For example, as people are not used to full redistribution in reality, they may also abstain from it in the lab. Some could further suppress the aptitude to maximize personal earnings because of strong political views on the illegitimacy of expropriating taxation.

Nonetheless, while these restraints can explain part of the evidence, the effect strength is also dependent on condition and context. In the high inequality/effort treatment of Phase 2, the tax rate preferred by the majority of individuals approaches 100%. This evidence bears further relevance for the debate on the perceived legitimacy of income differences. Some scholars have argued that when income depends on effort rather than luck, earnings are perceived more legitimate, and thus demand for redistribution sinks. Our results in the South Africa treatments find the opposite. When incomes are not assigned randomly but through the effort task, demand for redistribution is higher although not significantly, the distributional conflict is more polarized, and the income gap between poor and rich closes almost entirely through increasing tax rates (Table 2.5). A possible interpretation for less demand in the random treatment is the following: when pre-tax inequality is high prevailing a jackpot, luck is perceived as a form of fairness.

2.5.2 An inequality trap?

We witness increasing social inequality yet decreasing demand for fiscal redistribution dating back to the 1970s. The present analysis shows that greater income inequality nudges people to overestimate their chances of obtaining a high position on the income ladder, leading to declining demand for taxation. Inequality continues to grow. On top of that, because the wealth ambitions of rich aspirants are better preserved under a rigid than a mobile transition matrix, we find that demand for redistribution is further depressed with decreasing mobility. Thus, what can be called an inequality trap does not only affect social welfare today. Its vicious circle works like a downward spiral against the 'self-regulating power' of democracy - economic redistribution - to adjust inequality tomorrow.

Is there a way out? Our analysis offers some answers to this question. First of all, it reveals that overconfidence grows with pre-tax inequality. Second, it highlights that when misperception and reasons for overconfidence resolve, demand for redistribution builds up and reduces post-tax inequality. Although there exists little empirical data for a reality check on the finding of inequality-driven overestimation, the results of the few conducted studies back up such considerations.[32] Survey samples of relatively equal societies exhibit less overestimation of relative income, mobility, or general inequality than samples representing less equal populations. Reducing pre-tax inequality and raising awareness for people's relative income position in society are two measures that can invert the vicious circle. Other surveys and lab studies also attest changes in perception and framing an influence on demand for redistribution (Norton and Ariely, 2011), on voting behavior and acceptance of tax rates (e.g., *fiscal illusion;* Sausgruber and Tyran, 2005, 2011), or even on individual work effort levels (He, 2020; Weber and Schram, 2017).

2.5.3 Limitations and future research

The present study is not without limitations. We thus propose future research to consider mainly two matters. First, our results encourage further exploration of the effect that overconfidence exerts on people's demand for redistribution. While overconfidence represents "perhaps the most robust finding in psychology of judgment" (De Bondt and Thaler, 1995), it remains a complex phenomenon. Our findings confirm the need to separate between stated beliefs, which appear particularly inflated by men, and actual redistributive preferences, where we observe no gender differences in the effort treatments.

Placing the real-effort task after eliciting preferred tax rates in Phase 1 limits task-specific effects on performance expectations. Although this design choice comes with the drawback that tax choices could affect subjects' effort level, respectively stated beliefs, we find no such influence in the present data.

Second, our experiment has implemented a two-period model with exogenous pre-tax inequality and exchange mobility. It would be interesting to consider a longer dynamic horizon in which inequality can change with economic growth and economic mobility interacts with structural mobility. This would test the findings herein in a richer setting and advance the external validity of the analysis.

[32] Cruces et al. (2013) and Karadja et al. (2017) document little overestimation of relative income in their surveys of Argentinian (Swedish) households. Argentina (Sweden) ranks among the least unequal countries in South America (worldwide). Cojocaru (2014) finds greater prospects of upward mobility for EU member states than non-EU, former Socialist economies. Western European countries expose generally greater inequality levels than ex-Soviet states. Kraus and Tan (2015) report considerable mobility overestimation in a US sample. The US is among the least equal societies in the world.

2.5.4 Concluding remarks

This study examines individual preferences for redistribution relative to the level of pre-tax inequality. We divide subjects into two experimental societies. One-half is allotted the scaled-down pre-tax income distribution of South Africa while the other half faces the distribution of Switzerland. Subjects exhibit significantly more optimism about their income position in the high inequality condition (South Africa). The highest income standing out may seduce people to overestimate the small probability of ending up there. Moreover, we find individuals primed with the less equal pre-tax income distribution to accept also more inequality in the post-tax distribution. We construe both findings as practical examples of an inequality trap: unequal societies suffer from greater overestimation of relative income and tolerance of inequality, both leading to reduced demand for redistribution. Tolerance of inequality declines, and the distributional conflict becomes more polarized as people learn their income position. The danger of a society's further destabilization once stuck in an inequality trap should not be dismissed. More empirical and experimental research is needed to confirm these findings and their applicative value for policymakers.

Appendix 2.A Experimental instructions

Translated from Italian. Example: South Africa/effort treatment.

You and the other 19 participants take part in this experiment. The experiment lasts about one hour and consists of two phases. You will receive a 5 € participation fee. In addition, you have the chance of further earnings based on the decisions you make during the experiment. A random mechanism at the end of the experiment determines for which of the two phases in the experiment you are paid immediately after the experiment ends. Your decisions and earnings are kept confidential.

We start with an experimenter reading aloud the instructions for the first phase. The instructions for the second phase will be explained later. At the end of these instructions, you can ask questions before the experiment starts. The experiment concludes with a short questionnaire. During the experiment, it is not permitted to speak or communicate in any form with other participants. Comprehension of the following instructions is critical to maximize your earnings beyond the 5 € for your participation. If you have a question at any time during the experiment, you can raise your hand, and an experimenter will come to your desk and assist you.

Instructions for the first phase

In this phase of the experiment, there are two periods. In period 1, all participants are assigned a gross income from a specific income distribution. This income constitutes the gross income of period 1. In period 2, all participants will receive the gross income of period 2. The gross income in period 2 is determined by a potential increase or decrease of the gross income from period 1 according to an income transition table. This table transforms the gross income of period 1 into the gross income of period 2. Your additional payment in the experiment is the net income you receive in period 2. The net income for period 2 is determined by applying a tax and transfer rate to your gross income of period 2.

Your decision is choosing the tax and transfer rate to be applied to the gross incomes of period 2. This choice will be made before knowing the gross income that was assigned to you in period 1. To determine your period 1 gross-income from the initial income distribution, you and the other participants will compete in a skill task. Your performance relative to the other participants in the skill competition will determine your gross income for period 1. Prior to the skill competition, we will ask you for a self-assessment of your following performance. (*Alternatively in the random-assignment treatment:* In period 1, you will be randomly assigned by the computer to a gross income from the income distribution.) Now we describe the various parts of the experiment.

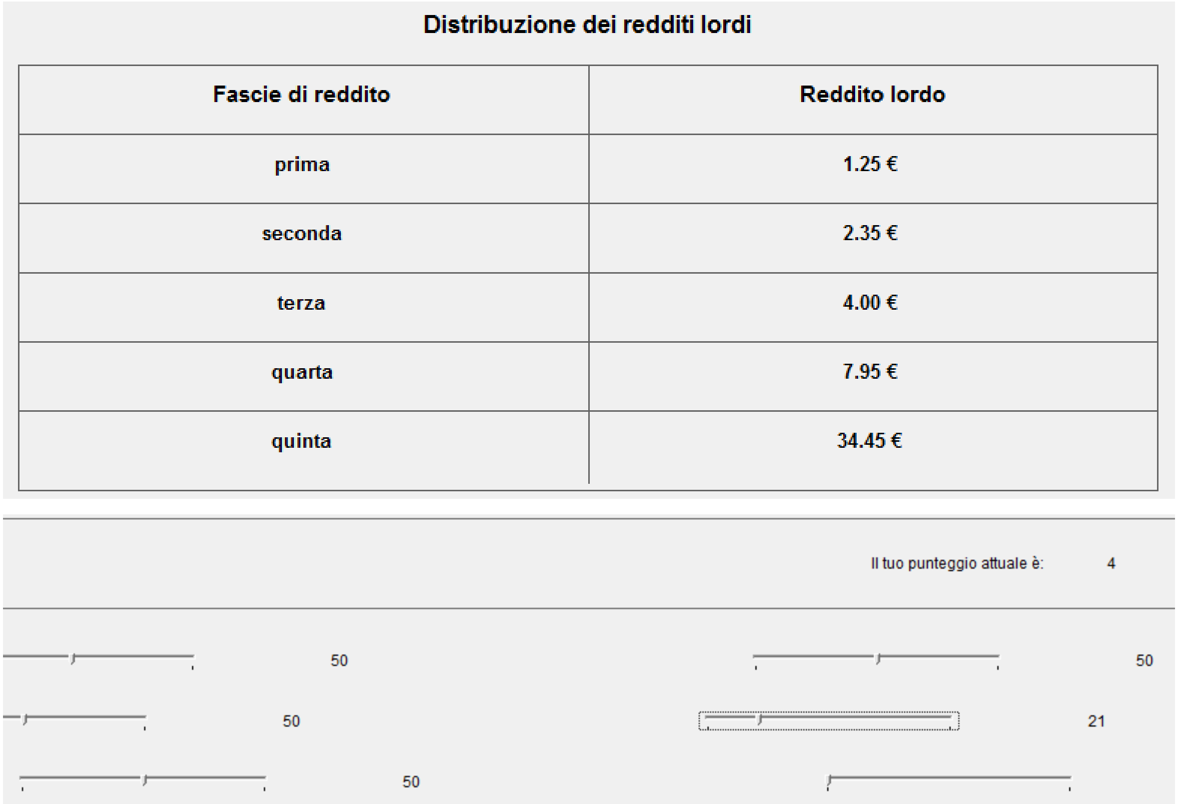
The gross income distribution in period 1

The following table shows the gross income distribution of one of the 40 largest economies in the world, scaled down to an average gross income of € 10.00. This distribution is made up of 5 income classes, containing 4 participants each. Hence, the participants' incomes resemble the gross income distribution of a particular country. The quantities in the right column of the table represent the gross incomes assigned to the various income classes in period 1. According to the table, this means that four participants receive an income of 1.25 €, four an income of 2.35 €, four an income of 4.00 €, four an income of 7.95 €, and the remaining four participants will be assigned an income of 34.45 €.

How is gross income assigned in period 1?

Your income in period 1 will be one of those in the table. As indicated above, your exact gross income in period 1 will be determined based on your performance in relation to the performance of the other participants in a skill task.

In the skill competition, the computer screen will display a series of "sliding



Distribuzione dei redditi lordi

Fascie di reddito	Reddito lordo
prima	1.25 €
seconda	2.35 €
terza	4.00 €
quarta	7.95 €
quinta	34.45 €

bars". Using the mouse and the directional keys of the keyboard, you need to place the cursor, initially sitting on the left end of the bar, on the central value "50" for as many bars as possible. The skill task lasts 120 seconds. The four participants with the highest score (most complete bars), i.e., the four participants on positions 1 to 4, receive an income of 34.45 €. An income of 7.95 € is assigned to the four participants ranking on positions 5 to 8. An income of 4.00 € is awarded to the four participants on positions 9 to 12. An income of 2.35 € is awarded to the four participants on positions 13 to 16. An income of 1.25 € will be awarded to the four participants on positions 17 to 20. In the event of a tie, for example, for 12th and 13th place, the computer will randomly assign each participant to one of the two income classes in question. The skill task takes place at the end of the first phase of the experiment, such that the choices for the tax and transfer rate in period 2 will take place before knowing the income assigned in period 1.

(*Alternatively, in the random-assignment treatment:* Your income in period 1 will be one of those you find in the table. It will be randomly assigned to you by the computer. Specifically, the computer will randomly assign an income of 34.45 € to four participants, an income of 7.95 € to four, an income of 4.00 € to four, an income of 2.35 € to four, and an income of 1.25 € to the remaining four participants. The random assignment will take place at the end of the first phase of the experiment, such that the choices for the tax and transfer rate in period 2 will take place before knowing the income assigned in period 1.)

Income transition table

In period 2, the positions in the gross income distribution are reassigned according to a transition table. The table does not change the gross income distribution, only the positions among participants. The transition table specifies the probability of reaching a certain gross income in period 2 given a certain gross income in period 1. For each cell of the transition table, the row indicates the income class of period 1 and the column the corresponding income class in period 2. There will be two transition tables that can be applied to the gross income distribution in period 1 for two different participant reassignments of income classes in period 2. The figure below exemplifies a screenshot similar to what you will find in the experiment. *(Note, the actual instructions show the tables also without red markings.)*

Tabella A di transizione dei redditi

	a prima fascia (1.25 €)	a seconda fascia (2.35 €)	a terza fascia (4.00 €)	a quarta fascia (7.95 €)	a quinta fascia (34.45 €)
da prima fascia (1.25 €)	75%	25%	0%	0%	0%
da seconda fascia (2.35 €)	25%	50%	25%	0%	0%
da terza fascia (4.00 €)	0%	25%	50%	25%	0%
da quarta fascia (7.95 €)	0%	0%	25%	50%	25%
da quinta fascia (34.45 €)	0%	0%	0%	25%	75%

To make an example for table A, assume a participant with gross income of 2.35 € in period 1 corresponding to the second class. In period 2, her or his gross income will be 2.35 € with 50% probability (see below), 1.25 € with 25% probability, 4.00 € with 25% probability, 7.95 € with 0% probability, and 34.45 € with 0% probability. Instead, to exemplify table B, let us imagine a participant with a period 1 gross income of 34.45 €, in period 2, her gross income will be 1.25 € with probability 0% (see below), 2.35 € with probability 25%, € 4.00 with 25% probability, € 7.95 with 0% probability, and € 34.45 with 50% probability.

Tabella B di transizione dei redditi

	a prima fascia (1.25 €)	a seconda fascia (2.35 €)	a terza fascia (4.00 €)	a quarta fascia (7.95 €)	a quinta fascia (34.45 €)
da prima fascia (1.25 €)	50%	25%	0%	25%	0%
da seconda fascia (2.35 €)	25%	25%	25%	25%	0%
da terza fascia (4.00 €)	25%	0%	25%	25%	25%
da quarta fascia (7.95 €)	0%	25%	25%	25%	25%
da quinta fascia (34.45 €)	0%	25%	25%	0%	50%

Remember that the transition table does not change the gross income of the five classes nor their distribution, but only the reassignment of participants to the five income classes between period 1 and period 2.

Taxes and transfers

We will now describe how taxes and transfers transform the gross income distribution into net incomes. First, each participant's gross income is taxed at a proportional tax rate t. The rate may vary between 0% and 100% in increments of 10%. Second, the collected taxes are redistributed equally among all participants. Each participant will thus receive an identical amount, independent of what she or he personally contributed. (The appendix contains the formula to determine net income based on a given tax rate). Each participant chooses her preferred tax rate t to be applied proportionally to all gross incomes, the proceeds of which will be distributed equally among all participants. To facilitate the choice, a table similar to the following shows all possible tax rates from 0% to 100% in increments of 10%, which, applied to the gross income distribution generates net incomes.

For example, if the tax rate is t = 0%, then no tax is collected. Therefore each participant earns exactly her or his gross income of period 2 (see below). E.g., t = 20% implies that 20% of each participant's gross income in period 2 will be divided among all participants, and each person will receive an equal share of the tax revenues plus the remaining 80% of his or her gross income in period 2 (see below). If t = 100%, each participant will contribute 100% of her period 2 gross income into the tax fund and everyone will then receive an identical net income equal to 10.00 €.

Distribuzioni dei redditi netti per ciascuna aliquota fiscale possibile

L'aliquota fiscale	*t=0%*	*t=10%*	*t=20%*	*t=30%*	*t=40%*	*t=50%*	*t=60%*	*t=70%*	*t=80%*	*t=90%*	*t=100%*
prima fascia	1.25 €	2.13 €	3.00 €	3.88 €	4.75 €	5.63 €	6.50 €	7.38 €	8.25 €	9.13 €	10.00 €
seconda fascia	2.35 €	3.12 €	3.88 €	4.65 €	5.41 €	6.18 €	6.94 €	7.71 €	8.47 €	9.24 €	10.00 €
terza fascia	4.00 €	4.60 €	5.20 €	5.80 €	6.40 €	7.00 €	7.60 €	8.20 €	8.80 €	9.40 €	10.00 €
quarta fascia	7.95 €	8.16 €	8.36 €	8.57 €	8.77 €	8.98 €	9.18 €	9.39 €	9.59 €	9.80 €	10.00 €
quinta fascia	34.45 €	32.01 €	29.56 €	27.12 €	24.67 €	22.23 €	19.78 €	17.34 €	14.89 €	12.45 €	10.00 €

The choice for tax rate t must be made for each of the two transition tables presented in the experiment. At the end, one of the two transition tables will be randomly selected by the computer. The tax rate for that table indicated by a participant randomly drawn by the computer determines the net incomes in period 2, which are paid to the participants in case this phase of the experiment is chosen for payment. Since your stated tax rates for each table can be the one drawn, it is in your own interest to think properly about your choices and report them as accurately as possible. The identity of the drawn participant will not be made public during or after the experiment, nor will you be informed if you were the one selected.

We have reached the end of the instructions for phase 1. To verify your understanding of the transition table please fill in the test table in the next window. (Your answers in the test table do not affect your earnings.) Upon completion, we ask you to answer a few comprehension questions to ensure that you understand the general instructions of phase 1. We invite you to ask for clarification of the instructions in case of any doubts. After all questions are answered, the experiment begins.

Appendix:
The formula to determine the participants' net income given a specific tax rate is the one below. In the formula X_i represents the gross income of participant i in period 2; Y_i is her net income in period 2 (after taxes and transfers), and t the tax rate.

$$Y_i = X_i - tX_i + \frac{1}{20}t\sum_{h=1}^{20} X_h$$

Appendix 2.B Non-parametric tests

This Appendix section reports results from non-parametric tests. In the tables below, we use acronyms for the 16 treatments constructed as follows. The first digit U/C refers to uncertainty (Phase 1)/certainty (Phase 2) of initial positions, the second digit R/E to random/effort income assignment, the third and fourth digits LM/HM to low/high mobility, and the fifth and sixth digits CH/ZA to Switzerland (low inequality)/South Africa (high inequality). E.g., U_R_LM_CH stands for the treatment with uncertain income positions (Phase 1), random income assignment, low mobility, and low inequality (Switzerland).

Table B1 documents evidence from unmatched Mann-Whitney-Wilcoxon tests conducted to check for potential session effects. Recall that we ran two separate sessions with 20 participants for each treatment. The tests show that for all but three treatments, tax choices stated by the participants of the two corresponding sessions in each treatment can be considered as drawn from the same distribution. The three exceptions are: the high mobility condition in Phase 1 (uncertainty of initial positions) of the ZA-random and the CH-effort treatments, and the high mobility condition in Phase 2 (certainty) of the ZA-random treatment. Furthermore, we test for the validity of the regression results reported in the text (Tables 2.6, 2.7, and 2.8) by including alternatively only one of the two sessions of each treatment and confirm that results are robust. We also control for specific effects, e.g., gender, between each session pair but do not find any irregularity.

Table B2 shows the results of one-sample Kolmogorov-Smirnov tests of the null hypothesis that the distributions of participants' choices in each experimental treatment can be considered as uniformly drawn. The test is based on the statistic $D = \sup_x \|F_0(x) - F_{data}(x)\|$ where $F_0(x)$ is the hypothesized distribution and $F_{data}(x)$ the empirical distribution function of the observed data. Since subjects choose tax rates from the discrete value set $\{0, 0.1, ..., 1\}$, we use the statistic D for discrete uni-

Table B1: Session effects, unmatched Mann-Whitney-Wilcoxon rank-sum tests.

Uncertainty			Certainty		
Treatment	*z*	*p*	*Treatment*	*z*	*p*
U_R_LM_CH	-1.469	.142	C_R_LM_CH	-0.509	.611
U_R_HM_CH	-1.590	.112	C_R_HM_CH	-0.478	.633
U_R_HM_ZA	1.150	.250	C_R_LM_ZA	0.999	.318
U_R_HM_ZA	1.715	.086*	C_R_HM_ZA	2.143	.032**
U_E_LM_CH	0.136	.891	C_E_LM_CH	1.288	.227
U_E_HM_CH	2.081	.037**	C_E_HM_CH	0.873	.383
U_E_LM_ZA	0.438	.661	C_E_LM_ZA	0.029	.977
U_E_HM_ZA	0.861	.390	C_E_HM_ZA	-0.195	.845

* $p < .1$, ** $p < .05$, *** $p < .01$.

Table B2: One-sample Kolmogorov-Smirnov test for discrete distributions.

Uncertainty			Certainty		
Treatment	*D*	*p*	*Treatment*	*D*	*p*
U_R_LM_CH	.182	.142	C_R_LM_CH	.243	.017**
U_R_HM_CH	.205	.070*	C_R_HM_CH	.143	.385
U_R_HM_ZA	.139	.425	C_R_LM_ZA	.259	.009***
U_R_HM_ZA	.189	.116	C_R_HM_ZA	.255	.011***
U_E_LM_CH	.179	.152	C_E_LM_CH	.159	.263
U_E_HM_CH	.129	.513	C_E_HM_CH	.114	.679
U_E_LM_ZA	.230	.030**	C_E_LM_ZA	.409	.000***
U_E_HM_ZA	.238	.021**	C_E_HM_ZA	.342	.000***

* $p < .1$, ** $p < .05$, *** $p < .01$.

form distributions (Arnold and Emerson, 2011).[33] As discussed in the text, the tests indicate that empirical distributions in Phase 1 are mostly consistent with the uniform distribution. Instead, very high p values reject the null distribution in all ZA treatments of Phase 2.

Table B3 reports matched Wilcoxon signed-rank tests for the within-subject treatments, that is uncertainty (Phase 1) versus certainty (Phase 2) of initial positions and Low vs. High mobility. Table B4 reports unmatched Mann-Whitney-Wilcoxon rank-sum tests for the between-subject treatments, i.e., CH vs. ZA and random vs. effort-based income assignment. (In the tables, the star $*$ in the digit position of the acronyms indicates the testing condition: e.g., $*$ in the first position refers to the test of uncertainty vs. certainty). The tests confirm the main visual conclusions discussed in Section 2.4. Regarding the positive effect in the ZA effort treatments

Table B3: Matched Wilcoxon signed-rank tests.

Uncertainty vs. Certainty			Low vs. High mobility		
Treatment	*z*	*p*	*Treatment*	*z*	*p*
*_R_LM_CH	-0.401	.688	U_R_*_CH	0.839	.401
*_R_HM_CH	-1.245	.213	U_R_*_ZA	-1.425	.154
*_R_LM_ZA	-2.415	.015**	U_E_*_CH	-0.251	.802
*_R_HM_ZA	-3.638	.003***	U_E_*_ZA	-2.910	.004***
*_E_LM_CH	-0.122	.906	C_R_*_CH	-0.271	.787
*_E_HM_CH	-1.070	.284	C_R_*_ZA	-0.512	.608
*_R_LM_ZA	-3.162	.002***	C_E_*_CH	-0.744	.457
*_R_LM_ZA	-3.244	.001***	C_E_*_ZA	-1.130	.258

* $p < .1$, ** $p < .05$, *** $p < .01$.

Table B4: Unmatched Mann-Whitney-Wilcoxon rank-sum tests

Switzerland vs. South Africa			Random vs. Effort		
Treatment	*z*	*p*	*Treatment*	*z*	*p*
U_R_LM_*	0.763	.445	U_*_LM_CH	1.235	.217
U_R_HM_*	-1.053	.292	U_*_HM_CH	0.325	.745
U_E_LM_*	0.044	.965	U_*_LM_ZA	0.763	.446
U_E_HM_*	-1.047	.295	U_*_HM_ZA	0.141	.888
C_R_LM_*	-0.893	.372	C_*_LM_CH	1.064	.287
C_R_HM_*	-1.420	.156	C_*_HM_CH	0.698	.485
C_E_LM_*	-2.191	.029**	C_*_LM_ZA	-0.573	.567
C_E_HM_*	-2.711	.006***	C_*_HM_ZA	-0.839	.401

* $p < .1$, ** $p < .05$, *** $p < .01$.

[33] In the most standard Kolmogorov-Smirnov test, the distribution $F_0(x)$ is continuous, so that the distribution of D does not depend on the hypothesized distribution. In case of discrete null distributions, the distribution of D depends on the null model and requires considerably more effort to obtain. Nevertheless, there exist methodologies to compute the statistics in the discrete cases, too. The tests reported here are based on the methodology proposed by Arnold and Emerson (2011).

of Phase 1 with respect to the high vs. low mobility control, it is also worthwhile pointing out that whereas the ZA effort condition of Phase 1 is the only treatment exhibiting a significantly negative effect, all but one of the individual tests report negative signs. This is contrary to the POUM hypothesis. Note that the Wilcoxon test conducted on the whole treatments provide overall evidence against the POUM ($z = -2.2$, $p = .025$).

Tables B5 and B6 report matched and unmatched tests between treatment and control group of each treatment condition. Besides confirming the results in the main text, they provide further qualifications, in particular emphasizing the stronger impact on the sub treatments by the high inequality (ZA) manipulation compared to the low inequality (CH) one. The tests show again that the difference between the two inequality conditions is more articulated under certainty (Phase 2) than uncertainty (Phase 1). This discrepancy grows stronger in the effort than in the random treatments and rather under certainty than uncertainty of initial positions. This is supportive of the results derived by the regression analyses in the paper and summarized in Result 1 (Section 2.4.2) and Result 2 (Section 2.4.3).

Table B5: Matched Wilcoxon signed-rank tests (aggregate).

Uncertainty vs. Certainty			Low vs. High mobility		
Treatment	*z*	*p*	*Treatment*	*z*	*p*
*_R_	-3.547	.000***	U_*_	-1.898	.057*
*_E_	-3.769	.000***	C_*_	-1.233	.217
*_	-5.159	.000***	_*_	-2.238	.025**
*_CH	-1.270	.204	_R_*	-0.596	.551
*_ZA	-6.004	.000***	_E_*	-2.505	.012**
*_LM	-3.154	.002***	_*_CH	-0.223	.823
*_HM	-4.425	.000***	_*_ZA	-3.179	.015**
*_R_LM	-1.955	.051*	U_R_*	-0.321	.748
*_E_LM	-2.494	.012**	U_E_*	-2.238	.025**
*_R_HM	-3.280	.010***	C_R_*	-0.552	.581
*_E_HM	-3.007	.003***	C_E_*	-1.190	.234
*_R_CH	-1.096	.273	U_*_CH	0.427	.669
*_E_CH	-0.699	.485	U_*_ZA	-3.206	.001***
*_R_ZA	-4.103	.000***	C_*_CH	-0.685	.493
*_E_ZA	-4.496	.000***	C_*_ZA	-1.156	.248
*_LM_CH	-0.400	.689	_R_*_CH	-0.449	.653
*_HM_CH	-1.616	.106	_R_*_ZA	-1.363	.173
*_LM_ZA	-3.925	.000***	_E_*_CH	-0.792	.423
*_HM_ZA	-4.740	.000***	_E_*_ZA	-3.033	.002***

* $p < .1$, ** $p < .05$, *** $p < .01$.

Table B6: Unmatched Mann-Whitney-Wilcoxon rank-sum tests (aggregate)

Switzerland vs. South Africa			Random vs. Effort		
Treatment	*z*	*p*	*Treatment*	*z*	*p*
U_*_	-0.711	.477	U_*	1.121	.263
C_*_	-3.502	.001***	C_*	0.285	.774
_*_	-2.961	.003***	_*_	0.967	.334
R_*_	-1.221	.222	LM_*_	1.243	.214
E_*_	-2.928	.003***	HM_*_	0.100	.921
LM_*	-1.337	.181	_*_CH	1.676	.094*
HM_*	-3.021	.003***	_*_ZA	-0.373	.709
U_R_*	-0.212	.832	U_LM_*	1.486	.137
U_E_*	-0.763	.445	U_HM_*	0.146	.884
C_R_*	-1.580	.114	C_LM_*	0.348	.728
C_E_*	-3.368	.001***	C_HM_*	0.010	.992
U_LM_*	0.517	.605	U_*_CH	1.107	.268
U_HM_*	-1.510	.131	U_*_ZA	0.550	.582
C_LM_*	-2.207	.027**	C_*_CH	1.254	.210
C_HM_*	-2.907	.004***	C_*_ZA	-1.010	.313
_R_LM_*	-0.162	.871	LM_*_CH	1.653	.098*
_R_HM_*	-1.666	.096*	LM_*_ZA	0.060	.952
_E_LM_*	-1.675	.094*	HM_*_CH	0.677	.498
_E_HM_*	-2.575	.009***	LM_*_ZA	-0.524	.600

* $p < .1$, ** $p < .05$, *** $p < .01$.

Appendix 2.C A model of structural preferences

While the identification of a single theory able to explain subjects' behavior is not the purpose of this work, it is nonetheless worthwhile investigating the extent to which a sufficiently general model can fit the data. With a similar purpose, Durante et al. (2014) adapt the social preferences model by Charness and Rabin (2002) to test their data. Doing the same we compare our estimates to those of Durante et al. (2014).

Charness and Rabin (2002) propose a simple theory of social preferences, valid for both strategic games and multiperson situations, in which people's preferences are expressed by a convex combination of selfish and social motivations. In the original model the selfish motivation represents a person's own expected payoff, while the disinterested social component is formed by a convex combination of concerns for equity, expressed in the Rawlsian form of caring for the worst-off person in the society, and efficiency. Remember, redistributive taxation occurs at no cost in our experiment such that efficiency considerations can be excluded. Durante et al. (2014) extend the model to include the individual risk attitude in the personal motivation. The resulting utility function for subject h in our experiment can be written as:

$$V_h = (1-\lambda)\left[(1-\gamma)\,E[Y] + \gamma(-\sigma_Y)\right] + \lambda Y^{min}, \tag{2.2}$$

where $E[Y]$ and σ_Y denote the expectation of own (post-tax) income, respectively its standard deviation. The term Y^{min} expresses the Rawlsian equity concern for the person with the lowest post-tax income.[34] According to the model, the structural preference parameters of interest in our experiment are the utility weights λ and γ.

The model encompasses as special cases the three main criteria of selecting tax rates discussed in the text (see Table 2.4), namely: I) maximization of expected earnings, occurring when λ and γ equal both zero; II) risk aversion, in the form of a mean-variance utility, when $\lambda = 0$ and $\gamma \in (0,1)$; and III) inequity aversion (with no risk aversion), when $\lambda \in (0,1)$ and $\gamma = 0$.

To derive the theoretical predictions more formally, let the three terms in the utility function, $E[Y]$, σ_Y, Y^{min}, be expressed as functions of tax rate τ, that is:

$$E[Y] = \sum_{j=1}^{5} \pi_{hj} X_j (1-\tau) + \tau \overline{X}, \tag{2.3}$$

$$\sigma_Y = \sqrt{\sum_{j=1}^{5} \pi_{hj} \left(X_j(1-\tau)\right)^2 - \left[\sum_{j=1}^{5} \pi_{hj} X_j (1-\tau)\right]^2}, \tag{2.4}$$

$$Y^{min} = X^{min}(1-\tau) + \tau \overline{X}, \tag{2.5}$$

where π_{hj} is the subjective probability of subject h being in income quintile j before taxation, X_j is the pre-tax income in quintile j, and $\overline{X}$ and X^{min} the society's mean and minimum pre-tax income. Clearly, while $\overline{X} = 10$ in all treatments, π_{hj}, X_j and X^{min} depend on treatments and subjective probabilities, as discussed in the paper.

Using the three expressions in the utility function (2.2) and taking the derivative with respect to τ gives:

$$\begin{aligned} \frac{\partial V_h}{\partial \tau} &= (1-\lambda)(1-\gamma)\left[\overline{X} - \sum_{j=1}^{5} \pi_{hj} X_j\right] + \\ &\quad \gamma(1-\lambda)\sqrt{\sum_{j=1}^{5} \pi_{hj} X_j^2 - \left[\sum_{j=1}^{5} \pi_{hj} X_j\right]^2} + \lambda\left[\overline{X} - X^{min}\right] \end{aligned}$$

which delivers the predictions discussed in section 2.3.3 given at least one of the two structural parameters λ and γ equals 0. Yet, if λ and γ are inside the interval $(0,1)$, all three motivations play a role in subjects' preferences. Nevertheless, note that

[34] Remark the extra term σ_Y in the utility function concerning the specification used for Table 2.4 in the main text representing the extension proposed by Durante et al. (2014) to Charness and Rabin (2002) original specification.

since utility function (2.2) is linear in the tax rate, the model always predicts corner solutions at $\tau = 0$ or 1, depending on the values of π_{hj}'s, X_j's, etc. in treatments and on the utility parameters λ and γ. In this sense, the model can be considered as benchmark limiting case.

Estimation of the structural parameters λ and γ can be obtained following the method of Durante et al. (2014), which is based on the conditional logit model of McFadden (1973). The method requires constructing an observation for each subject h in any treatment for each possible tax rate $\tau \in K = \{0, 0.1, ..., 1\}$. It maximizes utility function (2.2) making random errors, which, under conditions of "type I extreme value" distribution, imply that the probability for subject h choosing a certain tax rate $\tau_h = \tau$, with $\tau \in K$, is given by:

$$P(\tau_h = \tau) = \frac{e^{V_{h\tau}}}{\sum_{k \in K} e^{V_{hk}}}.$$

The resulting likelihood function is then maximized to estimate the parameters β_1, β_2 and β_3 the in utility function (2.2) written as:

$$V_h = \beta_1 EY + \beta_2 \sigma_Y + \beta_3 Y^{min}$$

The estimates of the $\beta_l's$ are finally used to obtain the structural preference parameters of function (2.2) according to the transformations:

$$\lambda = \frac{\beta_3}{\beta_1 + \beta_2 + \beta_3}; \qquad \gamma = \frac{\beta_2}{\beta_1 + \beta_2}$$

The estimates of the model obtained in our experiment are reported in Table D1. They exhibit great variability between treatments of Phase 1 (uncertainty of initial positions) and Phase 2 (certainty). Model (1) pools the data of all experimental treatments. All coefficients show the expected sign: those on personal variables, namely own expected income and standard deviation, are highly statistically significant; the coefficient on the society's minimum income is mildly significant. Models (2) and (3) investigate treatments of Phase 1: Model (2) regards all treatments, while Model (3) alone the effort treatments. The random treatments of Phase 1 are not reported because the maximization algorithm does not converge when the data of the random treatments are estimated alone. Nevertheless, results of Models (2) and (3) show that the utility specification in equation (2.2) is unsuitable for explaining subjects' behavior in Phase 1: note, in particular, parameter β_3 on the society's minimum income exposes a negative sign, contrary to the model predictions. As a consequence, also the utility weights λ and γ of the utility function are outside the predicted domain $[0, 1]$. Model (4) pools the data for the treatments of Phase 2. All coefficients are highly statistically significant and with the expected signs. Models (5) and (6) separate between the random and the effort treatments of Phase 2. Concerning the former treatments, convergence requires not using the data from the CH treatment with low mobility (possibly due to little variation in the data, particularly standard deviation of own incomes). The separate regressions (5) and (6) confirm the predicted signs of the coefficients, even if their significance is generally weaker than in the pooled model (4) because of fewer observations.

Table D1: Estimation of Charness and Rabin (2002) - Durante et al. (2014) utility

	(1) Phase 1 & Phase 2	(2) Phase 1	(3) Phase 1 Effort	(4) Phase 2	(5) Phase 2 Random	(6) Phase 2 Effort
Expected pers. income (β_1)	0.326***	0.194***	0.244***	0.423***	0.261***	0.610***
	(0.088)	(0.057)	(0.063)	(0.084)	(0.083)	(0.041)
St. dev. of pers. income (β_2)	-0.082***	-0.265***	-0.345***	-0.190***	-0.074***	-0.298***
	(0.240)	(0.010)	(0.014)	(0.050)	(0.099)	(0.489)
Minimum soc. income (β_3)	0.104*	-0.434***	-0.393***	0.131**	0.137*	0.085
	(0.054)	(0.083)	(0.016)	(0.061)	(0.071)	(0.118)
λ	0.203***	-17.461	-1.974***	0.176*	0.290	0.086
	(0.060)	(109.260)	(0.732)	(0.099)	(0.219)	(0.108)
γ	0.201***	0.577***	0.584***	0.309***	0.221	0.329***
	(0.023)	(0.063)	(0.072)	(0.014)	(0.176)	(0.038)
(1-λ)	0.797***	18.461	2.974***	0.824***	0.710***	0.914***
	(0.060)	(109.260)	(0.732)	(0.099)	(0.219)	(0.108)
Cases	640	320	160	280	120	160
Log-pseudolikelihood	-1408.897	-719.404	-361.323	-563.653	-249.211	-302.255

Note: Standard errors in parentheses. * $p < .1$, ** $p < .05$, *** $p < .01$.

Overall, the estimated models in Table D1 indicate that subjects follow very different behavioral rules in Phase 1 and Phase 2 of the experiment. Especially tax choices in Phase 1 depart substantially from models that include social concerns. This is consistent with the evidence summarized in Result 1 (Section 2.4.2). On the other hand, choices in Phase 2 are overall more in line with the utility model (2.2). The results from model (4), possibly the most reliable due to the largest degree of freedom, indicate that subjects place a relative weight on personal motivations about 4.7 times higher than the one placed on social concern (i.e., $\frac{1-\lambda}{\lambda} = \frac{0.824}{0.176} = 4.68$). Concern for own payoff is affected by risk aversion (ratio of about $\frac{1}{2.4}$ as measured in terms of a negative concern for the standard deviation of own payoff relative to the expected value, i.e., $\frac{\gamma}{1-\gamma} = \frac{0.309}{0.691}$). We find similarities to Durante et al. (2014). For example, Table 3, column (2) in Durante et al. (2014), reporting results for the treatments most similar to ours, finds estimates of $\lambda = 0.134$ and $\gamma = 0.110$, which are not too far from ours. Nevertheless, the whole set of regressions reported in their Table 3 also confirm that the structural weights in equation (2.2) vary substantially between contexts, pointing out that subjects' tax choices depend on considerations which go beyond those underlying the utility specification.

Appendix 2.D Additional tables

Table D2: Tobit regressions - Overconfidence about income quintile in effort treatments

	Phase 1 and Phase 2				Phase 1			
	(1) Overcon-fidence	(2) OC Female	(3) Critical OC	(4) Crit. OC Female	(5) Overcon-fidence	(6) OC Female	(7) Critical OC	(8) Crit. OC Female
Dependent variable	τ_h	τ_h	τ_h	τ_h	τ_h	τ_h	τ_h	τ_h
Income uncertainty	-0.167***	-0.171***	-0.154***	-0.181***				
	(0.054)	(0.054)	(0.055)	(0.053)				
High inequality (ZA)					0.024	0.028	0.016	0.075
					(0.066)	(0.065)	(0.063)	(0.061)
Low mobility					-0.080**	-0.080**	-0.080**	-0.081**
					(0.036)	(0.036)	(0.036)	(0.036)
Female	0.140*	0.135*	0.139*	0.123	0.063	0.030	0.064	-0.052
	(0.078)	(0.076)	(0.078)	(0.080)	(0.064)	(0.064)	(0.064)	(0.067)
OC (expected - actual quintile)					-0.004	-0.045**		
					(0.016)	(0.020)		
OC × Female						0.075**		
						(0.030)		
Critical OC							-0.068	-0.264***
							(0.076)	(0.075)
Critical OC × Female								0.534***
								(0.139)
Phase 1 × OC	-0.002							
	(0.016)							
Phase 1 × OC × Female		0.031						
		(0.024)						
Phase 1 × critical OC			-0.072					
			(0.083)					
Phase 1 × critical OC × Fem.				0.193				
				(0.132)				
Constant	0.551***	0.553***	0.551***	0.558***	0.446***	0.466***	0.461***	0.480***
	(0.072)	(0.072)	(0.072)	(0.072)	(0.056)	(0.056)	(0.060)	(0.060)
Observations	320	320	320	320	160	160	160	160
Log-pseudolikelihood	-243.383	-242.871	-243.072	-242.483	-66.068	-63.287	-65.535	-57.323
Pseudo R-squared	0.034	0.036	0.035	0.038	0.033	0.074	0.041	0.161

* p < .1, ** p < .05, *** p < .01.

Chapter 3

Pushed to perform - time pressure in long-run asset pricing experiments

3.1 Introduction

Dating back to the 17th century, bursting speculative bubbles have shattered economies continually. In 1637, the Dutch Republic - the then-largest economy in the world - caught the "tulip fever", which led to one of the first-ever documented financial crashes (Garber, 1989). Within no time, demand for the newly introduced tulip bulbs heated up to price levels worth annual wages. Centuries have since passed, but speculative bubbles have remained and characterize markets to this day, with the unprecedented stock market crash in response to the COVID-19 pandemic being its freshest sign of life.[35] The economic literature has evolved around describing and understanding market behavior since its very beginning. Yet, not before the end of the previous century, the economic society opened up to methodologies from other social sciences and started to exploit the potential of experimental work. The present chapter extends the strand of experimental literature, which studies the underlying forces favoring market volatility.

We run a so-called Learning-to-Forecast Experiment (LtFE) to investigate the emergence and robustness of bubbles in asset prices. Our work is innovative in the following aspects. First, we contribute to the growing literature on time pressure in decision-making. While related work, reviewed in Section 3.2.5, centers mostly on decision time in individual decision-making, we study time pressure in a market experiment with multiple players putting special focus on expectation formation.

[35]The Dow Jones Industrial Average suffered its three largest one-day point drops in history during the global outbreak of the pandemic in March 2020 (Statista, 2020).

Next, we are the first to examine long-term dynamics (beyond 65 periods) in LtFEs. As a consequence, this enables us, third, to distinguish elicited forecasting strategies of participants between early and mature market stages.

The contributions to the literature derive along those three dimensions. First, we take note that markets in the baseline condition with moderate time pressure become less volatile in the long run and often converge to the underlying price (Fig. 3.1, upper left panel) after typically starting with large fluctuations commonly found in LtFE studies (see Section 3.2.1). We find no evidence, though, that similar price patterns throughout the first 50 periods necessarily lead to similar dynamics in the long run. This means that previous and on-going LtFE research based on around merely 50 periods (or 65 periods in Bao et al., 2012) may suffer from overemphasizing this transitory phenomenon.

Second, price dynamics depend starkly on the level of time pressure participants

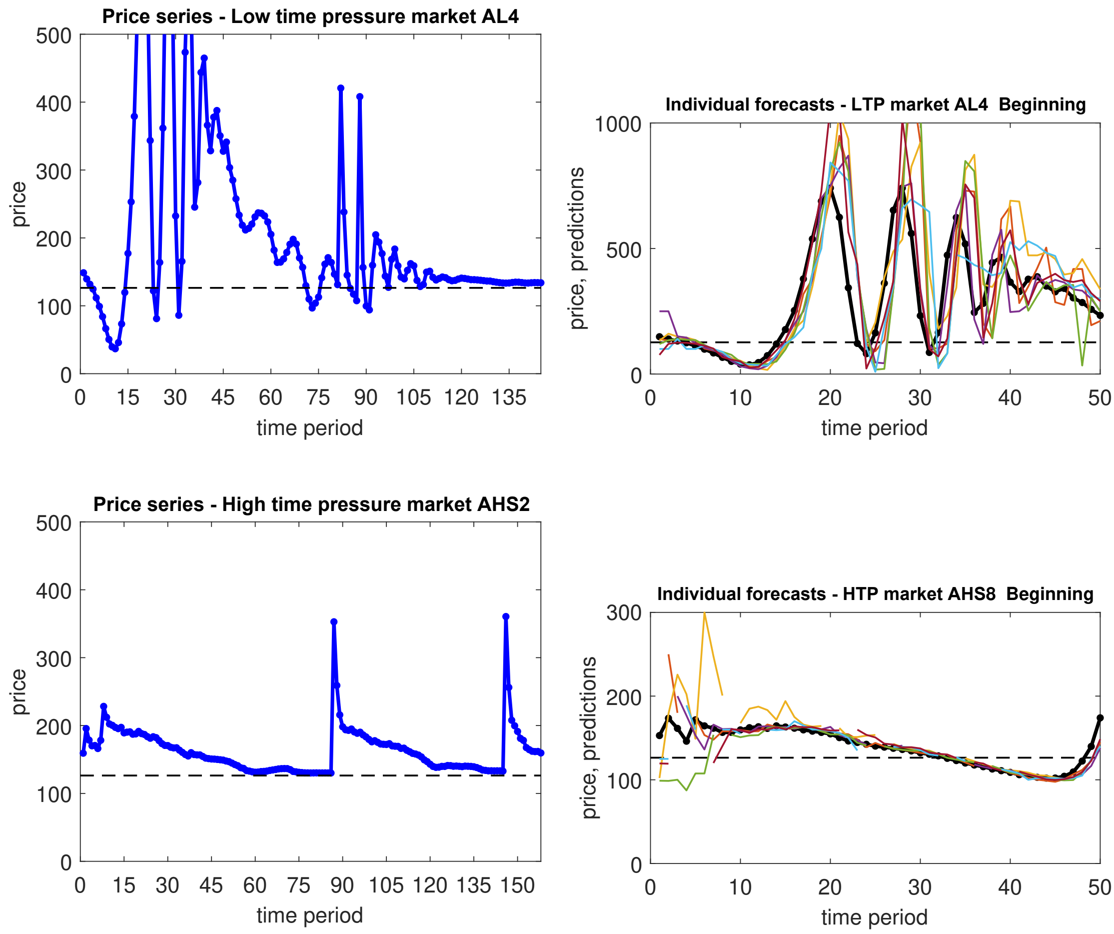


Figure 3.1: *Left panels.* Effect of time pressure on price dynamics.
Right panels. Treatment difference in coordination between individual forecasts.

experience. Specifically, markets are *less* prone to bubble and crash patterns under increased time pressure (compare left panels in Figure 3.1). Underlying these findings, we encounter, third, forecasting strategies differ between low and high time pressure conditions and can even be subject to change over time. Reduced time pressure seems to foster more complex forecasting rules such as trend-following and rapid coordination between subjects (Fig. 3.1, upper right panel), while the share of simple rules occurs to grow with time. Naïve strategies, for instance, perform better in this task. These behavioral changes can explain large volatility in the beginning and its decline toward the end (Fig. 3.1, upper left panel). Enhanced time pressure appears to nudge individuals immediately toward utilizing simple strategies. In line with the cognitive dual-process framework distinguishing decision making into intuition driven or effortless *fast thinking* (Type 1) versus effortful *slow thinking* (Type 2), time pressure may restrict cognitive effort from engaging in more complex strategies (Kahneman, 2011; Li et al., 2017a). With less decision time, individual forecasts start coordinating later in the market (Fig. 3.1, lower right panel). The observed forecasting behavior under increased time pressure produces stable markets and induces rapid convergence to the fundamental value. However, two different paths emerge in the long run. Prices either remain stable until the end of the experiment or leave the equilibrium state and become volatile. The data provides suggestive evidence that market volatility can be influenced by manipulation of decision time.

The remainder of this paper is structured as follows. In Section 2, we review previous market experiments and related literature and further formulate the hypotheses for this study. Section 3 lays out the experimental procedure and design while Section 4 describes and analyzes the data. The paper concludes with a discussion of the findings and an outlook for future research.

3.2 Related literature

The experimental literature maintains a long tradition of asset pricing experiments starting with the seminal work by Smith et al. (1988). Since then, experimentalists have set out to simulate asset markets in the laboratory. Still, most literature has dismissed two decisive characteristics of asset markets: *time pressure* and *long runtime*.

First, trading free of time constraints is likely the exception to the rule on financial markets regarding the hectic hustle and bustle on physical trading floors around the globe. While these images portray the stereotypical human response from time pressure, the nature of decision-making considering abundant information in a fast-changing environment creates an ideal playing field for algorithm-based, computerized trading. What is known as high-frequency-trading has found growing relevance on financial markets in terms of transaction volume and order-to-trade ratio (Aldrich and Vargas, 2020). Yet, it remains an open question how forecasting behavior in the lab changes under time pressure. Controlling market parameters experimentally further aims at better understanding their influence and functions as test-bed to inform regulators on the effect of potential interventions.

Second, albeit various scholars identified common patterns, such as *hog cycles* (Haas et al., 1925; Hanau, 1927), *Kuznets swings* (Kuznets, 1930), or *Kondratiev waves* (Kondratieff, 1979), economists around the globe have struggled repeatedly to foresee market crashes. Queen Elizabeth II herself wondered about the financial crisis in 2008 "why did no one see it coming" (Skidelsky, 2016). This can be understood as a royal request for feeding learning models with extensive experimental time series. Analysis of short time series is prone to overlook the bigger picture and thus limited in developing early warning systems. The scope of most studies rarely exceeds 50 periods. Although recent work (Friedman et al., 2015) suggests that behavior in interactive environments changes substantially with the number of feedback periods. Alongside the afore-mentioned double auction (Smith et al., 1988) and the closely related call market (e.g., Kopányi-Peuker and Weber, 2019), a third strand of market experiments has paved its way onto the landscape of the asset pricing literature, so-called Learning-to-Forecast experiments.

3.2.1 Learning-to-Forecast experiments

In LtFEs, subjects usually form markets of six participants and predict the price of an asset repeatedly over several periods (see Hommes et al., 2005). The forecast accuracy in each period determines earnings, while the asset price in any given period is a function of the average predictions. Markets can be characterized either by positive or negative feedback. Positive feedback implies that markets expecting increasing (decreasing) prices produce de facto increasing (decreasing) prices in a self-fulfilling manner (also known as demand-driven markets or strategic complements). This applies to asset-trading of all kinds, whether it is stocks or real estate. Instead, negative feedback suggests that expectations for rising (falling) prices result in falling (rising) prices such as in crop markets (supply-driven markets or strategic substitutes). Heemeijer et al. (2009) compare the dynamics of negative to positive feedback markets and conclude that the former are much more stable. They argue that positive feedback is prone to rapid coordination of trend-extrapolating strategies among forecasters evolving into bubble-and-crash dynamics, while lacking coordination between subjects' forecasts, representative for negative feedback markets, stabilizes prices.

Essentially, LtFEs can be modeled as a continuous *Keynesian Beauty-contest* game (BCG, i.e., *guessing game*, see Mauersberger and Nagel, 2018; Sonnemans and Tuinstra, 2010) in which the subject whose guess is closest to the average over all guesses, multiplied by a factor, wins.

3.2.2 Number of decision periods in LtFEs

LtFE studies usually assess markets over 50 periods (e.g., Bao et al., 2019; Heemeijer et al., 2009; Hommes et al., 2005). As one of the few exceptions, Bao et al. (2012) extend the standard LtFE model for another 15 decision periods, while in Evans et al. (2019), subjects continuously predict the asset price for the subsequent ten periods. In all these studies, markets typically follow four different price patterns: a) constant fluctuations around the fundamental value, b) amplifying, c) decreasing oscillations over time, or, d) convergence, from above or below, to the fundamental value. By manipulating subjects' experience in a first block of 36 periods (bubbly or stable prices), Hennequin (2018) shows that the nature of initial dynamics determines market volatility thereafter. Following a multi-block design, too, yet in constant subject configuration, Kopányi-Peuker and Weber (2019) run three rounds with each about 30 decision periods. They find no learning effects but boom-and-bust cycles across all rounds, even if full information about the fundamental value (FV) is provided. Nonetheless, it remains unclear whether the described patterns persist in a market beyond 50 periods or whether other phenomena emerge in the long run.

3.2.3 Experimental evidence on long-run decision-making

There are very few studies on asset pricing that substantially exceed the standard horizon (15 decision periods) set by Smith et al. (1988). Investigating double auctions and call markets for 100 periods, Hoshihata et al. (2017) report "flat bubbles" in both settings. Akin to what is generally observed across 15 periods, assets also trade throughout 100 periods significantly above the FV. It is not until close to the final periods that bubbles burst for one or another reason. Kopányi-Peuker and Weber's (2019) LtFE is part of a comparison study with a call market. Running call markets likewise over three rounds with each approx. 30 periods, the authors note flat bubbles and no more accurate pricing in later rounds, thus no experience effects even if full information about the FV is provided. Measuring convergence to the FV, Lahav (2011) finds that within-market experience being gained in a long-run market over 200 consecutive periods, is (even) less effective than between-market experience gained from multiple sessions.

Friedman et al. (2015) add intriguing evidence to the debate on whether subjects are able to adopt superior strategies over lengthy market experiments. They exploit a Cournot duo- and triopoly in which subjects (grouped in markets of two, respectively three) decide upon their output each round.[36] Subjects setting a higher output, without considering its effect on the price, will make higher profits than the other(s) in the same market, as long as prices exceed marginal costs. Consequently, market output approaches the *Walrasian equilibrium* during the first 25 rounds - all subjects submit maximum output and make zero profits - as documented in previous literature. What is new, markets continue for another 1175 rounds. Shortly after round 25, curiously enough marking the terminal round in earlier studies, subjects realize that they are following a sub-optimal strategy. They switch strategies and as if by magic, align on non-equilibrium collusion, which maximizes joint profits.

Following the insights on strategic learning by Friedman et al. (2015), we conjecture that substantially extending the market length enables LtFE markets to discover the fundamental asset value. Incentivized subjects should understand that stable price dynamics make forecasting easier and facilitate higher earnings.

Hypothesis 1: *Markets converge to their fundamental value in the long run.*

3.2.4 Decision time in LtFEs

Similar to the question of market length, implications from time pressure have not yet been studied in LtFEs. Most designs use, if any, soft timers of two minutes after which subjects are prompted to submit a prediction, while a few employ hard timers of 120 (60) seconds for the first ten periods and 60 (45) seconds thereafter (Hommes et al., 2020; Kopányi et al., 2019; Kopányi-Peuker and Weber, 2019). Reported average decision times in previous LtFEs equal approximately 30 seconds, irrespective of the number of periods passed (Hommes et al., 2020; Kopányi-Peuker and Weber, 2019). Kocher and Sutter (2006) study repeated BCGs over 24 periods and find that increasing time pressure (15 vs. 120 seconds per period) leads to slower convergence

[36]For the unfamiliar reader, note that the total output of all subjects determines the price of the homogeneous good in Cournot markets.

to the Nash Equilibrium, equal to the FV, and consequently lower earnings.[37] Agranov et al. (2015) introduce a *choice process* protocol to BCGs. Subjects play one-shot games in which they can change their guesses indefinitely often within 120 seconds. They find that tracked guesses approach the steady state (FV) as time passes.

3.2.5 Decision-making under time pressure

Behavioral implications of time pressure have been studied in psychology already for a long time. The common narrative tells that a time constraint complicates non-interactive decision making. Various scholars advocate a dual-process framework for an intelligible interpretation of handling cognitive information. System 1 (Kahneman, 2011) or the "Doer" (Thaler, 2018) represents intuitive or emotional thinking, while System 2 or the "Planner" endorses the idea of more elaborated deliberation suppressing or overwriting System 1 if necessary. More recently, neural research postulates a related alternative model. Li et al. (2017a) argue that decision-making is best characterized by low (Type 1) versus high cognitive engagement (Type 2). Depending on the framing, the decision problem is processed in two distinct areas of the brain. They detect a framing effect on response time in their data.

To our knowledge, Moritz et al. (2014) is the only experimental work that investigates financial forecasting under time pressure. In contrast to the present study, subjects forecast predefined time series of demand data over 50 periods. The authors measure subjects' decision times and forecasting errors. Subjects taking little time to reach a decision tend to perform worse. This goes in hand with the concept of System 1 or Type 1 and is considered as under-thinking. Interestingly, also those subjects taking considerably longer to submit predictions suffer from underperforming. Such overthinking may lead to sub-optimal strategies in specific tasks. Considering large numbers of decision attributes can result in cognitive overload and ineffective weighting of these attributes. In a second step, Moritz et al. (2014) enforce under-thinking by imposing a maximum decision time (∞/7/15 secs) and overthinking by a minimum decision time (15/0/7 secs).[38] Successful replication of prior performance levels in this manipulation treatment confirms their initial findings. Reportedly, the condition combining a maximum decision time of 15 seconds with a minimum of 7 seconds produces the smallest forecasting errors. Moreover, subjects with higher cognitive ability, measured by a cognitive reflection test (CRT,

[37] In contrast to LtFEs, in which the FV is inside the value range (usually 60), the equilibrium state in this study equals the zero lower bound.

[38] Although time pressure is not the object in Friedman et al. (2015), we note for referential purpose that their design limits subjects to a decision time of four seconds per period.

Frederick, 2005), are less prone to under- and overthinking in terms of decision time and commit fewer forecasting errors.

Rieskamp and Hoffrage (2008) study the effect of time pressure on the choice of inference strategies in an experimental setup. Subjects select the best option among four alternatives on the base of six cues over a total of 60 consecutive rounds. In the high time pressure (HTP) treatment, subjects need to decide within a time limit of 20 seconds, while the low time pressure (LTP) treatment permits 50 seconds. Subjects in the HTP treatment devote most of their attention to the most relevant cue (so-called non-compensatory heuristics). This finding is consistent with our conjecture that subjects in our HTP condition concentrate on the most focal cue (i.e. the last observed price). During the LTP treatment, instead, the share of rather time-consuming, *alternative-wise information search* (compensatory strategies), in which decision-making is based on the sum of weighted cues, increases.

Accordingly, we expect three main effects from time pressure in LtFEs, all working in the same direction. First, greater time constraint complicates the use of more sophisticated heuristics, such as trend-following, that seem to be required for bubble-and-crash dynamics. Second, coordination on trend-extrapolation among subjects ought to reduce as the time window to observe price trends shrinks. Interestingly, high forecasting coordination have been shown to drive bubble formation and price volatility in positive feedback markets (e.g., Hommes et al., 2005; Sonnemans and Tuinstra, 2010). Third, considering increased time pressure, subjects may encounter difficulties to submit predictions during each period. This may result in asynchronous updating of price expectations, which can bear a detrimental effect on coordination as well. Theoretically, this should favor market stability.

Hypothesis 2: *Time pressure reduces price volatility.*

3.3 Experiment

We designed an experiment in oTree (Chen et al., 2016) to test our hypotheses and answer the inherent research questions. The experiment took place in the CREED laboratory at the University of Amsterdam in May 2019. A total of 186 subjects participated in three treatments (see Table 3.1).[39] The sample is fairly gender-balanced (54% females) with an age range between 18 and 59 years (*M*=22.05, *SD*=4.58). For the most part, subjects were enrolled in the economic and business faculty (65%)

[39]One participant (AL12 market) left the experiment during period 57 due to personal reasons.

and combined reportedly little experience in financial trading (80% no experience). Mean earnings were 28 €, ranging between 13 € and 40 € (Appendix Table 3.4).

Upon arrival in the laboratory, an experimenter apprised participants of the ethical lab principles including fully anonymous participation through randomized seat allocation. Moreover, we permitted no form of communication between participants. Once seated, they found paper instructions on their desk and in digital form on the screen (see Appendix 3.A). The instructions tell the following story, standard in the LtFE literature. There is a market with a risk-free asset and a risky asset. Acting as financial advisors to large funds, participants have to forecast the price of the risky asset over several consecutive periods. Higher price forecasts lead to increasing investments in the risky asset by the fund. Some of these funds are advised by a participant to the experiment. Others use a fixed investment strategy (see fundamentalist traders in Section 3.3.1). Market equilibrium derived by the computer determines the price of the risky asset. The exact price generating mechanism (Section 3.3.1) and the total number of funds participating in the market remain unknown to participants. Earnings are based on participants' forecast accuracy. The closer the predicted price is to the realized market price for that period, the higher are the earnings for that period (see Section 3.3.3).

3.3.1 Price generating mechanism

The risk-free asset pays a commonly known interest rate r, held constant over time. Its return rate R is hence given by $1 + r$. The infinitely lived, risky asset instead pays an uncertain dividend y_t in each period t being independent and identically distributed with a commonly known expected dividend $\overline{y}$. The fundamental value p^f (FV) of the risky asset computes then as the discounted value of future expected dividends,

$$p^f = \frac{\overline{y}}{r}. \tag{3.1}$$

We can immediately see that the fundamental price stays constant over all periods. The price generating mechanism for the risky asset follows a standard two-period ahead forecasting model (Hommes et al., 2005):

$$p_t = \frac{1}{1+r}[(1 - n_t)\overline{p}^e_{t+1} + n_t p^f + \overline{y}], \tag{3.2}$$

where the realized price p_t in period t is a function of the sum of the weighted average between participants' average forecast $\overline{p}^e_{t+1}$ for period $t + 1$ and p^f, and the expected dividend $\overline{y}$, multiplied by a positive feedback parameter. The feedback parameter discounts expectations about tomorrow's price by the interest rate to the price of the current period. Equation (3.2) follows from market equilibrium for the risky asset. If traders expect the price in $t+1$ to move up, they will invest in the asset already in period t which - by market clearing - already drives up the market in period t. Note that forecasters observe realized prices only up to period $t-1$ when predicting the price of the next period $t+1$.

Robots take on the price stabilizing role of fundamentalist traders in the present design. Their influence on the asset price in the current period t is determined by the distance between the price in the last period p_{t-1} and the fundamental price p^f. The weight of fundamentalists grows with the level of aggregate mispricing (difference between the realized and the fundamental price) as follows

$$n_t = 1 - exp(-\frac{1}{20}|\frac{p_{t-1} - p^f}{p^f}|). \tag{3.3}$$

As bubbles blow up in real markets and overvaluation becomes apparent, a growing number of traders switch to fundamentalist strategies, which eventually brings bubbles to burst. The adjusting fraction of robot traders simulates such interdependencies. Moreover, earlier LtFEs without robot traders featured bubbles growing until they reached an artificial upper bound, invalidating the data collected after that point is reached. With robot traders bubbles burst due to an endogenous mechanism, which allows using all periods for analysis. Also note that the effect of robot traders is rather limited: n_t remains below 5% if the price is not more than twice the fundamental value, and n_t is only above 20% if the price becomes more than five times the FV.

3.3.2 Experimental design

This section outlines how the experimental design implements the two novel features - time pressure and extensive market length. Moreover, we describe further properties and the computer program in detail.

Decision time

For reasons of treatment comparability (see Table 3.1), we limit decision time also in the baseline condition. Accounting for actual decision times in earlier LtFEs (see Section 3.2.4), the low time pressure (LTP) condition allows 25 seconds, including 10 seconds *waiting time* in each period to submit a prediction about the price of the risky asset in the next period. It is critical to exclude the under-thinking fallacy (Moritz et al., 2014), possibly occurring in our high time pressure (HTP) condition, for the LTP condition by adding a waiting time of 10 seconds before participants are allowed to submit their prediction. During this waiting time, subjects can access all available information provided by the computer program and already place a prediction. However, the prediction remains open to changes. It can be finally submitted not before the waiting time is up and not later than the maximum decision time of 25 seconds is up (see Section 3.3.2 for more details). Note that subjects in the present experiment use on average only 10 to 12 seconds, including the waiting time, to submit a forecast. As there may still exist some form of time pressure, we refer to the baseline condition as *low* instead of no-time pressure. Based upon decision times applied in related literature (Section 3.2.5), the HTP condition permits subjects a decision time of six seconds per period to submit a prediction, featuring no waiting time.

Periods

We recognize that it requires substantially more periods compared to previous LtFEs, to answer questions associated with long-run behavior in asset markets. Upper bound restrictions, arising naturally, object exceeding a total length of 150 minutes. Considering the time for instructions, comprehension tests, questionnaire, and payment, this leaves subjects with 60-90 minutes to conduct the actual treatment. After satisfying the previous accounts on decision time and symmetric session duration between treatments (see Section 3.3.2), we suggest an approximately threefold increase of the runtime in earlier LtFEs, a total of around 150 periods.

Blocks

To prevent comparison issues from varying session length due to different condition duration, we split sessions into two blocks and let all subjects experience both time pressure conditions. We control for experience effects by switching the block order between-subject while switching the fundamental price within-subject. Regardless of the block order, the FV equals 126.4 price units (PU) in whatever condition is carried out during the first block (A) and 71.2 PU in the second block (B) condition (see Table 3.1).[40] To minimize subjects' bias, we deviate in both blocks from the FV of 60 used in earlier LtFEs. Controlling further for within-subject experience

[40]Note, analysis of the second block data is part of a further study. Since both blocks are relevant for the data collection, this chapter reports the entire experimental design and procedure.

effects, the HTP condition terminates after 159 periods, whereas the LTP after 146 periods, irrespective of the corresponding block. The two-block structure allows for possible changes between blocks regarding the set of funds active in the market (and thus the group of financial advisors/subjects) mentioned in the instructions. In line with Kopányi-Peuker and Weber (2019), leaving the exact termination period vague, subjects are further informed about an unspecified block duration of 120-180 periods. The unknown group size of six and the interest rate of 5% are held constant between blocks to grant comparability. Obtaining the FV defined in equation (3.1) fixes expected dividends at 6.32 PU in the first and 3.56 PU in the second block.

Experimental software

Participants can raise their hand and ask for help at any time during the experiment. In such a case, an experimenter walks up to the subject's computer terminal for assistance. Upon the instructions page and comprehension test, a practice round familiarizes subjects with the computer program. The respective time limit, interest rate, average dividend, and waiting time, if applicable, are announced immediately before the start of each block. After entering the forecasting stage, the computer interface displays a graph which plots participant's predictions along a blue line contrasted against a red line of realized asset prices (see Fig. 3.2). While this graph shifts continuously, showing the past 20 periods, a small graph in the upper left corner depicts the entire time series. Note, the price generating mechanism has the

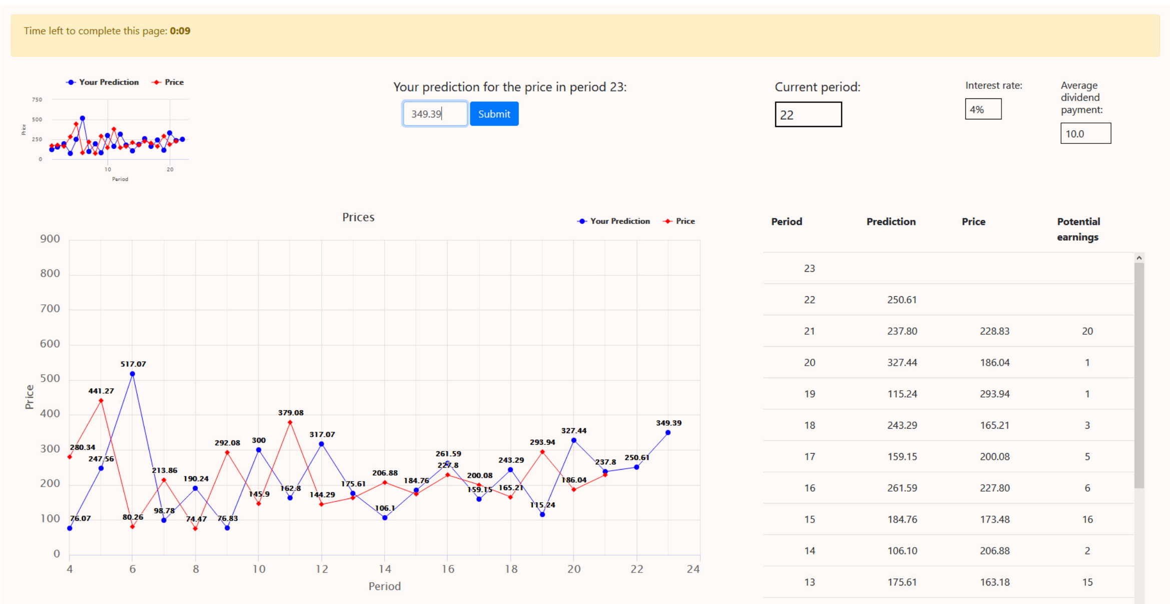


Figure 3.2: Participant screen.

last realized price lag behind the latest prediction. The vertical axis of both graphs rescales when prices or predictions increase beyond its current scale. Next to the graphical representation, a table lists past prices and predictions such that subjects can scroll up and down to see the prices and predictions from previous periods.

Predictions, up to two decimal places, can be made by typing in a submission box or by clicking with the computer mouse in the main graph. The choice set spans over the interval (0,10000]. Once clicked, the value appears as the current prediction in the graph and the box on top of the screen, changing automatically whenever a new value is clicked. Within the given time limit, subjects are free to change predictions any number of times. In the LTP condition, the program allows clicking values already during the waiting time. However, only after the waiting time is up the box will appear and show the last clicked value ready for submission. The remaining time in the current period is displayed at the top of the screen. If the time is up before a prediction has been actively submitted, either by clicking the "Submit" button or pressing "Enter" on the keyboard, the last value shown in the box will be automatically taken, and the next period starts. Interest rate and average dividend are displayed in the upper right corner.

Treatments

We run ten sessions with HTP as the first block condition and LTP as the second block condition (H treatment), and another 12 sessions in reversed order (L treatment, see Table 3.1). Sessions count either 12 or 18 participants as one experimental market contains six subjects. AL2, for instance, refers to the data of the second LTP group (market) for which LTP is the first block condition indicated by A. BH6 refers to the sixth HTP market for which HTP is the second block condition indexed by B.

After collecting the data, we worried that the software design could have been responsible for the observed treatment effect. It shows downward spikes within the

Table 3.1: Treatments of the experiment.

within-subject \ between-subject		**L**	**H**	**HS** (submit button)
First Block (A)	*Decisions*	146 periods, 25 sec/period	159 prds, 6 sec/prd	159 prds, 6 sec/prd
(r=0.05, y=6.32, p^f=126.4)	*Markets*	AL1-AL12	AH1-AH10	AHS1-AHS9
Second Block (B)	*Decisions*	159 prds, 6 sec/prd	146 prds, 25 sec/prd	146 prds, 25 sec/prd
(r=0.05, y=3.56 , p^f=71.2)	*Markets*	BH1-BH12	BL1-BL10	BLS1-BLS9
Number of groups		12	10	9

prediction series of a few subjects exclusively for HTP groups. Written comments in the questionnaire suggest indeed that some of these downward spikes, resulting from single-digit predictions, are because subjects ran out of time. They might be cut off after typing the first digits of their prediction in the box, which is then taken as their submitted forecast. Such downward outliers, which do not occur in the LTP groups, could settle dynamics at an early stage and stabilize prices after that. Although human trading failure likely represents an appropriate simulation of real market behavior,[41] we seek to rule out such reasoning as the main driver of the present results. We, therefore, run a robustness check with the following modification. If a subject neither clicks the "Submit" button nor presses "Enter" on the keyboard before the time is up, the computer will record no prediction for the subject during this period. The average over *actively submitted* predictions enters the price generating mechanism. In case no subject in a group submits a prediction, which has not happened once, the average over the submitted predictions in the last period will be considered. Since LTP markets do not seem to be affected by the *digit issue*, we conduct nine extra sessions with the "Submit" modification in which the HTP block precedes the LTP one (HS treatment, see Table 3.1).

3.3.3 Earnings

The computer randomly chooses ten periods from the first and ten periods from the second block for payout at the end of the experiment. If it selects a period in which no prediction was made, earnings for that period will be zero. The number of points $e_{i,t}$ earned for a period t is given by the absolute forecast error between a subject's prediction $p^e_{i,t}$ and the realized price p_t in that period, according to the below formula.[42] Earnings are transformed at a rate of 100 points per € and paid out privately at the end of the experiment.

$$e_{i,t} = \frac{200}{1 + |p_t - p^e_{i,t}|} \tag{3.4}$$

[41] So-called *fat-finger errors* through keyboard input or mouse click mistakes in human placed orders, are common in financial trading (Financial Times, 2019).

[42] Note that, as opposed to the hyperbolic scoring rule (3.4), a quadratic scoring rule - often used in earlier LtFEs (see, e.g., Assenza et al., 2019) - is very flat for small forecasting errors, and therefore heavily punishes the like. While related literature finds little difference between the two, we intended to raise incentives for precise forecasting by choosing the former.

3.4 Analysis and results

A glance at Figs. 3.3-3.5 provides a first qualitative description of the market dynamics complemented by an analysis of some descriptive statistics (Figs. 3.6 and 3.7). It follows a quantitative estimation of the aggregate expectations using regression analysis. For a more profound comprehension of market behavior, the analysis encompasses individual predictions and their difference in treatments. Deconstructing predictions brings out individual strategies. We thereby put special focus on the period intervals 11-50 and 106-145. Comparative analysis between behavior at the beginning and the end of the market allows us to answer the research question about long-run behavior. We choose the initial phase 11-50 for reasons of comparability with previous LtFE studies. It has become standard to dismiss the first ten periods due to initial learning effects (see e.g., Hommes et al., 2005) and analyze the remaining 40 periods instead. The final phase, periods 106-145, matches the observation span of 40 periods and is coherent between blocks of varying length (146 vs. 159 periods). The section concludes with a semantic analysis of strategy descriptions.

3.4.1 Prices

Eyeballing the prices series in Figs. 3.3-3.5, spots the largest bubble-and-crash cycles in terms of size and frequency, in the markets with LTP.[43] All AL groups break out into relatively large price fluctuations right from the beginning, while most of them calm down and settle near the FV towards the end. In contrast, HTP markets are less volatile in the beginning, particularly AHS markets and remain either stable around the FV or start fluctuating after a while. The following paragraphs describe aggregate price data based on various standard measures in the LtFE literature.

First, we aim to quantify price volatility and mispricing. Furthermore, we define criteria to assess bubbliness and convergence. To compare different market phases, we separate price series into three blocks, where I (Beginning) and III (End) correspond to the aforementioned period intervals 11-50 and 106-145, and II (Middle) to periods 59-98. Figures 3.6 and 3.7 visualize the statistical value of the respective measure (y-axis) for each group, separated by treatment (x-axis). It includes the mean (square) and median (circle) of each statistic. We provide separately in blue the values for phase I, and in red those for phase III. They are plotted to the left, respectively to the right of the main data column covering all periods 1-145 (black).

[43]As merely the first-block data is analyzed, this chapter uses LTP and AL synonymously. Analogously, HTP represents both AH and AHS groups.

Low time pressure

Decreasing price volatility generally characterizes LTP markets. Yet, some form of oscillation remains in most cases until the end. The key measure for price volatility, the interquartile range (IQR[44]), reduces significantly in the long run (left panel in Fig. 3.6; Appendix Table 3.8, columns 2-5). While large initial fluctuations have

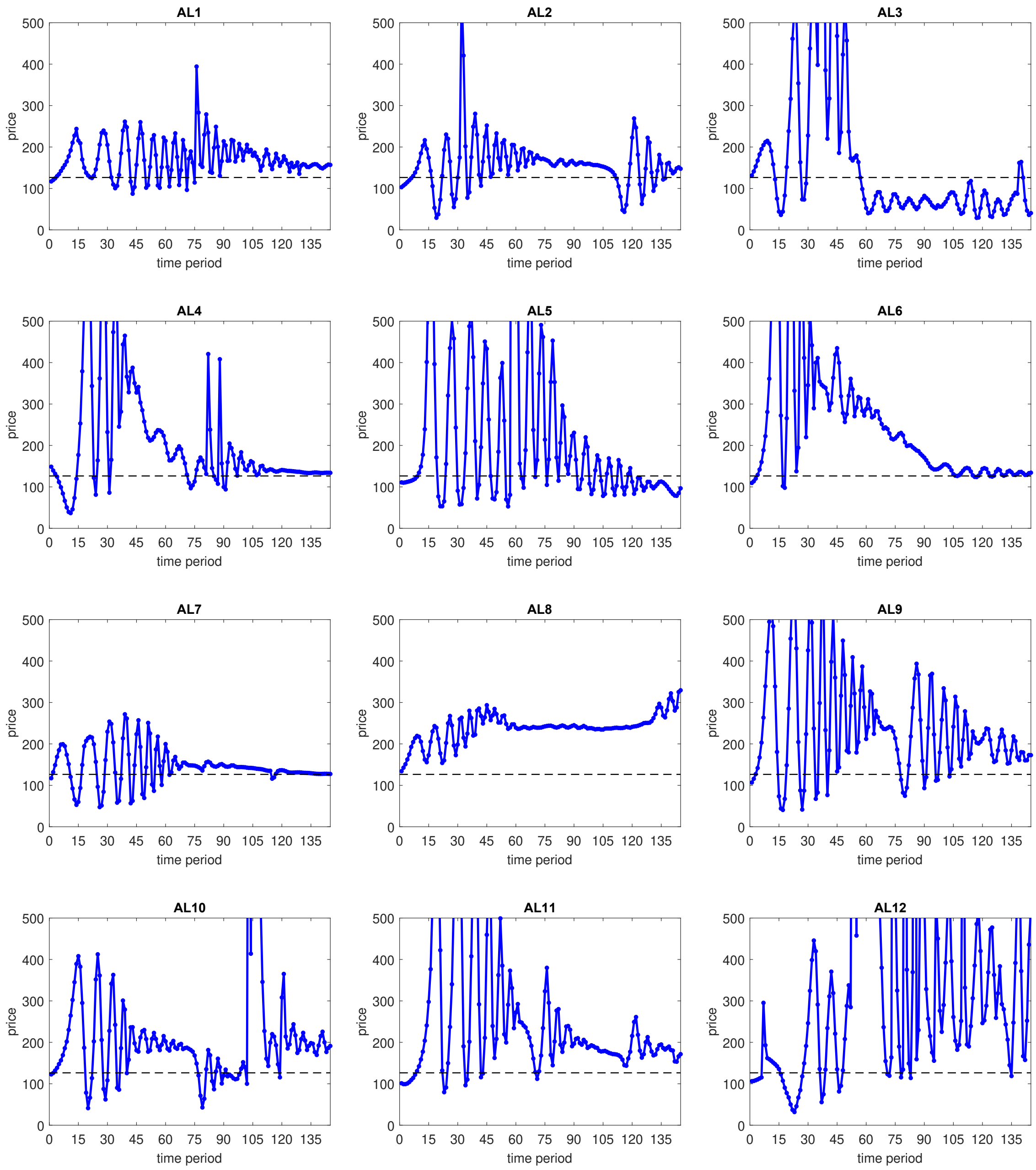


Figure 3.3: Prices in AL markets, dashed line indicates FV (126.4 PU).

[44]The interquartile range is the length of the interval containing half of the (ordered) data points.

been documented already in past LtFE studies, a general decline in volatility, including partial convergence to the FV has not been observed yet. In two-thirds of the markets price volatility decreases monotonically over time. Then again, there are a few markets in which volatility is lower in the second II than in the first interval I going up again in the third interval III (AL2-3,AL8) and one market in which volatility even peaks in II (AL12). For all markets except AL12, price fluctuation

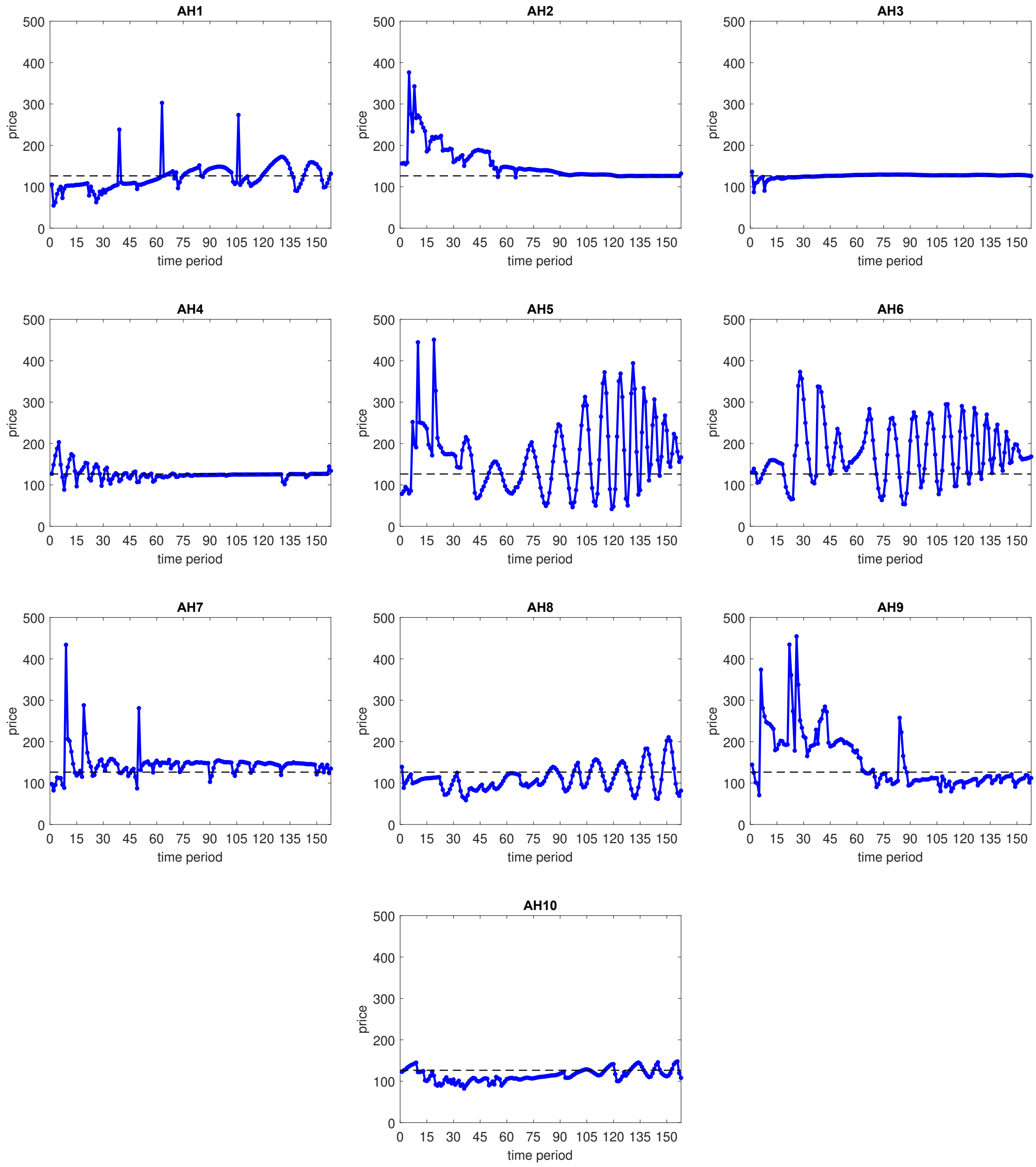


Figure 3.4: Prices in AH markets, dashed line indicates FV (126.4 PU).

declines considerably in either II or III. Standard deviation, an alternative measure for price volatility reaches the same conclusion (Appendix Table 3.8, column 6-9).

Next, we introduce the median of the *Relative Absolute Deviation* (RAD) from the FV as the principal indicator for the level of mispricing in the market (right panel in Fig. 3.6; Appendix Table 3.8, column 10-13). RAD per period is defined as $|p_t - p^f|/p^f$.[45] Its median reduces significantly between beginning and end across most markets, except AL8 and AL12. Two-thirds of the markets show monotonously decreasing levels of mispricing over the three intervals. This resembles exactly the previously described trend of price volatility. However, mispricing tells a different story for a few markets. For example, IQR in AL2 exhibits significant volatility growth in III after dynamics slow down in II. On the other hand, the median RAD takes on comparably large values in the mid interval but reduces toward the

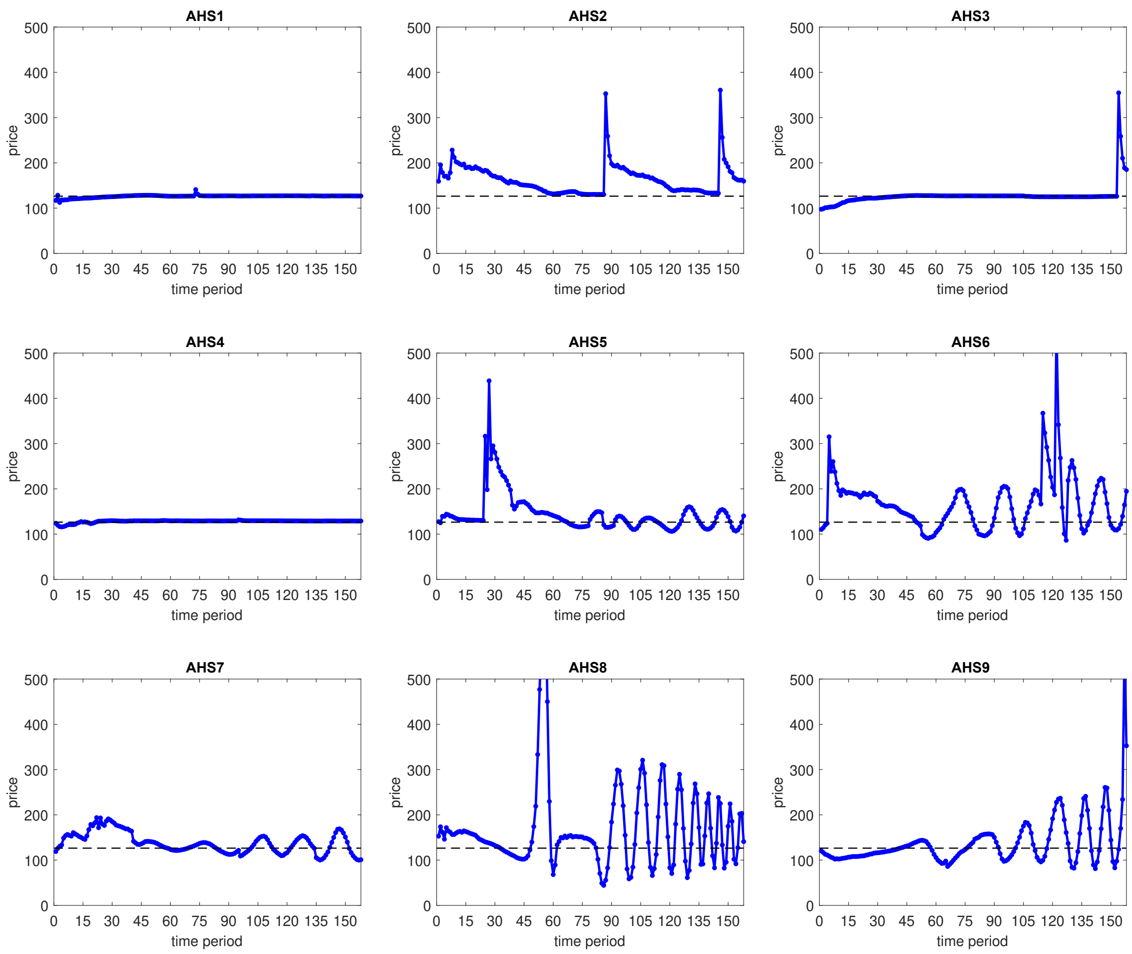


Figure 3.5: Prices in AHS markets, dashed line indicates FV (126.4 PU).

[45]This measure is introduced in Stöckl et al. (2010). Whereas they consider the mean RAD, we focus on the median, which is more robust to outliers.

end. Regarding AL2, Figure 3.3 illustrates both market stabilization during phase II and later convergence to a level much higher than the steady state, which explains substantial mispricing. Similar considerations underlie the discrepancy between volatility and mispricing in the markets AL3, AL8, and AL10.

Related to mispricing, we further assess the price series on bubbles. We define large (medium, small) bubbles for realized prices higher than $5\times$ ($2\frac{1}{2}\times$, $1.25\times$) p^f during at least four consecutive periods, and no bubble otherwise. The left panel in Figure 3.7 illustrates existing bubble sizes accordingly (3, 2, 1, 0). Bubbles appear in each LTP market, with large bubbles in four of these markets. The descriptive statistics in Appendix Table 3.8 (column B) report medium-size bubbles as average and median. Besides, we compute C as a measure of convergence, which is the longest sequence of prices such that all prices in the sequence lie within 5% of the

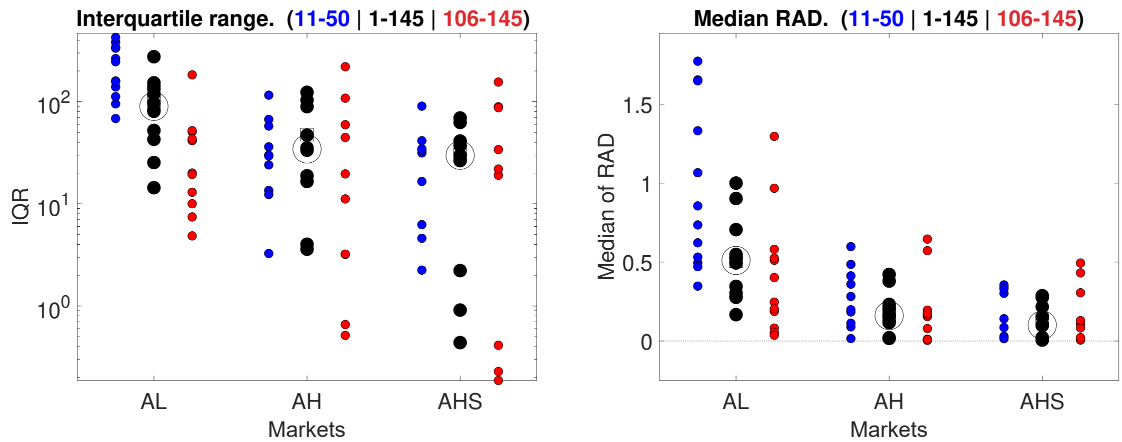


Figure 3.6: Measures for price volatility (IQR) and mispricing (Median of RAD).

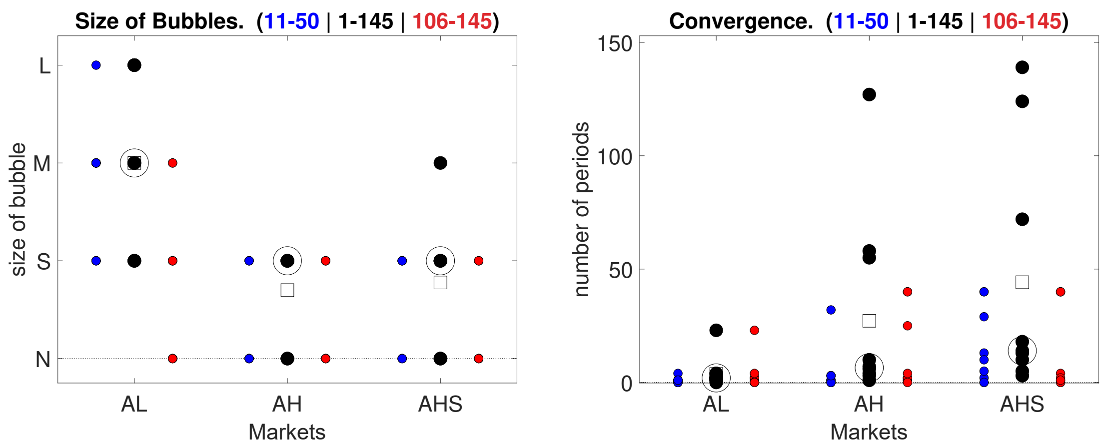


Figure 3.7: Measures for bubbliness and longest convergence sequence.

fundamental value p^f. The right panel in Figure 3.7 indicates that LTP markets exhibit longer convergence sequences to the FV only in one case (AL7, C=23 periods). We find relatively short converging sequences for the other AL markets (Appendix Table 3.8, column C). See Appendix 3.D for further data quantification measures.

We would like to draw the reader's attention to two additional phenomena. First, we find upward turns of prices before downward turns in 11 of 12 LTP markets. This finding reflects the famous presumption about "homegrown expectations" materializing, e.g., when people typically first place a positive bubble before a negative one in randomly drawn time series of stock markets (Huber and Kirchler, 2012).

Second, we frequently observe in a sub-sample of LTP markets (AL2,AL7,AL9-11) what could be called "step-wise convergence": fluctuations decline and prices converge to a value significantly above the FV. Soon after, these temporarily stable dynamics dissolve and engage in a new cycle of oscillations until they converge again, either to the FV or to a value between the first convergence point and the FV.

Apart, we denote two special cases. First, AL8 stabilizes on a price above the FV until close to the end. One participant (red line) happens to drive these dynamics by constantly predicting above the price level. The second outlier represents AL12 in which the blue participant manipulates the market through multiple, seemingly random, and extreme predictions, including the lower and upper bound integers of 1 and 10,000. This behavior causes the market to fluctuate unpredictably until the end. It appears meaningless to estimate market expectations (Section 3.4.2) or individual forecasting rules (Section 3.4.3) in this case. Nonetheless, we neither exclude these data points, nor any other outlier data from the analysis.

High time pressure

The HTP condition serves to understand the effect of severe time pressure on forecasting markets. Appendix Table 3.4 reports in detail on the perceived time pressure across both conditions. Four in five subjects (82.2%) feel time pressure in the HTP block, but only one in ten (11.3%) state so for the LTP block. It appears that almost all AH markets are affected by the inability of a few subjects to submit predictions in time. Analysis of the prediction series of the subjects in question suggests the program software has submitted merely these digits of a prediction which have been entered before the timer was up. We remind the reader that the H treatment does not feature a submit button. However, the similarity in plots of Figure 3.4 (AH) and Figure 3.5 (AHS) suggests a negligible impact from the design modification on the results between the two HTP treatments. Pair-wise testing of the descriptive statistics

(IQR, SD, median RAD) over period intervals 1-145, 11-50, and 106-145, by means of three non-parametrical tests further justifies data pooling (Kolmogorov-Smirnov, Mann-Whitney-Wilcoxon, Fischer-Pitman permutation; Appendix 3.C). All tests are two-sided, the latter performed with 1000 permutations. Our main findings do not discriminate between the HTP treatments AH and AHS, but rather between the two time pressure conditions HTP and LTP. It is therefore that the remaining sections refer to the AH and AHS treatment as joint HTP treatment versus the LTP treatment. Tables and figures still report the two HTP treatments separately.

Price volatility in HTP markets is lower than in LTP markets across all three phases I, II, and III (left panel in Fig. 3.6; Appendix Table 3.9, column 2-9). We observe that also subjects in the majority of HTP markets learn to stabilize prices over time, including immediate convergence to the FV, but temporal differences are not as accentuated as in the LTP markets.

Thus, HTP markets are typified by either immediate price stability or reduced levels of price volatility, except market AH4 evolving straight away into large oscillations. Its moderate amplitudes flatten out around period 70, resembling the typical pattern observed in LTP markets. All other markets exhibit limited volatility at least until period 50. After that, twofold dynamics emerge (Appendices 3.E.2 and 3.E.3). Roughly one-half remains stable, including frequent and fairly exact convergence to the FV (AH1-3,AH7,AH9,AHS1-4). The other half leaves the equilibrium path for a zigzag course of price levels (AH5-6,AH8,AHS5-9). We observe with this a copy of dynamics known from past LtFEs: reinforcing (AH8,AHS7,AHS9), converging (AH5-6,AHS8), or constant (AH10,AHS5-6) oscillations.

Volatility in terms of IQR (similar for SD) decreases monotonically in 9 of 19 markets. In two markets, it first rises and falls after that (AH6,AHS2), whereas for the rest, dynamics become more volatile towards the end. Three of these markets display increasing followed by declining fluctuations (AH5,AH10,AHS7). The remaining five markets exhibit monotonically growing volatility (AH1,AH8,AHS6,AHS8-9). The analysis of the median RAD in Appendix Table 3.9 (column 10-13) reveals less mispricing during the final periods (106-145) than during the beginning (11-50) in 11 of 19 markets. The level of mispricing remains constant over time in two groups and augments in the other six. Further analysis detects small bubbles in a majority of HTP markets but no medium or large bubbles. The average bubble size is considerably smaller than in LTP markets. In contrast, the number of sequential periods fulfilling the convergence definition from the previous Section 3.4.1 is significantly higher than in the LTP condition (right panel Fig. 3.7, Appendix Table 3.9). While the other reported measures do not differ significantly between AH and AHS markets, we note on average longer convergence sequences for the latter.

The treatment seems to have an unexpected implication on the genesis of price dynamics. Contrary to LTP markets, price series trend upward after market launch almost as often as the other way around. Time pressure arguably affects people's inclination of expecting price series to go up first.

We would like to remind the reader again that previous LtFE studies terminate after 50 periods. Interestingly enough, this proves to be a critical period length with regard to the present data. Analysis based on the first 50 periods alone would have resulted in a misleading characterization of price dynamics in multiple markets (AH2,AH4-5,AHS2,AHS5-6,AHS8-9).

3.4.2 Market expectations

We identify distinct behavior in both time pressure conditions through the following regression model:

$$\overline{p^e_{t+1}} = \alpha + \sum_{k=1}^{m} \beta_k p_{t-k} + \nu_t \,,$$

in which $\overline{p^e_{t+1}}$ averages participants' forecasts for period $t+1$ in a specific market, p_{t-k} represents the $t-k$ past market price lag, and ν_t the error term. This model includes *naïve* ($\beta_1 = 1$, all other parameters 0) and *adaptive* ($\alpha = 0$, β_k geometrically declining) expectations, and *trend extrapolation rules*:

$$\begin{aligned}\overline{p^e_{t+1}} &= \alpha + \beta_1 p_{t-1} + \beta_2 p_{t-2} + \nu_t \\ &= \alpha + \vartheta_0 p_{t-1} + \vartheta_1 (p_{t-1} - p_{t-2}) + \nu_t,\end{aligned}$$

with $\vartheta_0 = \beta_1 + \beta_2$ and $\vartheta_1 = -\beta_2$. The value of ϑ_1 (or $-\beta_2$) therefore measures the extent to which subjects are extrapolating past trends in the data. The following analysis is based on merely two price lags (Table 3.2), while Appendix Table 3.11 provides estimations of an extensive model with j = 4 lags. Yet, the reduced model ought to facilitate interpretation of the varying dynamics between treatments.

Low time pressure

The parameter β_1, referring to the last observed price (lag 1), exhibits positive and statistically significant values in all 12 LTP markets during the first interval I and remains so throughout the last interval III (upper section in Table 3.2). Except for AL3, β_2, corresponding to the second last observable price, is also statistically significant, yet negative in I. In the final interval III, β_2-values lose significance in two of the 11 markets (AL7,AL12). More remarkable is the decline absolute values of β_2, and hence in importance, for eight of the remaining nine markets. The weight of β_2 increases only in AL2. Our results suggest that trend extrapolation is much stronger in the early stages of the experiment and that it tends to decrease as time goes by. A smaller trend extrapolation coefficient explains the smaller fluctuations typically observed in the last 50 periods. In the LTP condition, trend extrapolation rules seem to be the norm, in particular for the first 50 periods, where β_2 is typically close to -1. This explains larger price volatility in that phase (under trend extrapolation the dynamics are unstable for $\beta_2 < -1.05$).

High time pressure

From the lower parts of Table 3.2 we can see three differences between HTP and LTP markets. First, α bears statistical significance for both analyzed intervals I and III in six of the 19 HTP markets. Instead, the LTP condition denotes significant alphas for both intervals in nine out of 12 markets. Second, we note increasing β_1-values in 13 of the 19 HTP markets (68%), though diminishing β_1-values in 75% of the LTP markets. Third and most relevant to distinguish expectation formation between the two conditions: the role of the second last observable price (β_2) is statistically insignificant in almost two-thirds of the HTP markets during phase I and remains so in merely one third during phase III. There are eight markets in which β_2-values switch from significance in I to insignificance in III. Only one market exerts the opposite (AHS4). In summary, the last observed price predominantly determines price expectations in the beginning, while past prices beyond one lag come to matter once markets have passed the initial phase. In the HTP condition, trend extrapolation, therefore, is much weaker than in the LTP condition. In particular, aggregate forecasting behavior is close to naïve expectations $\overline{p^e_{t+1}} = p_{t+1}$ for a substantial number of markets. Remarkably, as time goes by the tendency to extrapolate trends becomes stronger in a number of groups. This presents suggestive evidence that rather simple rules, including naïve or adaptive learning, dominate price formation in the beginning until a general tendency towards trend-following rules takes over at a later stage. Perceived time pressure may reduce after a while, supposedly giving room to cognitively more advanced strategies.

Table 3.2: Market expectations between treatments.

AL

Group	Phase	Const	Past Prices lag 1	Past Prices lag 2	LB	Obs
1	**11-50**	61.24	1.69	−1.03	0.58	38
	106-145	117.91	0.62	−0.35	0.50	38
2	**11-50**	76.57	1.46	−0.83	0.94	36
	106-145	59.46	1.46	−0.88	0.13	38
3	**11-50**	-	1.11	-	0.99	35
	106-145	20.97	1.39	−0.75	1.00	38
4	**11-50**	174.68	1.53	−0.91	0.69	36
	106-145	72.13	0.93	−0.45	0.03	38
5	**11-50**	89.35	1.84	−1.04	0.20	38
	106-145	84.75	0.97	−0.77	0.08	38
6	**11-50**	272.61	1.38	−0.91	0.06	36
	106-145	74.03	1.35	−0.90	0.53	38
7	**11-50**	77.99	1.46	−0.95	0.27	38
	106-145	34.85	0.73	-	0.98	38
8	11-50	82.41	1.32	−0.63	0.00	38
	106-145	-	1.59	−0.50	0.00	38
9	11-50	170.54	1.45	−0.97	0.01	38
	106-145	163.12	0.84	−0.64	0.13	38
10	**11-50**	101.66	1.46	−0.87	0.64	38
	106-145	-	1.15	−0.22	0.06	38
11	**11-50**	150.12	1.80	−1.02	0.31	36
	106-145	62.34	1.48	−0.80	0.45	38
12	11-50	29.46	1.87	−0.99	0.04	38
	106-145	234.72	0.44	-	0.55	38

AH

Group	Phase	Const	Past Prices lag 1	Past Prices lag 2	LB	Obs
1	**11-50**	-	0.95	-	0.96	36
	106-145	-	1.67	−0.72	1.00	37
2	**11-50**	-	0.85	-	0.39	38
	106-145	-	1.82	−0.81	0.66	38
3	**11-50**	-	1.00	-	0.08	38
	106-145	-	1.81	−0.85	0.02	38
4	**11-50**	80.51	0.88	−0.51	0.52	38
	106-145	59.66	1.04	−0.52	0.95	38
5	**11-50**	-	0.78	-	1.00	38
	106-145	112.54	1.48	−1.00	0.07	38
6	**11-50**	-	1.36	−0.53	0.09	38
	106-145	123.24	1.35	−0.96	0.05	38
7	**11-50**	146.72	-	-	0.98	38
	106-145	66.85	0.55	-	0.38	38
8	**11-50**	12.91	1.58	−0.74	0.29	38
	106-145	19.36	1.89	−1.06	0.67	38
9	**11-50**	137.42	0.44	-	0.29	36
	106-145	60.41	0.42	-	0.77	38
10	**11-50**	37.01	0.61	-	0.70	38
	106-145	28.67	1.53	−0.76	0.82	38

AHS

Group	Phase	Const	Past Prices lag 1	Past Prices lag 2	LB	Obs
1	**11-50**	−2.85	1.02	-	0.99	38
	106-145	-	0.77	-	0.69	38
2	**11-50**	−9.40	1.07	-	0.38	38
	106-145	-	1.46	−0.43	0.24	38
3	**11-50**	-	0.98	-	0.56	38
	106-145	-	1.49	−0.57	0.59	38
4	**11-50**	-	1.35	−0.43	0.98	38
	106-145	-	0.99	-	0.98	38
5	**11-50**	-	0.89	-	1.00	34
	106-145	7.78	2.01	−1.07	0.51	38
6	**11-50**	−18.30	1.11	-	0.53	38
	106-145	-	0.71	-	0.92	34
7	**11-50**	-	1.00	-	0.74	38
	106-145	-	1.85	−0.90	0.68	38
8	**11-50**	-	2.69	−1.66	0.31	38
	106-145	86.96	1.53	−0.99	0.39	38
9	**11-50**	−11.04	1.10	-	0.63	38
	106-145	26.41	1.90	−1.05	0.43	38

Note: Phases are marked bold if Ljung-Box (LB) test finds no significant autocorrelation in the data (p>.05).

3.4.3 Individual predictions

The preceding paragraph relates market expectations to price patterns. To verify the findings, the analysis proceeds by examining expectation formation at the micro level. Individual predictions ought to elucidate what is driving markets to create bubbles and crashes after all. In this regard, we first showcase the individual prediction series of representative markets for each condition (Figs. 3.8 and 3.9) before the section concludes by eliciting individual strategies in more detail.

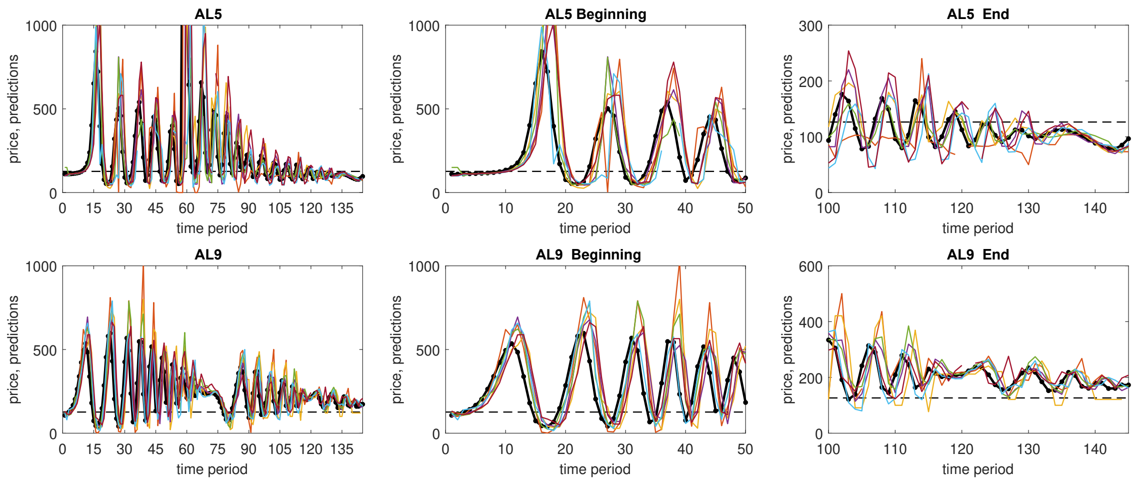


Figure 3.8: Prices and predictions in LTP markets in different phases.

Coordination and stability

Figures 3.8 and 3.9 present the individual prediction series of all six subjects in some exemplary markets. The left-hand panels exhibit periods 1-145, the middle panels periods 1-50, and the right-hand panels periods 96-145. In doing so, we extend interval I for the very first ten periods to capture crucial dynamics of between-subject coordination in the very beginning. Consistently, we also extend interval III for periods 96-105 to ensure comparability between same interval lengths. The thin colored series represent subjects' predictions, the bold black line represents prices, and the horizontal dashed line the FV (126.4 PU).

Individual predictions, except for some outliers, seem to reliably match the shape of the price line. Little is new in this observation, as prior studies commonly find that LtFEs with positive feedback are typified by a high degree of coordination. However, taking a closer look at the second column panels in Figures 3.8 and 3.9, one can detect a striking difference between LTP and HTP markets. Most LTP markets demonstrate quasi-immediate coordination of individual forecasts when past

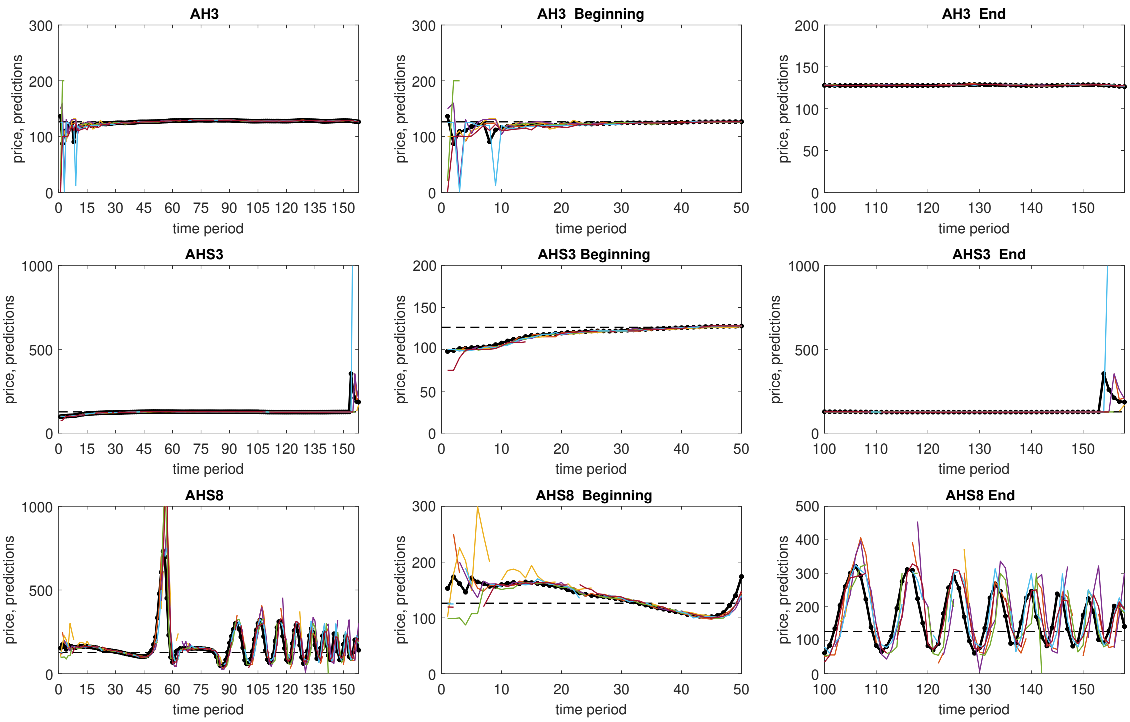


Figure 3.9: Prices and predictions in HTP markets in different phases.

prices become observable for the first time in period three. In contrast, the predictions in most HTP markets start converging to the black price line a few rounds later between periods five and 10. We rank the mean in standard deviations between subjects' forecasts of the first five periods after price disclosure, from highest to lowest. Out of the 31 markets in scope, HTP markets occupy the first 12 ranks. Appendix Table 3.10 confirms significant differences in forecast coordination between the two conditions (Mann-Whitney-Wilcoxon test, α=.05). The dispersion in expectations around the FV leads to an average expectation close to the FV, thereby promoting convergence to the steady state (cf. Muth, 1960).

Instead, it appears that ample time constitutes fertile ground for rapid coordination. Greater initial coordination implies that subjects in LTP markets match the price line earlier. Essentially, such coordination has been shown to produce trend-chasing behavior in forecasting, which in turn causes fluctuations in LtFEs (see, e.g., Hommes et al., 2005).[46] Some markets, such as AL9, break out into large fluctuations right away. In others, oscillations evolve with delay (i.e., AL5). Trend-chasing behavior describes strategies in which subjects observe recent price changes and

[46]Heemeijer et al. (2009) consider LtFEs with positive and negative feedback. Under positive feedback, subjects quickly coordinate their forecasts, and persistent deviations from the FV emerge. For negative feedback, on the other hand, the initial dispersion between predictions is substantial (as in our HTP condition) and, as a consequence of this disagreement, prices quickly converge to the FV.

extrapolate their forecasts on future prices thereupon. The strategy analysis below provides evidence for larger shares of trend-following rules during the early market stage I in the LTP condition. "Missing" initial coordination in HTP markets seems to prevent prices from adopting a clear trend early on. Often they still fluctuate but rather juggle in a narrow range, as if there was an invisible hand stabilizing the price line. The initial stability arguably facilitates stable dynamics afterwards. Indeed, we observe higher levels of forecast coordination in LTP markets throughout the entire experiment (Appendix 3.E.4.)[47]

Forecasting rules

This paragraph investigates further behavioral determinants beyond coordination with the following forecasting model:

$$p^e_{h,t+1} = \alpha + \sum_{k=1}^{m} \beta_k p_{t-k} + \sum_{\ell=0}^{n} \gamma_\ell p^e_{h,t-\ell} + \nu_t \,,$$

regressing price expectations $p^e_{h,t+1}$ of subject h not only on past prices as incorporated in the market expectations model (Section 3.4.2), but on own past predictions $p^e_{h,t-\ell}$, too. As before, it includes naïve ($p^e_{h,t+1}$=p_{t-1}), adaptive expectations ($p^e_{h,t+1} = \gamma p_{t-1} + (1-\gamma)p^e_{h,t}$), and trend extrapolation rules:

$$\begin{aligned} p^e_{h,t+1} &= \alpha + \beta_1 p_{t-1} + \beta_2 p_{t-2} + \nu_t \\ &= \alpha + \vartheta_0 p_{t-1} + \vartheta_1 (p_{t-1} - p_{t-2}) + \nu_t, \end{aligned}$$

with $\vartheta_0 = \beta_1 + \beta_2$ and $\vartheta_1 = -\beta_2$. Table 3.3 presents a comprehensive summary of differences in expectation formation between treatments. It compares the average over the *largest past price lag* that subjects take into account to form predictions. The largest past price lag refers to the non-zero $\beta-$coefficient, corresponding to the observed market price dating back farthest. During the early stages of LTP markets, participants' forecasts depend on past prices dating back close to three periods into the past. Throughout the market, these subjects seem to slowly shift their focus toward the most recent price. Questionnaire comments reveal that several subjects correctly relate constant forecasts with better predictability, and thus earnings (see semantic analysis in Section 3.4.4). Participants in HTP markets exploit data on past

[47] We further observe discontinuous prediction series in HTP but not in LTP markets. Its implications can be associated with the concept of asynchronous expectation updating. The scope of this paper, though, does not allow studying its isolated effect on market behavior.

prices significantly less. On average, they look back just slightly further than the last two prices. In the very beginning, subjects focus even on the most recent price alone. Soon after, they double the scope of past price information. Albeit the difference in largest lags between treatments continues to reduce over time, we note a significant and lasting effect of time pressure on strategy formation.

Table 3.3: Mean largest significant past price lag used for predictions by period interval.

	1-10	***11-20***	***21-30***	***31-40***	***41-50***	***51-60***	***61-70***	***71-80***	***81-90***	***91-100***	***101-10***	***111-20***	***121-30***	***131-45***
AL	3.07	2.89	2.79	2.69	2.15	2.12	2.26	2.54	2.31	1.81	2.07	1.83	2.28	2.63
AH	0.85	2.10	1.73	1.87	2.28	1.58	1.60	2.47	2.22	2.17	1.78	2.12	2.43	2.30
AHS	1.13	1.61	1.44	1.56	1.31	1.52	1.52	1.56	1.83	1.33	1.70	1.52	1.57	2.00
HTP	0.98	1.87	1.60	1.72	1.82	1.55	1.56	2.04	2.04	1.77	1.75	1.83	2.03	2.16

Note: Number of past price lags = 4. HTP corresponds to the weighted average over the AH and AHS markets.

3.4.4 Semantic analysis

Among other things, the questionnaire asks subjects to describe the prediction strategies they followed in each condition. Semantic analysis detects that subjects use 20% more words to verbalize their strategies in the LTP than in the HTP condition. One participant writes *"Too long"* regarding its LTP strategy. Although this exemplary comment consists of few characters only, it illustrates the high degree of sophistication representative for expectation formation in LTP markets. Ranking the 20 most frequent words in these comments brings forward additional evidence for the evolution of simple forecasting rules when time pressure is higher. The usage frequency of the word "price" is 40% greater in HTP than in LTP strategies, while "prices" is used more than twice as frequently in LTP compared to HTP comments. The predominantly singular use of the term "price" in HTP strategies points toward *one* focal point in expectation formation. There are good reasons to suspect that this refers to the last observed price. On the other hand, the plural use of "prices" likely mirrors the multitude of past prices subjects in LTP markets base their strategy on. The exclusive presence of the word "period" in HTP strategies plus the 50% greater usage frequency of "periods" in LTP comments tell the same story.

> *"I tried to mimic the price line and extrapolate based on its volatility"*

> *"I was mainly following previous patterns of the graph, but then I tried to stay in a stable price"*

These comments, both describing LTP strategies, are exemplary for the trend-following behavior observed in these markets. Analysis of the written comments demonstrates that the terms "trend", "follow", and "pattern" appear significantly more often in LTP comments, with the latter being found even exclusively there. Such cumulative occurrence of key words reveals the extended scope of past prices considered in forming predictions when subjects face only limited time pressure.

In contrast, the unique presence of the word "same" in HTP comments provides suggestive evidence for the presence of the last price as the sole reference point when time pressure is high. Subjects mention the term "time" more frequently in HTP than in LTP comments. This reflects subjects' greater struggle with time in HTP markets (Table 3.4). The fact that the term "click" appears exclusively in HTP comments suggests that subjects make greater use of the clicking feature in the computer program, which seems reasonable given the intense time constraint. Elevated usage frequency of terms such as "just", two times higher than in LTP comments, further attest that subjects exhaust information about past price trends to a lower degree. This materializes in more straightforward strategies, including naïve and less trend-following rules. The below comments illustrate these findings.

"I had no time to think, so I almost always filled in the previous price"

"Panic and try to click something in time"

3.5 Discussion

The final section ruminates on the above results in responding to the initial research hypotheses of Sections 3.2.3 and 3.2.5. Besides, it comprises limitations of the present study and an outlook on follow-up research.

3.5.1 Convergence of prices

The LTP condition in the present study provides comprehensive evidence that markets can learn the FV and even stabilize in the long run. Similar to previous LtFEs, we observe three different dynamics during the first 50 periods. While all LTP markets are characterized by an initial phase of profound volatility, subsequently price amplitudes remain roughly constant in the first type, grow in the second type, and decrease in the third type. However, none of the markets converges to the FV during the first 50 periods. It is only afterwards that price series display a general tendency of convergence. We conclude that a market length of fewer than 60 periods is too short to capture converging dynamics toward the FV reliably. When extended sufficiently, as done for the first time in the present study, markets prove capable of learning the FV and stabilizing around it in the long run.

Trend-following behavior in individual forecasting strategies favors large price fluctuations. As prices diverge from the FV-level, average accuracy in predictions drops by design. This results in dwindling earnings. A glance at the data promotes that by realizing this causality, subjects shift toward less trend-chasing behavior over time, that is, considering price information dating back less far in history (Table 3.3). Focusing primarily on the last observed price when forming predictions - essentially naïve expectations - triggers a chain reaction. Stable forecasting renders stable price dynamics, which generate higher earnings. This is because forecasting constant time series is an easier task. Hence, for the standard LtFE design with positive feedback and moderate time pressure, we find confirming evidence for Hypothesis 1. Prices converge to the FV in the long run.

By relating market maturation to volatility decline, this result shines in an ambiguous light considering both experimental and empirical evidence. The continuous debate surrounding a new bubble in the housing market, dangling like the sword of Damocles above the economy, questions if the market length in this study is enough to capture long-run dynamics. Lahav's experimental paper (2011) attributes less value to experience gained from a single long-run market than to experience gained from repeated short-run markets. Yet, evergreen bubble appearances in emerging markets - the cryptocurrency bubble being a recent example - do not attest a learning effect to previous observations of overheated markets. Although the present results link learning effects in the LtFE setup to long-run market duration, there are some cases in which price convergence occurs only temporarily until volatile dynamics break out again (Fig. 3.3, e.g., AL4,AL9-10). Note that none of these temporary market equilibria are steady states.

Reducing decision time in the HTP condition distorts the picture. Long-run dynamics become manifold. Some HTP observations resemble converging LTP markets. Others do not approach the FV but remain oscillating in constant two-cycles until the end, or even increase in fluctuations. Above all, we notice a striking difference at the beginning of HTP markets compared to what we know from experimental asset markets so far. This forms the starting point for the following paragraph.

3.5.2 Time pressure can stabilize prices

When comparing LTP and HTP conditions, we detect markets in the latter to be generally less volatile. Although differences in price volatility reduce upon market maturation, the second novel finding in this study represents the initial price stability of HTP markets. Seeming counter-intuitive at first, the impairing implications from increased time pressure exert a stabilizing effect on markets. The present study identifies mainly two reasons for this. First, analyzing the information base for expectation formation manifests that less price history is exploited in subjects' forecasting strategies than in LTP markets (Table 3.3). It seems logical that less decision time limits subjects' information base with the consequence of paying increased attention to the most relevant cue - the last observed price (e.g., Rieskamp and Hoffrage, 2008). Following the latest price primarily leads naturally to less forecast deviation from the current price level and the mitigation of destabilizing trend extrapolation. Volatility slows down. This effect becomes particularly obvious in the very beginning after market launch (Table 3.3).

What the study brings forward as the second reason is rather a consequence of the first one. The inability to collect thorough information in HTP markets makes individual predictions prone to differ within groups and periods. Lower degrees of coordination appear to counterbalance trend-extrapolative forecasting behavior, which would otherwise fuel bubble formation. An aggregated survey series on the U.S. housing market by Case et al. (2012) comments aptly: "when price trends are strong, there is little disagreement among respondents. When there is ambiguity, respondents seem, not surprisingly, to have a much less clear picture." This conclusion is in line with earlier LtFE studies, comparing the sign of the feedback parameter. As previously mentioned, Heemeijer et al. (2009) find a larger share of stable markets and less coordination among subjects' forecasts for negative than positive feedback markets.

As subjects get used to it, time pressure supposedly loses its severity in the long run. Table 3.3 underpins this rationale, stating a steep increase between periods 1-10 and 11-20 in terms of the mean largest past price lag. Whereas time pressure seems to impair subjects' information base in the very beginning, they incorporate more information on past prices in their forecasting rules soon after. How is such a behavioral shift reflected in the market? Several HTP groups leave the calm waters of stable pricing and transform into volatile markets. Such mutation may be triggered through a switch in forecasting strategies. Subjects no longer use simple rules but turn toward rather complex strategies, integrating recent trends. Although analysis of strategy identification states this tendency on the treatment level, it fails to discriminate on the group level. It seems logical that relaxing perceived time pressure creates room for reconsidering strategies. Participants switch forecasting strategies, misguided by the idea of extrapolating the current price trend to improve prediction accuracy. If followed by a critical mass, this can cause stable dynamics to drift toward amplifying volatility.

Except for the steady state, complementary markets are characterized by a signature price-prediction discrepancy. This is due to their nature, in which price expectations feed back into price formation. Trending prices, driven by trend-extrapolating expectations overshoot the like by definition. Subjects are constantly informed that their predictions are too conservative and, hence, pushed to adjust their expectations further in the direction of the recent trend line. This continues until the market realizes the imminent crash, just before the self-fulfilling mechanics of any demand-driven market structure let the bubble ultimately burst. The following downswing takes place even faster than the preceding uprise. By then, subjects avoid predicting too conservatively. Positive feedback markets are vulnerable to get stuck in vicious bubble-and-crash cycles. The list of examples is long (dot-com bubble, financial crisis 2008-09, 2018 cryptocurrency crash, 2020 stock market crash, ongoing Canadian property bubble, ...). Among these examples, the LTP condition and the described characteristics have been particularly compared to the housing market in previous work (e.g., Bao and Hommes, 2015), which Case et al. (2012) report as "a market largely driven by expectations. People seem to form their expectations from past price movements rather than having any knowledge of fundamentals" (p. 6), [and seem] "very much aware of trends in home prices... [as] there is a strong correlation between the respondents' stated perceptions of price trends and actual movements in prices" (p. 39).

Instead, we must not omit that a large portion of HTP markets remains stable until the end.

> *"I did what first came to my mind and didn't have time for second thoughts which ended up being a good thing"*

This HTP comment highlights the effect which time restriction exerts on experimental asset markets. Almost like Adam Smith's *invisible hand*, time pressure steers market prices toward their underlying values. At first sight coming off as a serious handicap in decision-making, time pressure turns out to be a blessing in LtFEs. This study provides suggestive evidence that by manipulating decision speed, forecasting performance can be improved, and market volatility reduced.

3.5.3 Limitations and future research

The discussion around repeated versus long-run experience is carried forward to a follow-up study comparing the data from the second block to the first one analyzed here. Besides that, further questions emerge from this study. One aims at the twofold long-run dynamics after common initial stability, which are observed in HTP markets. It remains unclear what triggers late-emerging volatility in some markets while others stay unaffected and stable throughout the end. Even though this study shows that asynchronous updating of price expectations results from increased time pressure, it remains questionable whether this bears an additional stabilizing effect on markets.

A future variation of the present design could be a joint treatment, in the sense that part of the subjects face HTP and the other part LTP. It would be interesting to see which of the behavioral implications observed in this experiment dominates and what it would mean for the aggregate dynamics.

Similarly interesting is how time pressure affects market dynamics relative to varying feedback strength in the price-generating mechanism (see Equation 3.2). Earlier positive feedback markets exhibit a negative relationship between convergence to the FV and the size of the feedback parameter (Bao and Hommes, 2015; Sonnemans and Tuinstra, 2010).

In negative feedback markets, on the other hand, there is reason to assume that time pressure increases volatility. If forecasters focus predominantly on the last price, much as in the present HTP condition, they get likely caught in constant fluctuations as predicted by the cobweb model and observed throughout the initial periods in previous LtFE experiments with negative feedback (e.g., Hommes et al., 2007).

Furthermore, it would be insightful to see how subject priming plays out on the design implications introduced in this paper. Hanaki et al. (2019) review experimental asset markets, where framing biases individual expectations.

Aggregate market dynamics in groups of six subjects appear to impact their forecasting rules substantially. More research is needed to understand the evolution of individual learning strategies. It may prove informative to study single-person markets with and without revealing the isolated market structure to subjects.

In this study, we advocate time pressure as an essential characteristic of financial markets. A follow-up study could introduce *continuous time* to LtFEs.[48] Subjects

[48] See Calford and Oprea (2017) for an interesting implementation of continuous time to strategic games through a *freezing protocol* that allows players to react "immediately" from a game-time perspective to opponents' strategy adjustments.

forecast asset prices in continuous time with their earnings being generated in the same fashion. They can change their forecast as often as they wish. However, each adjustment is costly, similar to broker fees in real financial markets or opportunity costs in general. This appeals to a rather fundamental discourse: Is it wise to wait and gather lots of information before making a decision, or is it likely that the right moment has then already passed?[49] Does fast decision making - reacting to the slightest hint to be the first - pay off, or does it necessarily imply a lack of rational situation assessment (see, e.g., Berninghaus et al., 1999; Magnani et al., 2016)? By introducing small costs when changing a price forecast, it is left up to the subject deciding how often to respond to price developments. We merely analyze the most profitable strategies.

3.5.4 Concluding remarks

In this paper, we advance the design of Learning-to-Forecast Experiments (LtFE) by two features, innate to real asset markets. So far overlooked, these parameters prove to affect results significantly. First, we extend the market length of standard LtFEs by tripling the number of periods. It shows that demand-driven markets are able to converge to the fundamental value when provided with sufficient feedback periods. Volatile dynamics, known from previous studies, can be regarded as some sort of "teething pain". Markets seem able to overcome the like in the long run. Second, we find that inducing time pressure gives rise to a surprising effect. Upon market launch, dynamics do not break out into large fluctuations, as known from prior LtFEs, but remain rather stable. Yet, markets may get used to time pressure and behave differently after perceived time pressure fades. While we note significantly lower volatility in high time pressure markets (HTP) throughout all stages compared to low time pressure (LTP) markets, it remains unclear why volatility in some of the HTP markets increases over time. Third, we know from earlier LtFE studies that trend-following behavior in forecasting strategies benefits bubble formation. This seems to be also the case in the present data. Looking at the participant level, we manifest that subjects in LTP markets base their strategies on a larger number of past prices than subjects in HTP markets. Strategies entail higher degrees of trend-chasing behavior. The number of considered past prices reduces over time and can explain declining volatility. More analysis is needed to confirm the identified reasons for the difference in behavior and to gain new insights from its implications on human forecasting strategies.

[49] See related TED talk by Ambühl (2016) on this dilemma in the context of marriage.

Appendix 3.A Experimental instructions

The main part of the experiment consists of two phases. In each phase, you have to make a large number of consecutive stock market predictions. After the main part of the experiment, but before the payments are made, you will have to do one short additional task and fill out a short questionnaire. **Please read these instructions carefully.** If at any moment in time you have a question, please raise your hand, and one of the experimenters will come to your seat to assist you.

General information: Your role is that of a **financial advisor** to a pension fund that wants to optimally invest a large amount of money. The pension fund has two investment options: a risk-free investment and a risky investment. The risk-free investment is putting money on a bank account, paying a fixed interest rate. The alternative, risky investment is an investment in the stock market. In each time period, the pension fund has to decide which fraction of its money to put on the bank account and which fraction of the money to spend on buying stocks. To make an optimal investment decision, the pension fund needs an accurate prediction of the price of the stocks. As its financial advisor, you have to predict the stock market price (in euro) during a number of consecutive time periods.

Information about the stock market: The stock market price is determined by equilibrium between the demand and supply of stocks. The supply is fixed during the experiment. The demand is determined by the aggregate demand of a number of large pension funds active in the stock market. Some of these pension funds are advised by a participant to the experiment, others use a fixed investment strategy. The price of the stocks is determined by market equilibrium, that is, the stock market price in period t will be the price for which aggregate demand equals supply.

Information about the investment strategies of the pension funds: The precise investment strategy of the pension fund that you are advising and the investment strategies of the other pension funds are unknown. The bank account of the risk free investment pays a fixed interest rate per time period. The holder of the stocks receives a dividend payment in each time period. These dividend payments are uncertain and may vary over time. Economic experts of the pension funds have computed the average dividend payments per time period. The return on the stock per time period for a pension fund is given by the dividend payment for that period and profits or losses from possible price changes of the stock. As the financial advisor of a pension fund, you are **not** asked to forecast dividends, but you are only asked to forecast the price of the stock in each time period. Based upon your stock price forecast, the pension fund that you advise will make an optimal investment decision. The higher your price forecast for the next period is, the larger will be the fraction of money invested by this pension fund in the stock market in the current period, so the larger will be its demand for stocks.

Forecasting task of the financial advisor: The only task of the financial advisors in this experiment is to forecast the stock market price in each time period as accurately as possible. For each period, there will be only a limited amount of time to make a prediction. After this **limited amount of time**, the next period starts. This **time limit** will be the same for each period in the same phase, but it may be different between the two phases. Additionally, there may be a **waiting time** in either one of the phases, which again will be the same for each period in the same phase. Submitting a prediction in a period is only possible after the waiting time for that period is over. The relevant time limit and, if applicable, waiting time will be announced immediately prior to the start of each phase.

In the first period of each phase of the experiment, you have to predict the price for the first period, and in the second period, you will have to predict the price for that second period. **Only after the second period has finished, the realized price for the first period will be announced.** After that, you have to give your prediction for the price in the third period. When the third period is finished, the realized stock price for the second period will be revealed, and so on. This process continues until the end of the phase. To forecast the stock price p_t for period t, the available information provides:

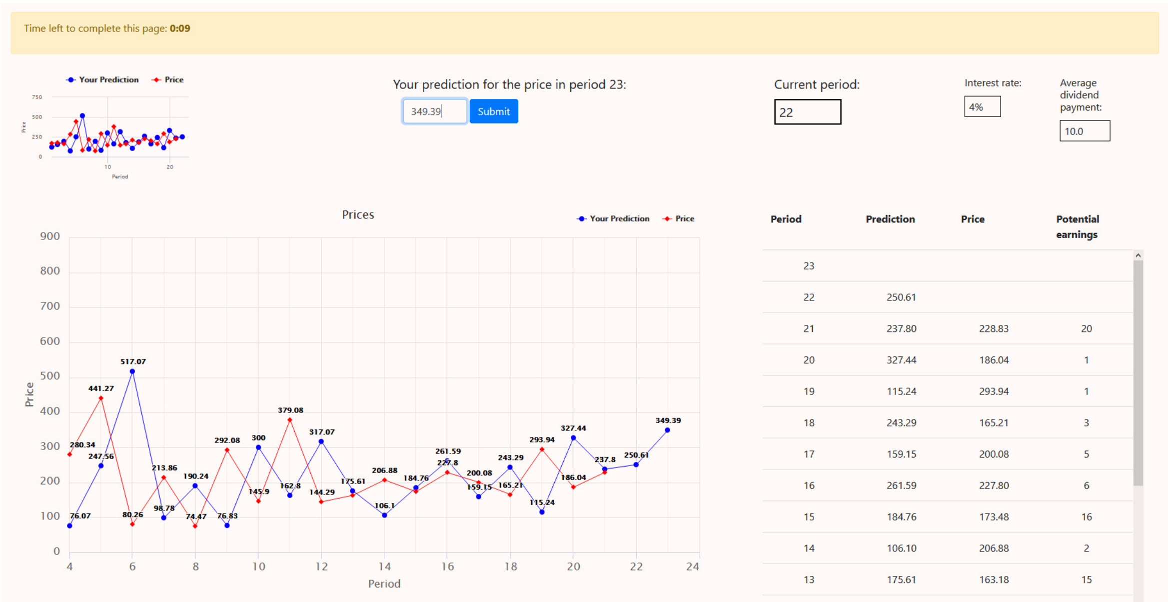
- past prices up to period $t-2$
- your own past predictions up to period $t-1$

Your past predictions and prices are represented both in a table and graphically, see the computer interface below for illustration. The main graph and the table shows prices (red) and your predictions (blue) in the last 20 periods. You can scroll back/up to see prices and predictions from earlier periods. In the upper left corner of the screen, a smaller graph shows all your predictions and prices over the entire horizon up to the current period, and in the upper right corner, you find the relevant interest rate and the average dividend payment. The vertical axis of the graphs rescales when prices or predictions increase beyond the current scale of the graph.

You can make a prediction by clicking with your **mouse** in the graph. The value you click will then be shown as your current prediction in the graph and the box on the top of the screen, and that value will automatically change if you use your mouse to click on another value. You can change that value as often as you want to, within the given time limit, by moving and clicking the mouse. During the waiting time, you can already click with the mouse, but only after the waiting time is over the box appears and shows a clicked value.

At the top of the screen, you see a timer with the remaining time for the current period. To submit your prediction, you can **either click the "Submit" button with the mouse or press "Enter" on the keyboard before the time for the period is up**. If the time is up and you have not pushed the "Submit" button or pressed "Enter" yet, you made **no prediction** for this period. Although using the mouse is the fastest way to submit a prediction, you can also directly type your forecast in the box.

The price of the stock must be positive. The vertical scale of the graphs does **not** represent an upper bound for prices or predictions: the vertical axis of the graphs will rescale when prices or predictions change. However, the **first two prices in the first phase** are likely to lie between 0 and 200, and the **first two prices in the second phase** are likely to lie between 0 and 100. After the instructions, you will enter a practice round to get familiar with the program.



<u>Differences between the two phases:</u> The experiment consists of two phases. Prices in the second phase will be independent of prices and predictions from the first phase. Your task, predicting the price of the risky asset, will be the same in both phases, but there are also some important differences between the two phases. In particular, the **time limit** for making a decision and the **waiting time** in each period may be different for the two phases, and the **interest rate** and **average dividend payment** might be different between the two phases as well. Just prior to the start of each phase, the time limit and waiting time for each period in that phase and the relevant interest rate and average dividend payment will be announced.

There are two other differences between the two phases. First, the set of pension funds (and therefore, the group of financial advisors/participants) active in your market may not be the same in the two phases. Second, the number of periods might be different between the two phases. In particular, the number of periods is unknown for each phase, but it will be between 120 and 180 periods for each phase.

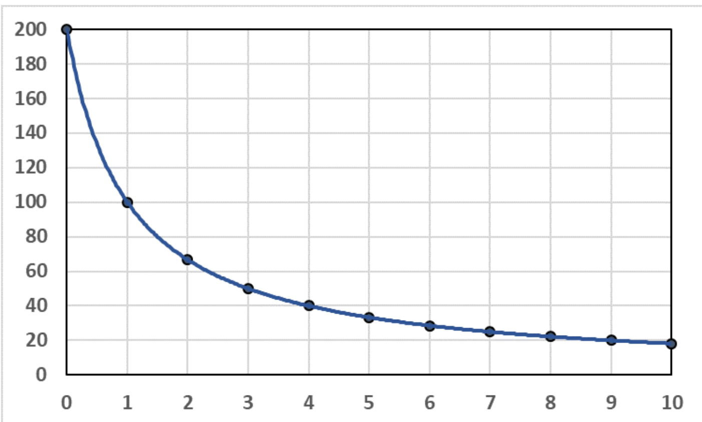
Earnings: Earnings will be fully determined by your forecasting accuracy. At the end of the experiment, 10 periods from the first and 10 periods from the second phase are randomly chosen. You will be paid for your forecasting accuracy in these 20 periods. The better you predict the stock market price in these periods, the higher your earnings will be. Note that if a period is selected in which you did not make a prediction, your earnings for that period will be zero. The number of points you get for a period is given by the following formula:

$$\textbf{Points for period } t = \frac{200}{1 + \textbf{Prediction Error}},$$

where **Prediction Error** is your prediction error in that period (the difference between your forecast for period t and the realized market price p_t for period t). The table below gives the relationship between the prediction error and the number of points for some particular cases:

Prediction Error	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Points	200	100	66.67	50	40	33.33	28.57	25	22.22	20	18.18

The figure below shows the relation between the number of points you score (vertical axis) and the prediction error. Notice that the table presents only some possibilities for your point earnings and that the number of points you earn decreases more slowly as your prediction error increases. The number of points earned in a period that is randomly selected to be paid will be transferred to euros at a rate of



100 points for one euro. For example, suppose for period 16 you predict that the price will be equal to 45, but the actual market price is 48, then your error is 48 – 45 = 3, and you will, therefore, earn 50 points. If period 16 is randomly selected as one of the periods for which you are paid, you will earn €0.50 for that period.

As a second example: suppose you predict a price of 90 for period 60, and the price in that period turns out to be 88.3. Now your prediction error is 90 – 88.3 = 1.7 and you would receive $\frac{200}{1+1.7} = 74.04$ points. This will give you €0.74 if that period is selected for payment. Note that if you are paid for a period in which you predicted the price correctly, you will obtain 200 points, or €2.00 euro, for that period. On the computer screen that you see during the experiment, you can find (under "potential earnings"), for each period, the number of points that you would get if that period is selected for payment. Your total earnings for the experiment will be the earnings from the 20 periods that are randomly selected, plus a participation fee of €10.00.

Appendix 3.B Questionnaire information

Table 3.4: Questionnaire information from 185 participants.

Demographics		**Shares**
Gender		
	Female	54.1%
	Male	45.9%
Mean age (SD), 18-59 yrs		21.96 (4.58)
Mean earnings, 13.09-39.96 €		27.64
Study program		
	Faculty of Econ and Business	65.4%
	Faculty of Law	7.0%
	Faculty of Science, Math, and CompSci	2.7%
	Faculty of Humanities	4.3%
	Faculty of Med/Dentistry	1.6%
	Faculty of BehSci other than Psych	13.0%
	Other	3.2%
Trading experience in financial markets		
	No	79.5%
	Yes, a little bit	17.8%
	Yes, I frequently trade	2.7%
Enough time to make a decision		
	No, I felt time pressure in some of the periods in HTP, but not in LTP	30.3%
	No, I felt time pressure in most of the periods in HTP, but not in LTP	40.5%
	No, I felt time pressure in almost all of the periods in HTP, and in some periods in LTP	8.6%
	No, I felt time pressure in most of the periods in both blocks	2.7%
	Yes, in all periods	17.8%

Appendix 3.C Merging data

Table 3.5: p-values of the Kolmogorov-Smirnov test comparing Interquartile Range (IQR), Standard Deviation (SD), Median of Relative Absolute Deviation (RAD), pairwise for different treatments. Data for 1-50, 1-145, and 86-145 periods are used.

IQR, 1-50	**AH**	**AHS**	**AL**	*IQR, 1-145*	**AH**	**AHS**	**AL**	*IQR, 86-145*	**AH**	**AHS**	**AL**
AH	X	0.7475	0.0001***	**AH**	X	0.5732	0.1582	**AH**	X	0.8283	0.1902
AHS		X	0.0000***	**AHS**		X	0.0102**	**AHS**		X	0.5246
AL			X	**AL**			X	**AL**			X
SD, 1-50	**AH**	**AHS**	**AL**	*SD, 1-145*	**AH**	**AHS**	**AL**	*SD, 86-145*	**AH**	**AHS**	**AL**
AH	X	0.1617	0.0017***	**AH**	X	0.9252	0.0076***	**AH**	X	0.8283	0.0706*
AHS		X	0.0002***	**AHS**		X	0.0102**	**AHS**		X	0.3306
AL			X	**AL**			X	**AL**			X
RAD, 1-50	**AH**	**AHS**	**AL**	*RAD, 1-145*	**AH**	**AHS**	**AL**	*RAD, 86-145*	**AH**	**AHS**	**AL**
AH	X	0.5305	0.0043***	**AH**	X	0.4892	0.0032***	**AH**	X	0.8971	0.0452**
AHS		X	0.0009***	**AHS**		X	0.0009***	**AHS**		X	0.1915
AL			X	**AL**			X	**AL**			X

Table 3.6: p-values of the Fischer-Pitman permutation test comparing Interquartile Range (IQR), Standard Deviation (SD), Median of Relative Absolute Deviation (RAD), pairwise for different treatments. Data for 1-50, 1-145, 86-145 periods are used.

IQR, 1-50	**AH**	**AHS**	**AL**	*IQR, 1-145*	**AH**	**AHS**	**AL**	*IQR, 86-145*	**AH**	**AHS**	**AL**
AH	X	0.2489	0.0003***	**AH**	X	0.3022	0.0429**	**AH**	X	0.9163	0.7871
AHS		X	0.0004***	**AHS**		X	0.0058***	**AHS**		X	0.8706
AL			X	**AL**			X	**AL**			X
SD, 1-50	**AH**	**AHS**	**AL**	*SD, 1-145*	**AH**	**AHS**	**AL**	*SD, 86-145*	**AH**	**AHS**	**AL**
AH	X	0.0650*	0.0008***	**AH**	X	0.6836	0.0037***	**AH**	X	0.8187	0.2437
AHS		X	0.0006***	**AHS**		X	0.0026***	**AHS**		X	0.3241
AL			X	**AL**			X	**AL**			X
RAD, 1-50	**AH**	**AHS**	**AL**	*RAD, 1-145*	**AH**	**AHS**	**AL**	*RAD, 86-145*	**AH**	**AHS**	**AL**
AH	X	0.2622	0.0011***	**AH**	X	0.3509	0.0015***	**AH**	X	0.8709	0.0749*
AHS		X	0.0005***	**AHS**		X	0.0005***	**AHS**		X	0.0606*
AL			X	**AL**			X	**AL**			X

Table 3.7: p-values of the Mann-Whitney-Wilcoxon test test comparing Interquartile Range (IQR), Standard Deviation (SD), Median of Relative Absolute Deviation (RAD), pairwise for different treatments. Data for 1-50, 1-145, 86-145 periods are used.

IQR, 1-50	**AH**	**AHS**	**AL**	*IQR, 1-145*	**AH**	**AHS**	**AL**	*IQR, 86-145*	**AH**	**AHS**	**AL**
AH	X	0.4470	0.0001***	**AH**	X	0.4002	0.0518*	**AH**	X	0.9682	0.2225
AHS		X	0.0001***	**AHS**		X	0.0077***	**AHS**		X	0.8036
AL			X	**AL**			X	**AL**			X
SD, 1-50	**AH**	**AHS**	**AL**	*SD, 1-145*	**AH**	**AHS**	**AL**	*SD, 86-145*	**AH**	**AHS**	**AL**
AH	X	0.0789*	0.0017***	**AH**	X	0.5490	0.0033***	**AH**	X	0.9048	0.0927*
AHS		X	0.0003***	**AHS**		X	0.0020***	**AHS**		X	0.2707
AL			X	**AL**			X	**AL**			X
RAD, 1-50	**AH**	**AHS**	**AL**	*RAD, 1-145*	**AH**	**AHS**	**AL**	*RAD, 86-145*	**AH**	**AHS**	**AL**
AH	X	0.2775	0.0014***	**AH**	X	0.4002	0.0022***	**AH**	X	0.9048	0.0698*
AHS		X	0.0004***	**AHS**		X	0.0006***	**AHS**		X	0.0428**
AL			X	**AL**			X	**AL**			X

Appendix 3.D Descriptive statistics and measures

Table 3.8: Descriptive statistics in LTP treatment for the 3 phases and all data.

	interquartile range				standard deviation				RAD (median)				B	C
AL	I	II	III	1-145	I	II	III	1-145	I	II	III	1-145		
1	95.13	63.85	12.98	52.33	52.30	55.40	12.75	44.59	0.35	0.40	0.25	0.30	1	4
2	112.10	12.65	50.82	42.80	100.82	14.00	49.73	61.69	0.49	0.30	0.19	0.28	1	2
3	382.51	21.85	41.75	119.05	387.86	14.60	33.21	253.97	1.65	0.50	0.51	0.53	2	1
4	264.41	61.69	4.85	97.25	195.45	67.11	4.31	140.24	1.65	0.30	0.08	0.28	2	2
5	334.23	238.22	20.01	141.54	204.49	370.17	23.07	245.34	0.62	0.63	0.20	0.35	3	2
6	245.96	75.89	10.03	152.97	227.12	48.32	6.60	170.49	1.77	0.71	0.05	0.55	3	4
7	139.54	7.27	7.44	25.37	72.61	14.21	5.20	42.90	0.53	0.17	0.04	0.17	1	23
8	68.42	3.95	41.64	14.39	41.77	2.63	27.88	34.02	0.86	0.90	0.97	0.90	1	0
9	340.97	124.11	52.05	132.94	183.37	88.54	39.07	121.20	1.07	0.87	0.52	0.71	2	1
10	159.67	66.66	43.14	82.60	105.83	43.42	289.50	239.98	0.74	0.40	0.58	0.49	3	3
11	425.36	52.04	19.40	80.51	249.89	55.84	24.09	163.54	1.33	0.65	0.40	0.53	2	2
12	157.32	822.14	183.22	275.52	113.53	688.91	178.85	472.68	0.47	1.97	1.30	1.00	3	1
av	227.14	129.19	40.61	101.44	161.25	121.93	57.86	165.89	0.96	0.65	0.42	0.51	2.0	3.8
md	202.82	62.77	30.83	89.92	148.45	51.86	25.99	151.89	0.80	0.56	0.32	0.51	2.0	2.0

Table 3.9: Descriptive statistics in HTP treatments for the 3 phases and all data.

	interquartile range				standard deviation				RAD (median)				B	C
AH	I	II	III	1-145	I	II	III	1-145	I	II	III	1-145		
1	13.52	19.25	44.60	35.33	24.71	29.13	33.04	32.45	0.18	0.09	0.15	0.15	1	4
2	36.18	9.36	3.20	47.01	26.13	6.50	1.63	42.91	0.49	0.11	0.00	0.12	1	55
3	3.27	0.72	0.66	3.60	2.34	0.43	0.44	5.74	0.02	0.02	0.01	0.02	0	127
4	24.05	2.75	0.52	4.02	18.31	3.57	5.02	14.49	0.09	0.02	0.01	0.02	0	58
5	67.02	88.42	219.85	123.48	71.81	57.01	111.26	88.26	0.41	0.34	0.65	0.42	1	1
6	115.82	118.06	108.17	103.65	86.59	70.02	65.45	71.33	0.36	0.44	0.57	0.38	1	1
7	29.91	4.67	3.22	16.57	39.70	10.63	6.51	34.79	0.11	0.18	0.17	0.17	1	3
8	29.44	25.98	59.55	33.49	17.20	15.92	34.01	24.92	0.28	0.12	0.20	0.20	1	6
9	57.73	26.65	11.17	89.44	65.50	35.28	9.66	70.68	0.60	0.15	0.18	0.23	1	7
10	12.33	6.41	19.63	18.82	10.13	5.25	12.41	14.32	0.20	0.14	0.08	0.13	0	10
av	38.93	30.23	47.06	47.54	36.24	23.37	27.94	39.99	0.27	0.16	0.20	0.18	0.7	27.2
md	29.67	14.31	15.40	34.41	25.42	13.27	11.04	33.62	0.24	0.13	0.16	0.16	1.0	6.5

	interquartile range				standard deviation				RAD (median)				B	C
AHS	I	II	III	1-145	I	II	III	1-145	I	II	III	1-145		
1	4.60	0.61	0.23	0.91	2.53	2.30	0.15	2.94	0.01	0.00	0.01	0.01	0	72
2	31.50	56.21	19.06	41.21	16.90	44.85	12.71	29.72	0.34	0.05	0.11	0.22	1	14
3	6.26	0.14	0.19	2.22	4.59	0.14	0.28	6.63	0.03	0.01	0.01	0.01	0	124
4	2.24	0.25	0.41	0.44	2.06	0.48	0.22	2.79	0.02	0.02	0.02	0.02	0	139
5	90.44	21.88	21.89	26.48	66.97	11.60	15.55	44.45	0.30	0.08	0.08	0.10	1	13
6	33.36	75.00	89.12	69.26	20.41	39.10	86.89	61.76	0.36	0.23	0.49	0.29	1	3
7	34.53	12.73	33.98	29.88	18.76	8.44	17.92	21.80	0.33	0.05	0.13	0.10	1	18
8	41.50	32.60	155.62	62.66	21.72	61.63	81.86	103.70	0.14	0.21	0.43	0.28	2	5
9	16.56	47.17	87.39	37.12	10.37	23.68	50.81	33.89	0.09	0.19	0.31	0.15	1	10
av	29.00	27.40	45.32	30.02	18.26	21.36	29.60	34.19	0.18	0.09	0.18	0.13	0.8	44.2
md	31.50	21.88	21.89	29.88	16.90	11.60	15.55	29.72	0.14	0.05	0.11	0.10	1.0	14.0

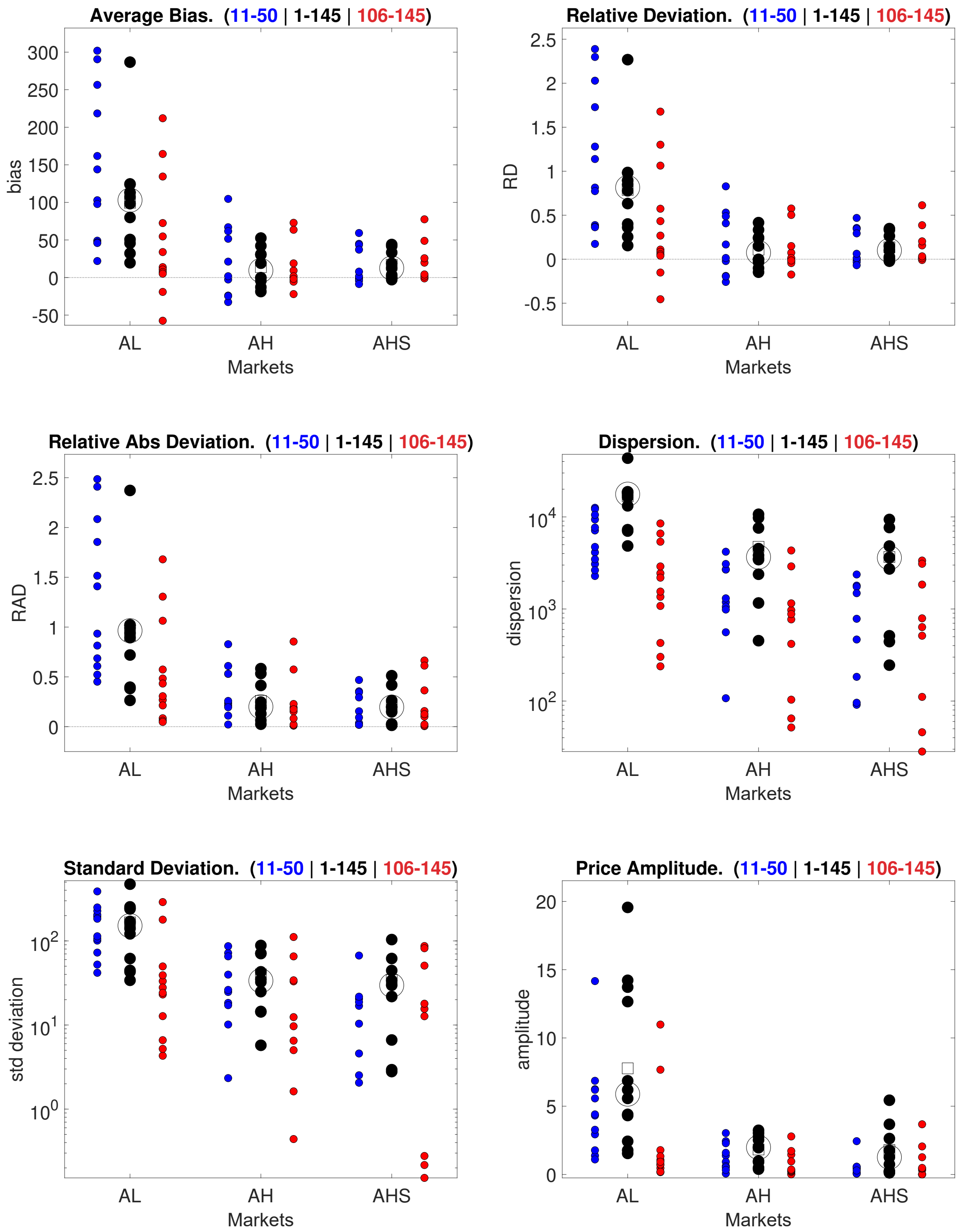


Figure 3.10: Additional measures: price amplitude: $\max_{t=1}^{T}\{(p_t - p^f)/p^f\} - \min_{t=1}^{T}\{(p_t - p^f)/p^f\}$; average bias: $\sum_{t=1}^{T}(p_t - p^f)/T$; mean relative deviation: $(\sum_{t=1}^{T}(p_t - p^f)/p^f)/T$; mean relative absolute deviation: $(\sum_{t=1}^{T}|p_t - p^f|/p^f)/T$; dispersion: $\sum_{t=1}^{T}|p_t - p^f|$. See e.g., Stöckl et al. (2010).

Appendix 3.E Price and predictions

3.E.1 AL markets

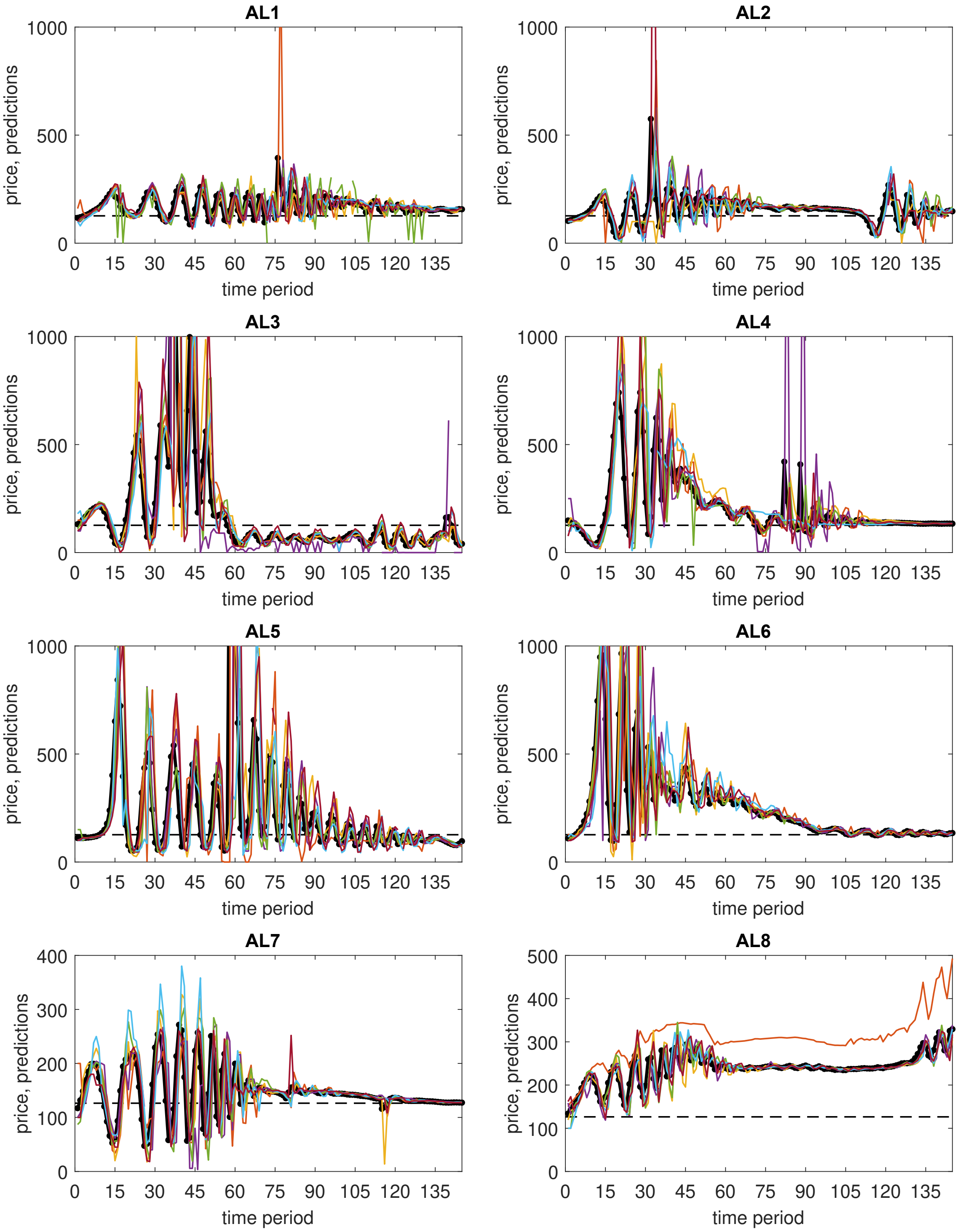


Figure 3.11: Prices and predictions in LTP markets AL1-8

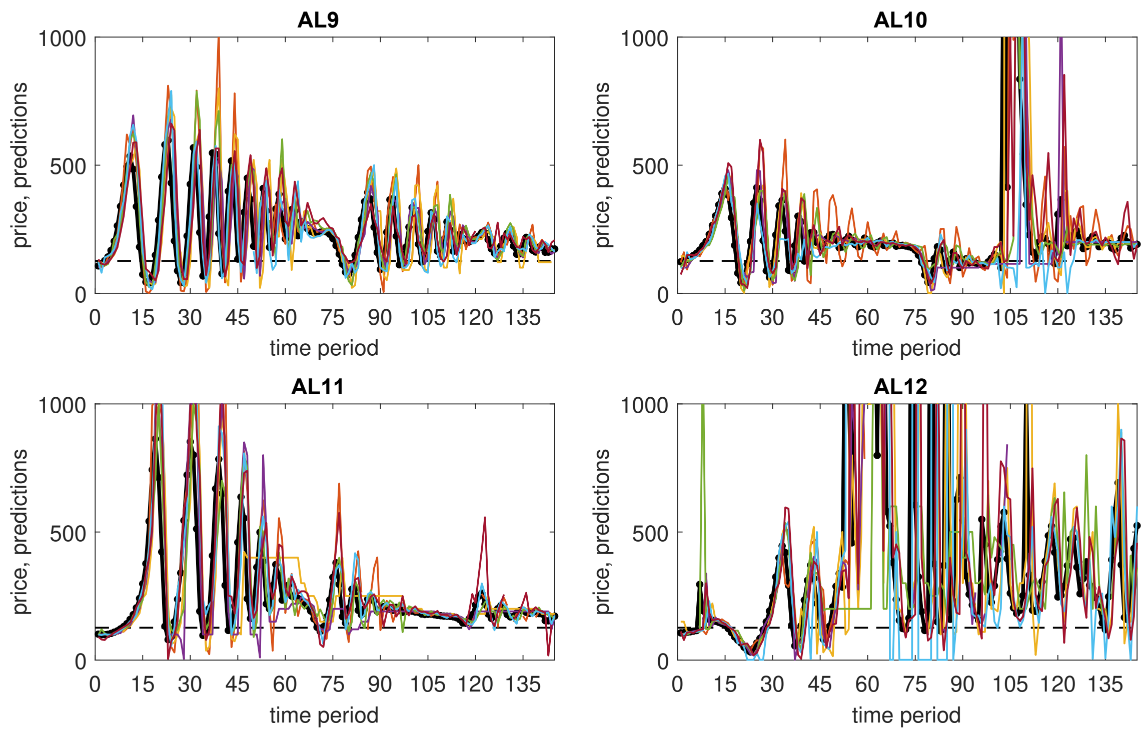


Figure 3.12: Prices and predictions in LTP markets AL9-12

3.E.2 AH markets

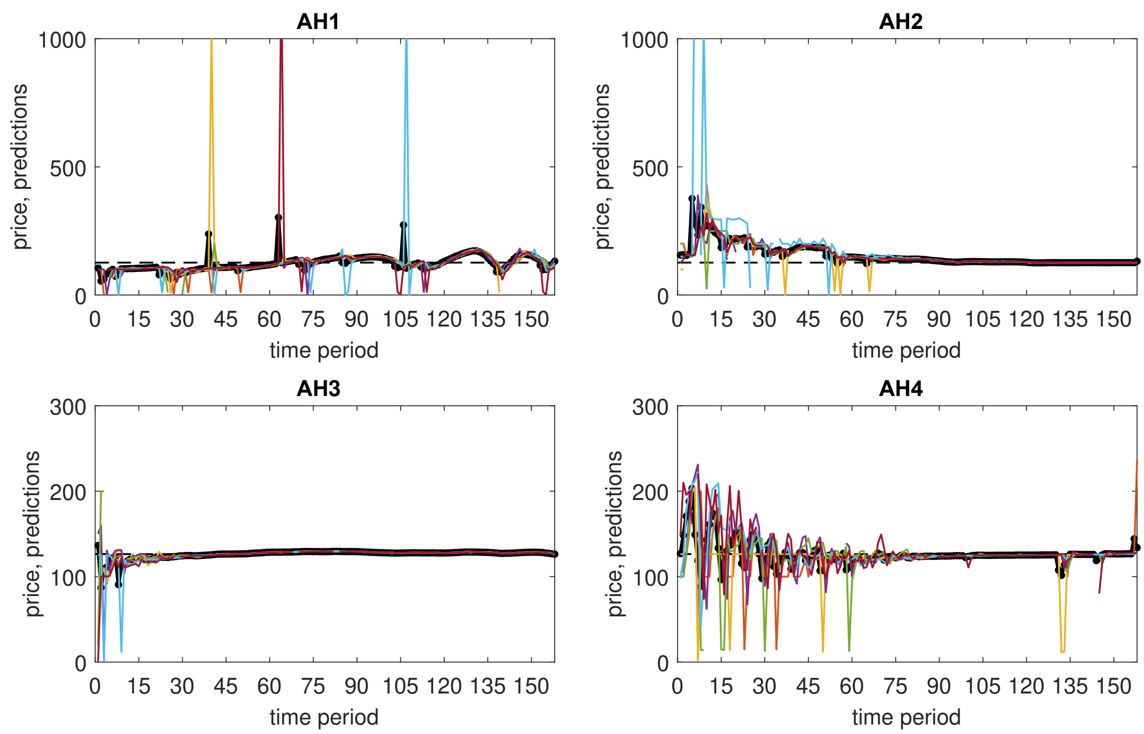


Figure 3.13: Prices and predictions in HTP markets AH1-4

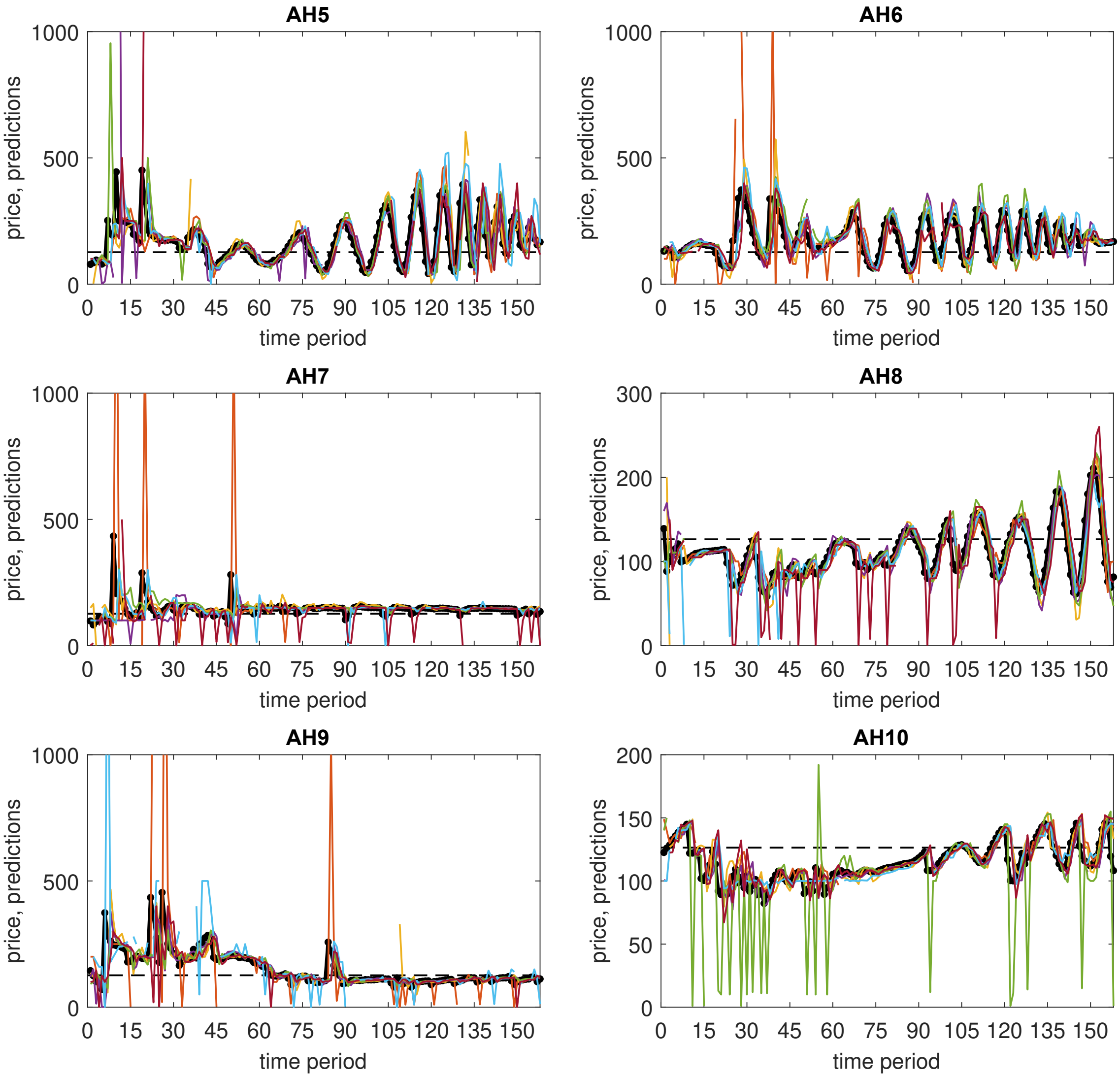


Figure 3.14: Prices and predictions in HTP markets AH5-10

3.E.3 AHS markets

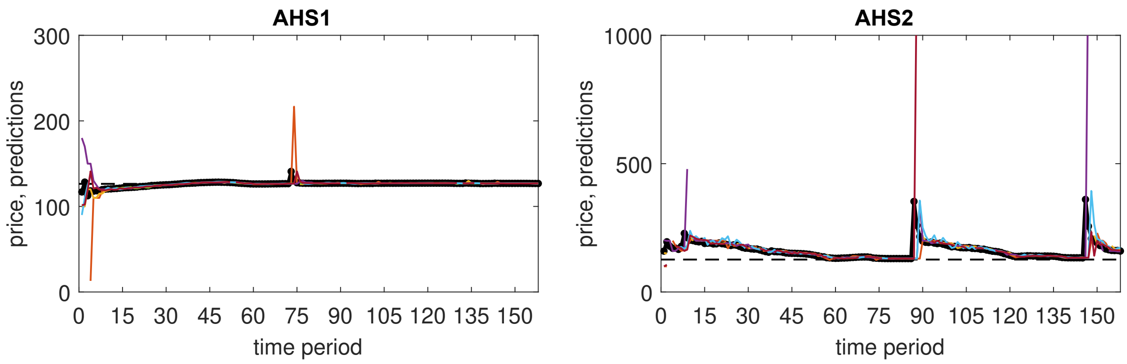


Figure 3.15: Prices and predictions in HTP markets AHS1-2

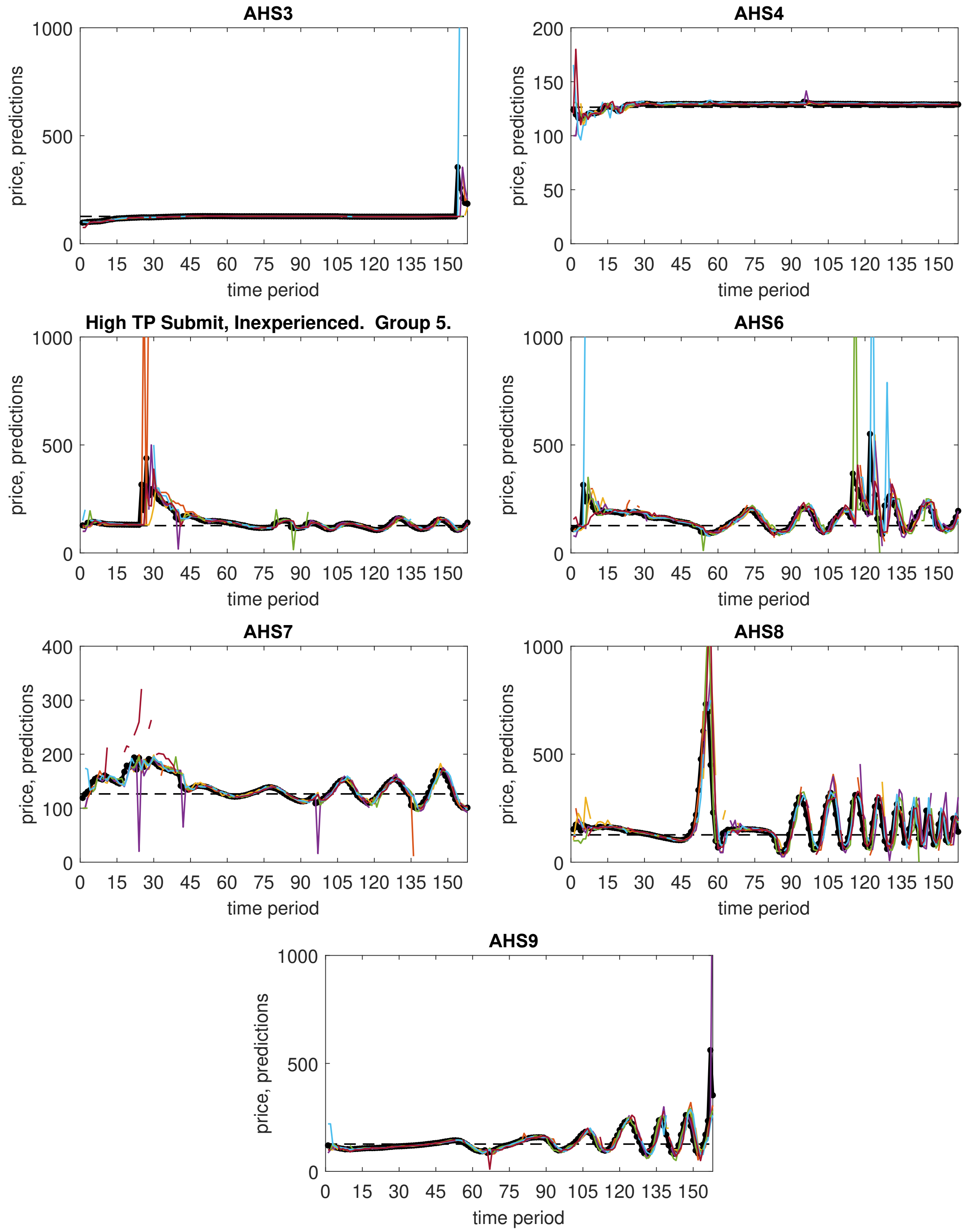


Figure 3.16: Prices and predictions in HTP markets AHS3-9

3.E.4 Coordination

Table 3.10: p-values of the Mann-Whitney-Wilcoxon test comparing the mean of standard deviations for various initial intervals, pairwise for different conditions.

3 periods	**LTP**	**HTP**	*5 periods*	**LTP**	**HTP**	*10 periods*	**LTP**	**HTP**
LTP	X	.0406**	**LTP**	X	.010**	**LTP**	X	.092*
HTP		X	**HTP**		X	**HTP**		X
25 periods	**LTP**	**HTP**	*50 periods*	**LTP**	**HTP**	*145 periods*	**LTP**	**HTP**
LTP	X	.033**	**LTP**	X	.001***	**LTP**	X	.000***
HTP		X	**HTP**		X	**HTP**		X

Note: LTP represents the AL markets. HTP corresponds to the weighted average over the AH and AHS markets.

Appendix 3.F Market expectations

Table 3.11: Market forecasting rules between treatments with 4 price lags.

AL

Gr	Phase	Const	Past Prices lag 1	lag 2	lag 3	lag 4	LB autoc	Obs
1	**1-50**	29.42	1.94	-	−1.27	0.19	0.38	46
	96-145	-	0.54	-	-	0.35	0.39	46
2	**1-50**	93.98	1.04	-	−0.57	-	1.00	46
	96-145	-	1.92	0.55	−1.63	-	0.68	46
3	**1-50**	-	2.24	0.59	−1.53	-	1.00	46
	96-145	21.18	1.40	-	−0.76	-	1.00	46
4	**1-50**	-	2.21	0.68	−1.81	-	0.66	46
	96-145	29.27	0.92	-	−0.44	0.31	0.66	46
5	**1-50**	-	2.56	1.27	−2.49	−0.33	0.85	46
	96-145	-	1.25	-	−0.82	0.35	0.01	46
6	**1-50**	-	2.07	0.56	−1.58	-	0.89	46
	96-145	-	1.72	-	−1.08	0.22	0.26	46
7	**1-50**	38.29	1.90	0.49	−1.62	-	0.39	46
	96-145	-	0.88	-	-	-	1.00	46
8	**1-50**	-	1.76	0.74	−1.52	-	0.17	46
	96-145	−24.73	2.55	2.44	−3.12	−0.71	0.40	46
9	1-50	86.42	2.46	1.84	−2.85	−0.65	0.05	46
	96-145	76.54	1.49	0.47	−1.31	-	0.72	46
10	**1-50**	49.24	1.82	0.44	−1.44	-	0.81	46
	96-145	-	0.72	-	0.84	−0.39	0.94	46
11	**1-50**	-	2.37	0.58	−1.86	-	0.33	46
	96-145	61.95	1.48	-	−0.80	-	0.39	46
12	**1-50**	33.10	1.53	-	−0.67	-	1.00	46
	96-145	392.12	0.42	−0.44	-	-	0.93	46

AH

Gr	Phase	Const	Past Prices lag 1	lag 2	lag 3	lag 4	LB autoc	Obs
1	**1-50**	99.71	-	-	-	-	0.25	46
	96-145	85.83	0.36	-	-	-	0.95	46
2	**1-50**	-	0.50	0.33	-	-	0.95	46
	96-145	-	1.80	-	−0.78	-	0.70	46
3	**1-50**	65.21	0.47	-	-	-	1.00	46
	96-145	9.49	1.41	−0.48	-	-	0.61	46
4	**1-50**	73.03	0.97	-	−0.53	-	0.10	46
	96-145	59.47	1.04	-	−0.52	-	0.94	46
5	**1-50**	68.04	0.66	-	-	-	0.66	46
	96-145	60.44	1.89	0.43	−1.57	-	0.12	46
6	**1-50**	51.71	1.17	−0.42	-	-	0.57	46
	96-145	111.60	1.43	-	−0.98	-	0.02	46
7	**1-50**	153.07	-	-	-	-	0.76	46
	96-145	78.68	0.81	-	−0.35	-	0.66	46
8	**1-50**	-	1.47	-	−0.62	-	0.81	46
	96-145	21.84	1.84	-	−1.02	-	0.95	46
9	**1-50**	145.97	0.43	-	-	-	0.09	46
	96-145	53.32	0.48	-	-	-	0.44	46
10	**1-50**	-	0.90	-	-	-	0.96	46
	96-145	28.32	1.53	-	−0.75	-	0.69	46

AHS

Gr	Phase	Const	Past Prices lag 1	lag 2	lag 3	lag 4	LB autoc	Obs
1	**1-50**	-	0.86	-	0.15	-	0.72	46
	96-145	35.04	0.72	-	-	-	0.88	46
2	**1-50**	-	1.01	-	-	-	1.00	46
	96-145	−7.21	1.21	-	-	−0.15	0.27	46
3	**1-50**	−3.66	1.60	-	−0.57	-	0.84	46
	96-145	-	1.15	-	-	−0.17	0.91	46
4	**1-50**	-	1.29	−0.66	-	0.36	0.15	46
	96-145	-	1.02	-	-	-	1.00	46
5	**1-50**	-	0.32	-	0.96	−0.41	0.96	46
	96-145	8.44	1.98	-	−1.04	-	0.57	46
6	**1-50**	66.51	0.65	-	-	-	0.96	46
	96-145	64.20	0.71	-	-	-	0.70	46
7	**1-50**	-	0.96	-	-	-	0.95	46
	96-145	22.02	1.25	-	-	−0.41	0.95	46
8	**1-50**	-	1.60	-	-	−0.61	0.79	46
	96-145	73.08	1.61	-	−1.00	-	0.18	46
9	**1-50**	−7.40	1.48	−0.42	-	-	0.54	46
	96-145	24.76	1.90	-	−1.04	-	0.67	46

Note: Phases are marked bold if Ljung-Box (LB) test finds no significant autocorrelation in the data (p>.05).

Table 3.12: Market forecasting rules between treatments with 2 price lags.

AL

Group	Phase	Const	Past Prices lag 1	Past Prices lag 2	LB autoc	Obs
1	**1-50**	52.73	1.71	−1.00	0.23	48
	96-145	50.08	0.71	-	0.01	48
2	**1-50**	90.58	1.05	−0.57	1.00	48
	96-145	59.35	1.47	−0.88	0.02	48
3	**1-50**	-	1.95	−0.87	0.70	48
	96-145	20.36	1.40	−0.76	1.00	48
4	1-50	82.11	1.72	−0.87	0.02	48
	96-145	75.70	0.88	−0.41	0.05	48
5	**1-50**	64.35	1.86	−1.01	0.06	48
	96-145	84.63	0.95	−0.72	0.00	48
6	**1-50**	143.84	1.67	−0.88	0.16	48
	96-145	40.35	1.53	−0.83	0.18	48
7	1-50	76.30	1.48	−0.95	0.05	48
	96-145	-	0.90	-	1.00	48
8	1-50	56.56	1.36	−0.57	0.00	48
	96-145	-	1.60	−0.49	0.00	48
9	1-50	144.42	1.53	−0.95	0.00	48
	96-145	157.84	1.10	−0.85	0.00	48
10	**1-50**	88.18	1.48	−0.85	0.31	48
	96-145	-	0.74	0.57	0.99	48
11	**1-50**	87.67	1.89	−0.99	0.20	48
	96-145	60.55	1.46	−0.77	0.27	48
12	**1-50**	31.68	1.53	−0.67	1.00	48
	96-145	297.42	0.60	−0.36	0.50	48

AH

Group	Phase	Const	Past Prices lag 1	Past Prices lag 2	LB autoc	Obs
1	**1-50**	67.66	0.31	-	0.92	48
	96-145	85.56	0.37	-	0.93	48
2	**1-50**	73.67	0.68	-	0.52	48
	96-145	-	1.86	−0.84	0.62	48
3	**1-50**	65.14	0.47	-	1.00	48
	96-145	-	1.81	−0.85	0.12	48
4	**1-50**	67.91	1.06	−0.58	0.31	48
	96-145	59.54	1.04	−0.52	0.93	48
5	**1-50**	62.89	0.68	-	0.68	48
	96-145	95.62	1.52	−0.97	0.00	48
6	1-50	-	1.37	−0.51	0.05	48
	96-145	107.40	1.44	−0.97	0.01	48
7	**1-50**	150.98	-	-	0.56	48
	96-145	75.72	0.82	−0.34	0.57	48
8	**1-50**	-	1.15	−0.27	0.23	48
	96-145	21.57	1.83	−1.02	0.88	48
9	**1-50**	123.47	0.51	-	0.10	48
	96-145	53.03	0.49	-	0.39	48
10	**1-50**	-	0.92	-	0.97	48
	96-145	27.28	1.52	−0.74	0.68	48

AHS

Group	Phase	Const	Past Prices lag 1	Past Prices lag 2	LB autoc	Obs
1	**1-50**	-	0.35	0.66	0.99	48
	96-145	33.34	0.74	-	0.88	48
2	**1-50**	-	0.98	-	0.99	48
	96-145	−5.08	1.04	-	0.00	48
3	**1-50**	−3.60	1.49	−0.46	0.86	48
	96-145	-	0.99	-	0.73	48
4	**1-50**	-	1.56	−0.59	0.14	48
	96-145	-	1.49	−0.49	0.86	48
5	**1-50**	-	0.30	0.65	0.77	48
	96-145	8.47	1.98	−1.05	0.49	48
6	**1-50**	58.18	0.69	-	1.00	48
	96-145	63.03	0.71	-	0.66	48
7	**1-50**	-	0.96	-	0.93	48
	96-145	-	1.86	−0.91	0.50	48
8	**1-50**	-	1.00	-	0.99	48
	96-145	73.08	1.62	−1.01	0.25	48
9	1-50	−8.26	1.81	−0.74	0.00	48
	96-145	23.10	1.90	−1.03	0.63	48

Note: Phases are marked bold if Ljung-Box (LB) test finds no significant autocorrelation in the data (p>.05).

Chapter 4

On subscription traps and preference reversals - the pigeonholing effect

4.1 Introduction

Dominance rule is obeyed when its application is transparent, but it can be masked by a frame in which the inferior option yields a more favorable outcome in an identified state of the world.

— Tversky and Kahneman (1986)

In the literature on decision-making under risk and uncertainty, consensus prevails that people follow the compelling principles of rational choice theory only if a problem is presented sufficiently transparent. However, facing disguised problems, decision-makers (DM) can lose orientation and violate rational paradigms they would not otherwise (e.g., Tversky and Kahneman, 1986). In consequence, DMs may select dominated choices or commit preference reversals.

This paper reports a novel phenomenon in the domain of *context-dependent* preference reversals.[50] Subjects face the hypothetical choice between two subscriptions of a service, for which the provided context conveys unlikely follow-up use (see Problem 1). In a between-subject design, the treatment condition poses a short-term subscription A vs. a costlier but longer alternative B. The control condition contrasts

[50]See Bleichrodt et al. (2019) for a recent discourse on the importance of reference-dependent generalizations of standard decision theories on the basis of Rabin (2000) paradox.

B against A′, a truncated version of A at the price of A. In a classroom experiment ran in late February 2020, a few days before the first COVID lock-down closed the university, we find that most students select option B in the treatment group but the between-treatment dominated option A′ in the control.

Problem 1 Imagine you are visiting New York City over the Christmas break. As last item on your agenda before flying back home, you have saved the panorama view from the Rockefeller Center. At the entrance, you have to register with your ID card and choose between two alternatives.

Treatment condition (23 subjects, shares in preferences are reported in brackets)

A: 6 months free entrance corresponding with registered ID for 5.00$	[26.1%]
B: 18 months free entrance corresponding with registered ID for 5.50$	[73.9%]

Control condition (18 subjects)

A′: Single entry corresponding with registered ID card for 5.00$	[66.7%]
B: 18 months free entrance corresponding with registered ID for 5.50$	[33.3%]

The varying subscription length of the short option between treatments (A vs. A′) may trigger distinct cognitive processes causing a biased likelihood assessment of using the subscription again. While the nature of risk and ambiguity forms per se the breeding ground for cognitive biases in decision-making[51], distinct signaling across the two choice menus may be a plausible explanation for the present paradox.

4.1.1 Availability heuristic

There is evidence that individuals evaluate the frequency of uncertain events by the ease with which past instances are retrievable to a judge's memory, the so-called *availability heuristic* (Tversky and Kahneman, 1973).

Analogously in Problem 1, the momentary ticket demand could lead a person to overestimate the general probability of visiting the Rockefeller building. Overweighting instances that have occurred more recently to a person can explain choosing B over A, even though the rational value of B is inferior to most people since it is improbable for a regular person to visit New York twice within 18 months. Yet, distorted probabilities cannot explain reversed preferences after the utility of the short option shrank (A′). It must be that something suppresses overweighting, e.g., caused by the availability heuristic, and instead puts the decision into objective context. Following Tversky and Kahneman (1973), there is reason to believe that certain

[51] The canonical examples in behavioral economics demonstrating the effect of risk, respectively ambiguity, are the Allais (1953) paradox and the no less compelling Ellsberg (1961) paradox. They commence an exhaustive review of related preference reversals in Section 4.5.

factors feed the inclination to overrate events that occurred more recently to a DM. It may be that the nature of time spans, e.g., six months, emphasize probability overweighting. Instead, the presence of single-use works like a counterspell against probability distortion and for rational cost-value analysis.

Another hypothesis sings the same tune in a different voice. A preference reversal (PR) is typically characterized by varying prominence of one or more decision aspects (see Section 4.5). The introduced PR may accentuate certain stimuli (e.g., marginal value, regret[52]) when time spans are compared to each other but deems them irrelevant as soon as single-use is available. Suppose in the absence of a single-use option, the mere duration of a 6-month subscription blurs its utility advantage over an 18-month one by drawing a DM's attention towards other less rational criteria. Hence, a DM may start comparing what Thaler (1985) coined *transaction utility*.

4.1.2 Transaction utility

The concept of transaction utility relates to the perceived fairness of a deal. Albeit the personally attached value of a good exceeds its asked price, a DM may still reject the deal, let alone carry away a bittersweet aftertaste from a successful transaction, if she associates a much lower reference price with the good (e.g., if she learns that the seller acquired it herself for much cheaper). Thus, Thaler (1985) describes the total value of a transaction by the sum of its parts: objective *acquisition utility* and subjective transaction utility. He quotes a two-night-minimum for hotel bookings on Super Bowl weekend as an example for obscuring market values. Demand typically balloons on Super Bowl night, and the market-clearing price would exceed 50% of the full weekend price by far. Compared to box fight promoters who charge incredibly high prices for blockbuster fights mirroring the true market value, hotel chains capitalize on their corporate image through a loyal customer base. Setting prices at the theoretical market value level would send a greedy signal and harm their brand value in the long run. Instead, hotel chains dress the high price on Super Bowl Sunday in an acceptable two-night bundle with the night before.

4.1.3 Pigeonholing hypothesis

Depending by definition on a reference point, the notions of transaction utility and of the availability heuristic are both related to the realm of similarity theories (e.g., Gilboa et al., 2002; Rubinstein, 1988; Tversky, 1969). Following similarity consid-

[52]See Section 4.3.7 for a brief discourse on the relation to regret theory.

erations, the idea of *pigeonholing*[53](PH) advocates "categorical thinking" whenever a DM processes the corresponding information between alternatives on a choice menu. Whether a DM interprets a subscription as a point in time or instead as time period deems one or the other utility dimension more salient. The PH hypothesis attributes an influential role to the length of the short subscription on the choice menu: the briefer the subscription, the greater the salience of acquisition utility, representing the rational dimension in an assessment. Given further use is unlikely to the degree that the short subscription yields a higher objective value, its potentially low transaction utility may be substantially dismissed. The effect size reaches capacity for single-use framed subscriptions. By excluding any risk, they grant the most straightforward rational cost-value assessment. Its presence on the choice menu nudges a DM to devote increased attention to rational criteria (expected utility). Conscious about the high chance to consume the particular good once only, a DM may choose the presumably superior single-use option.

Yet, in the absence of single-use, the comparability of transaction utility between the two choices grows with the length of the short subscription. The human mind craves an easily quantifiable alternative to the rational cost-value assessment. Attention shifts towards the next obvious criterion. In view of transaction utility, the marginal subscription value of a lottery (i.e., duration over price) could resemble such a benchmark in the present context. Correspondingly, a substantially longer subscription, marginally higher priced, outshines the shorter subscription with ease. Similar to how different elicitation procedures highlight distinct stimuli and thereby trigger particular evaluation heuristics (Tversky et al., 1988), it may be that varying subscription lengths accentuate distinct criteria leading to diverging evaluation.

The PH effect postulates that under specific conditions, the shorter alternative in a choice problem, once extended at no extra costs, may become *less* attractive in terms of preference shares or willingness-to-pay. The PR found in Problem 1 complies with this hypothesis under the premise that an extended subscription (say six months) at no extra costs is preferred to its truncated pendant (say single-use). Unless mentioned otherwise, this assumption holds for the remainder of this paper.

This work explores the PH heuristic in a series of experimental studies. The next section presents the findings from a pilot study, including Problem 1. Whereas the nature of lotteries lets most PRs described in the literature evolve along two dimensions (payoff and probability), the additional stimulus (length of the short option) obfuscates subjects in a manner that has yet not been reported to the best of our

[53]Negatively connoted, the term pigeonholing refers to systematizing an object into classes.

knowledge. In Section 4.3, we test our results against leading models in choice under risk. Their two-dimensional framework limits most theories (e.g., EUT; prospect theory, Tversky and Kahneman, 1979) in explaining the observed PR, yet regret theory (Bell, 1982; Loomes and Sugden, 1982) and configural weight theory (Birnbaum et al., 1992) provide in parts an applicable intuition of our paradox. Thereupon, we present a model that extends salience theory (Bordalo et al., 2012) by a third dimension to accommodate the PH effect. Section 4.4 reports additional studies that validate the model predictions experimentally. Albeit the data, totaling 687 independent observations, find overall mixed support for the PH hypothesis, the individual studies document in parts striking evidence of the construed phenomenon. Upon discussing the findings in an extensive review of context-dependent PRs in the literature, Section 4.5 concludes with insights for practitioners and an outlook on follow-up research.

4.2 Pilot study

We conducted a classroom experiment with 88 students from the economic faculty at the Ca' Foscari University of Venice in February 2020.[54]

4.2.1 Method

After entering the lecture hall, we briefed students on the role of experiments in studying human decision-making. We kindly asked for their support in participating in a paper-and-pen experiment and assured them that there was no possibility to trace back choices to subject identities. Furthermore, we notified them that neither were there right or wrong answers, nor was an intelligence test exerted on them. We encouraged the participants to answer the hypothetical questions as if those were real situations they were facing. Next, we handed out the questionnaire asking for age and gender, followed by a set of three choice problems.[55] We chose a between-subject design to control for experimenter demand effects, which represent a common concern to decoy studies (Frydman and Mormann, 2018). Thus, we framed all questions either as short-term vs. long-term (treatment group) or single-use vs. long-term (control group). On top of that, we divided the treatments into two subgroups each, to allow for some initial comparative parameter analysis.

[54]Paolo Pellizzari and Sebastiano Della Lena were kind enough to share precious teaching time literally up to the last second before the campus lock-down due to the pandemic outbreak.

[55]The original questionnaire in Italian can be provided on request.

4.2.2 Results

Problem 1 Imagine you are visiting New York City over the Christmas break. As last item on your agenda before flying back home, you have saved the panorama view from the Rockefeller Center. At the entrance, you have to register with your ID card and choose between two alternatives.

Control subgroup 1 (25 subjects)

A': Single entry corresponding with registered ID card for 5.00$ [80.0%]
B: 3 months free entrance corresponding with registered ID for 5.50$ [20.0%]

Control subgroup 2 (18 subjects)

A': Single entry corresponding with registered ID card for 5.00$ [66.7%]
B: 18 months free entrance corresponding with registered ID for 5.50$ [33.3%]

Although statistically not significant, note the effect of extending option B (3 months $\rightarrow$ 18 months) on the corresponding preferences between control subgroups 1 and 2 (20% vs. 33%). The utility of the short subscription (A') shrinks such that preferences move in the direction predicted by EUT.

Treatment subgroup 1 (22 subjects)

A: 2 weeks of free entrance corresponding with registered ID for 5.00$ [54.5%]
B: 3 months of free entrance corresponding with registered ID for 5.50$ [45.5%]

Treatment subgroup 2 (23 subjects)

A: 6 months free entrance corresponding with registered ID for 5.00$ [26.1%]
B: 18 months free entrance corresponding with registered ID for 5.50$ [73.9%]

We observe the conjectured PR between treatments. Albeit the short subscription gains two weeks, resp. 3 months, free entrance relative to the control, preferences reduce significantly in both treatment subgroups (80% vs. 55%, $t(45)= -1.9$, $p=.032$; 67% vs. 26%, $t(39)= -2.7$, $p=.004$). Note the larger decline in preferences in treatment subgroup 2, in which A and B are extended considerably, compared to subgroup 1.

Problem 2 Imagine you want to book a ride on a ride-sharing platform (e.g., BlaBlaCar). Before you can go ahead and pay the ride to the driver, you have to choose between two utilization fees.

Control subgroup 1 (25 subjects)

A': Single platform use for 3.00 € [56.0%]
B: 10 weeks free platform use for 7.95 € [44.0%]

Control subgroup 2 (18 subjects)

A': Single platform use for 3.00 € [50.0%]
B: 10 weeks free platform use for 4.99 € [50.0%]

Preferences between alternatives are fairly balanced in the control groups of Problem 2. Yet, we observe slight parameter sensitivity. The price for the long-term option B in subgroup 2 decreases by almost 3 €. Thus, a comparably larger proportion of subjects favors the long subscription herein (36% vs. 82%, $t(43)$= -3.5, p<.001).

Treatment subgroup 1 (22 subjects)

A: 1 week free platform use for 3.00 € [63.6%]
B: 10 weeks free platform use for 7.95 € [36.4%]

Treatment subgroup 2 (23 subjects)

A: 1 week free platform use for 3.00 € [17.4%]
B: 10 weeks free platform use for 4.99 € [82.6%]

Contrary to Problem 1, the PR occurs in Problem 2 only between subgroups 2 (50% vs. 17%, $t(39)$= -2.3, p=.013).

Problem 3 Imagine you need a book for your dear friend's birthday by tomorrow. The only chance to get it in time is shipping it via an express delivery service, which offers two alternatives.

Control subgroup 1 (25 subjects)

A′: Single delivery for 4.00 € [44.0%]
B: 8 weeks unlimited delivery service for 5.99 € [56.0%]

Control subgroup 2 (18 subjects)

A′: Single delivery for 6.00 € [33.3%]
B: 8 weeks unlimited delivery service for 9.95 € [66.7%]

In Problem 3, the majority prefers the long-term over the single-use option. Note that price parameters in subgroup 2 lower the attractiveness of single-use compared to subgroup 1 (33% vs. 44%). The difference is statistically not significant, though.

Treatment subgroup 1 (22 subjects)

A: 1 week unlimited delivery service for 4.00 € [22.7%]
B: 8 weeks unlimited delivery service for 5.99 € [77.3%]

Treatment subgroup 2 (23 subjects)

A: 1 week unlimited delivery service for 6.00 € [47.8%]
B: 8 weeks unlimited delivery service for 9.95 € [52.2%]

Preference shifts are similar to those witnessed in Problem 2. Subgroup 1, featuring the smaller price difference, exhibits a PR between treatments. Preferences for the short option halve once single-use replaces the weekly option (44% vs. 23%, $t(45)$= -1.5, p=.065). Yet, the larger price difference in subgroup 2 produces no PR.

Equipped with these insights, we develop a first model based on the PH hypothesis to test the findings from the pilot study against leading theories in the field.

4.3 Model

The present work explores a cognitive bias that causes PRs in certain types of decision problems. These problems are characterized by a choice between two risky prospects (lotteries), given the specific context also named subscriptions: A brief subscription or lottery L_b yielding a small, positive payoff with high probability and a comparably large, negative payoff with low probability. Contrasted against a long subscription L_l, which represents the corresponding shadow lottery paying a small, negative payoff with high probability and a large, positive payoff with low probability. The exact payoff structure is outlined below after formally introducing states and relative payoffs. We introduce the model framework based on Problem 1 with subscriptions' duration $d_b < d_l$ and prices $p_b < p_l$.

Problem 1 Imagine you are visiting New York City over the Christmas break. As last item on your agenda before flying back home, you have saved the panorama view from the Rockefeller Center. At the entrance, you have to register with your ID card and choose between two alternatives.

L_b: d_b free entrance for p_b

L_l: d_l free entrance for p_l

4.3.1 State space and payoffs

The problem discerns three states of the world. A subject can visit the building again within d_b (state 1), after d_b but within d_l (state 2), or not before d_l (state 3).

state 1: *next service use* $\in (0, d_b]$
state 2: *next service use* $\in (d_b, d_l]$
state 3: *next service use* $\in (d_l, \infty)$

State probabilities

We consider a uniformly distributed probability parameter π_d for any given time unit (e.g., day) to utilize the service, which is used at capacity once per time unit.

Thus, state probabilities are given by

$$\pi_s = \begin{cases} \pi_1 = \pi_d + (1-\pi_d)^1\pi_d + ... + (1-\pi_d)^{d_b-2}\pi_d = \pi_d\left(\frac{1-(1-\pi_d)^{d_b-1}}{1-(1-\pi_d)}\right) \\ \quad = 1-(1-\pi_d)^{d_b-1} \\ \pi_2 = (1-\pi_d)^{d_b-1}(\pi_d + (1-\pi_d)^1\pi_d + ... + (1-\pi_d)^{d_l-d_b-1}\pi_d \\ \quad = (1-\pi_d)^{d_b-1} - (1-\pi_d)^{d_l-1} \\ \pi_3 = 1-\pi_1-\pi_2 = (1-\pi_d)^{d_l-1}. \end{cases}$$

Payoffs

If state 2 realizes and a DM has chosen L_b, her subscription has already expired at the time. She then depends on renewal. Equation (1) highlights that her follow-up choice, L_b or L_l, determines payoffs in state 2, whereas payoffs in states 1 and 3 remain unaffected. In line with the majority of participants' feedback,[56] this paragraph treats the follow-up preference L_l as standard case and displays the follow-up-L_b case in brackets only if it matters. Let us denote by x_s^L payoffs in state $s \in \{s_1, s_2, s_3\}$ of lottery $L \in \{L_b, L_l\}$[57]

$$x_s^L = \begin{cases} x_1^{L_b} = -p_b; & x_1^{L_l} = -p_l \\ x_2^{L_b} = -p_b - p_l,\ (-2p_b); & x_2^{L_l} = -p_l \\ x_3^{L_b} = -p_b; & x_3^{L_l} = -p_l. \end{cases} \tag{4.1}$$

With state probabilities and payoffs defined, we can compute the value of a subscription $L \in \{L_b, L_l\}$ by

$$V(L) = \sum_{s\in S} v(x_s^L)\pi_s = \underbrace{v(x_1^L)\cdot\pi_1}_{\text{state 1}} + \underbrace{v(x_2^L)\cdot\pi_2}_{\text{state 2}} + \underbrace{v(x_3^L)\cdot\pi_3}_{\text{state 3}}.$$

Considering identical payoffs in states 1 and 3 for both subscriptions $L \in \{L_b, L_l\}$, we can rewrite

$$V(L) = \sum_{s\in S} v(x_s^L)\pi_s = \underbrace{v(x_1^L)\cdot(\pi_1+\pi_3)}_{\text{state 1 and 3}} + \underbrace{v(x_2^L)\cdot\pi_2}_{\text{state 2}}.$$

[56] Students usually stated that if they had to use the service again within the duration of the long subscription, they would buy the long subscription the next time irrespective of their first choice.

[57] A precise payoff structure demands a substantially finer subdivision of state 2, that is, considering a separate state for each time unit between d_b and d_l as demonstrated in App. 4.A for the simple case of d_b=2 days (two time units) and d_l=3 days. One can see that the complete payoff structure becomes infinitely long and approaches the payoffs in (4.1), including controls for price discounting.

As decision basis to the choice problem, let us compute the value of L_b over L_l by

$$\begin{aligned} V(L_b - L_l) = \; & \underbrace{\big(v(x_1^{L_b}) - v(x_1^{L_l})\big) \cdot (1 - (1-\pi_d)^{d_b-1} + (1-\pi_d)^{d_l-1})}_{\text{state 1 and 3}} \\ + \; & \underbrace{\big(v(x_2^{L_b}) - v(x_2^{L_l})\big) \cdot \big((1-\pi_d)^{d_b-1} - (1-\pi_d)^{d_l-1}\big)}_{\text{state 2}}. \end{aligned}$$

A second subscription set

To lay the groundwork for the introduced PR, let us substitute the short subscription L_b by a truncated version $L_{b'}$ to form a second lottery pair of the same Problem 1:

$L_{b'}$: $d_{b'}$ free platform use for p_b

L_l: d_l free platform use for p_l,

with $p_b < p_l$ and $d_{b'} = 1 < d_b < d_l$. We get the following payoffs x_s^L for lottery $L \in \{L_{b'}, L_l\}$ in state $s \in S$

$$x_s^L = \begin{cases} x_1^{L_{b'}} = -p_b - p_l, \; (-2p_b); & x_1^{L_l} = -p_l \\ x_2^{L_{b'}} = -p_b - p_l, \; (-2p_b); & x_2^{L_l} = -p_l \\ x_3^{L_{b'}} = -p_b; & x_3^{L_l} = -p_l. \end{cases}$$

Here payoffs are identical in states 1 and 2. The value of $L_{b'}$ over L_l computes as

$$\begin{aligned} V(L_{b'} - L_l) = V(L_{b'}|L_l) \; = \; & \underbrace{\big(v(x_1^{L_{b'}}) - v(x_1^{L_l})\big) \cdot (1 - (1-\pi_d)^{d_l-1})}_{\text{state 1 and 2}} \\ + \; & \underbrace{\big(v(x_3^{L_{b'}}) - v(x_3^{L_l})\big) \cdot (1-\pi_d)^{d_l-1}}_{\text{state 3}}. \end{aligned}$$

Keeping all other payoffs constant between the first and second lottery pair, from $x_1^{L_b} > x_1^{L_{b'}}$ it must follow that

$$\begin{aligned} V(L_b) &> V(L_{b'}) \\ V(L_b - L_l) &> V(L_{b'} - L_l). \end{aligned} \tag{4.2}$$

4.3.2 Problem illustration along 4 cases

To demonstrate the problem let us assume a majority of human decision-makers, in the following represented by a single DM, prefers the long subscription L_l over the short one $L_{b'}$ in the second lottery pair, i.e., $L_{b'} \preceq L_l$.

i) $L_l \succeq L_b$: The long subscription L_l is preferred in both pairs if the value surplus of L_l over $L_{b'}$ is that abundant, that L_b is less preferred than L_l despite the additional duration $(d_b - d_{b'})$ compared to $L_{b'}$.

ii) $L_l \preceq L_b$: Switching preferences to the short subscription L_b in the first lottery pair can be rational if the additional subscription $(d_b - d_{b'})$ tips over the utility advantage in favor of L_b. This case represents no PR.

Let us analyze the counterfactual: A DM prefers $L_{b'}$ over L_l in the second subscription pair, that is, $L_{b'} \succeq L_l$.

iii) $L_b \succeq L_l$: Extending the coverage by $(d_b - d_{b'})$ in L_b, i.e., adding positive value to the short subscription, ought to further advance its attractiveness compared to L_l. Preference for the short subscription remains.

iv) $L_b \preceq L_l$: Switching preferences from the short option in the second lottery pair to the long subscription in the first lottery pair represents an irrational PR since the relative value of L_l reduces in the first pair. Independent of the parameter set, the preferential pattern in *iv)*

$$L_b | L_l \preceq L_{b'} | L_l,$$

violates the principle of transitivity in rational choice theory, given

$$L_l \succeq L_b \succeq L_{b'}.$$

The reference point L_l bears relevance for the shift in preferences. Various models have advanced standard choice theory to unravel context-dependent PRs. Let us scrutinize this behavioral paradox through the lens of three profiled theories in the field.

4.3.3 Prospect theory

A well-known proposition of prospect theory (PT) advocates that "losses loom larger than gains" (Tversky and Kahneman, 1979, 1992). In the context of lotteries, it translates to evaluating positive state payoffs based on concave utility while assessing negative state payoffs with a convex utility function. What is more, state probabilities are distorted by a probability weighting function $w(\pi_s)$ that attaches reduced weight to almost certain events but inflates the perceived probability of unlikely events.[58] Applied to the state space herein, we get

[58]In accord with rank-dependent models, *cumulative* prospect theory (CPT) further distinguishes between $w^+(\pi_s)$ for the domain of gains and $w^+(\pi_s)$ for the domain of losses (Tversky and Kahneman, 1992). Since p is already allocated to index subscription prices in this paper, note the distinct notation of objective probabilities (π_s) compared to Tversky and Kahneman (1979, 1992; p_s).

$$V_{PT}(L) = \sum_{s \in S} v(x_s^L)w(\pi_s) = \underbrace{v(x_1^L)w(\pi_1)}_{\text{state 1}} + \underbrace{v(x_2^L)w(\pi_2)}_{\text{state 2}} + \underbrace{v(x_3^L)w(\pi_3)}_{\text{state 3}}.$$

The same state space in a vis-à-vis comparison between the two subscription sets $\{L_b, L_l\}$, $\{L_{b'}, L_l\}$ leads to equal probability weighting on the left-hand and right-hand side of equation (4.2). If $V(L_b - L_l) > V(L_{b'} - L_l)$, it must also apply

$$V_{PT}(L_b - L_l) > V_{PT}(L_{b'} - L_l).$$

Since monotonicity violation is necessary to produce $V(L_b - L_l) < V(L_{b'} - L_l)$, we conclude that PT fails to explain the PH effect. In reality, the state space shrinks for the second lottery set as states 1 and 2 become indistinguishable for a DM:

$$V_{PT}(L_{b'} - L_l) = \underbrace{\big(v(x_1^{L_{b'}}) - v(x_1^{L_l})\big)w(\pi_1 + \pi_2)}_{\text{state 1 and 2}} + \underbrace{\big(v(x_3^{L_{b'}}) - v(x_3^{L_l})\big)w(\pi_3)}_{\text{state 3}} \tag{4.3}$$

From $\sum_{s \in S} w(\pi_s)$=1 follows $1 - w(\pi_2) = w(\pi_1) + w(\pi_3) = w(\pi_1 + \pi_3)$. One can easily see that *state-splitting* in the present context has no effect on evaluation based on PT. Although the weighting functions in cumulative PT (Tversky and Kahneman, 1992) make it somewhat more complicated, an assessment by CPT does not differ from its original version due to the independence of state-splitting in common. Precisely due to this property, the payoff configuration here, containing merely two different outcomes, does not affect the rank-dependency underlying cumulative PT.

4.3.4 Configural weight model

Closely related to rank- and sign-dependent theories, the intuition behind Birnbaum (1974) configural weight model is the relationship heterogeneity between lottery outcomes. An outcome's relative weight depends on its alternatives. Conditional on the point of view V, the utility U_V of a binary gamble (x_1, π_1, x_2), with outcomes $x_1 < x_2$, and corresponding probabilities π_1; $\pi_2 = 1 - \pi_1$, is defined by

$$U_V(x_1, \pi_1, x_2) = \frac{Au(x_1) + Bu(x_2)}{A + B}.$$

A and B are *configural* weights of x_1, x_2 and depend, besides probabilities, on the configural weighting parameter a_v which reflects a DM's point of view as follows

$$A = a_v S_{x_1}(\pi_1); \qquad B = (1 - a_v)\big(1 - S_{x_1}(1 - \pi_2)\big).$$

Given $.04 < \pi < .96$, Birnbaum et al. (1992) specify two distinct S_{x_1} functions

$$S_{x_1}(\pi) = \begin{cases} .59\pi + .29 & \text{for} \quad x_1 > 0 \\ .74\pi + .14 & \text{for} \quad x_1 = 0, \end{cases}$$

implying $S_{x_1>0} > S_{x_1=0}$, in particular for small probabilities. This violates monotonicity since the zero lower outcome is given less weight than a small positive, lower outcome. Birnbaum (1997) blames underweighting of zero outcomes for the puzzling observation that subjects price a dominated lottery, yielding \$96 with 95% chance and \$0 with 5%, higher than its dominant version paying the same upside of \$96 with .95 but a higher downside of \$24 with .05. *Zero underweighting* places greater relevance to the upside in the dominated lottery, deeming it more attractive to a judge. Analogously, the single-use option may shift attention to rational criteria affecting a DM to prefer it indirectly (between-subject) over the dominant 6-month option. Modeling the present problem through configural weight theory requires normalization to the focal point x_1=0. Applied to the corresponding three-outcome model $(x_1, \pi_1, x_2, \pi_2, x_3, \pi_3)$ in Birnbaum et al. (1992) and relaxing the condition x_1<x_2<x_3, App. 4.B.1 formally proves the existence of a pigeonholing PR given

$$\begin{aligned} & a_v^2(-0.8732\pi_1^2 + 0.1984\pi_1 + 0.096\pi_2 + 0.1116) \\ + \quad & a_v(0.4366\pi_1^2 - 0.4776\pi_1 - 0.06\pi_2 - 0.1702) + 0.018(\pi_3 - \pi_2) > 0, \end{aligned}$$

which holds for small positive a_v-values, $a_v < 0$, and all combinations of (π_1, π_2, π_3).

4.3.5 Salience theory

Bordalo et al. (2012) explain a set of empirically observed PRs (e.g., Allais paradox) in a two-dimensional model that weighs probabilities of states according to their salience in payoff differences as follows:

Definition 1. The salience $\sigma(x_s^i, x_s^{-i})$ of a state $s \in S$ for lottery $L_i, i = 1, 2$, and payoffs x_s^i satisfies three conditions.

1. Ordering: If for states $s, \tilde{s} \in S$, $[x_s^{min}, x_s^{max}] \subset [x_{\tilde{s}}^{min}, x_{\tilde{s}}^{max}]$, then

$$\sigma(x_s^i, x_s^{-i}) < \sigma(x_{\tilde{s}}^i, x_{\tilde{s}}^{-i}).$$

2. Diminishing sensitivity: If $x_s^j > 0$ for $j = 1, 2$, then for any $\epsilon > 0$,

$$\sigma(x_s^i + \epsilon, x_s^{-i} + \epsilon) < \sigma(x_s^i, x_s^{-i}).$$

3. Reflection: For any two states $s, \tilde{s} \in S$ such that $x_s^j, x_{\tilde{s}}^j > 0$ for $j = 1, 2$,

$$\sigma(x_s^i, x_s^{-i}) < \sigma(x_{\tilde{s}}^i, x_{\tilde{s}}^{-i}) \iff \sigma(-x_s^i, -x_s^{-i}) < \sigma(-x_{\tilde{s}}^i, -x_{\tilde{s}}^{-i}).$$

An exemplary salience function represents

$$\sigma(x_s^i, x_s^{-i}) = \frac{|x_s^i - x_s^{-i}|}{|x_s^i| + |x_s^{-i}| + \theta}, \quad \text{where } \theta > 0. \tag{4.4}$$

Definition 2. Distorted decision weights of state probabilities are defined by the salience rank $k_s^L \in \{1, ..., |S|\}$ of state $s \in S$ in lottery $L_i, i = 1, 2$. Lower k_s^L indicate higher salience in the pairwise ratio of states $s, \tilde{s} \in S$

$$\frac{\pi_{\tilde{s}}^L}{\pi_s^L} = \delta^{k_{\tilde{s}}^L - k_s^L} \cdot \frac{\pi_{\tilde{s}}}{\pi_s}, \tag{4.5}$$

where $\delta \in (0, 1]$. States of equal salience are ranked alike, and ranking runs smooth (i.e., no jumps). After normalizing $\sum_{s \in S} \pi_s^L = 1$ and defining $\omega_s^L = \frac{\delta^{k_s^L}}{\sum_{s \in S} \delta^{k_s^L} \cdot \pi_s}$, the decision weight attached to state s in the evaluation of lottery L is given by

$$\pi_s^L = \pi_s \cdot \omega_s^L.$$

Per Definitions 1 and 2, a so-called *local thinker* (Bordalo et al., 2012) affected by probability distortion through salience differences, evaluates a lottery L then by

$$V_{ST}(L) = \underbrace{v(x_1^L) \cdot \pi_1^L}_{\text{state 1}} + \underbrace{v(x_2^L) \cdot \pi_2^L}_{\text{state 2}} + \underbrace{v(x_3^L) \cdot \pi_3^L}_{\text{state 3}} \tag{4.6}$$

$$= \frac{v(x_1^L) \cdot \pi_1 \delta^{k_1^L} + v(x_2^L) \cdot \pi_2 \delta^{k_2^L} + v(x_3^L) \cdot \pi_3 \delta^{k_3^L}}{\pi_1 \delta^{k_1^L} + \pi_2 \delta^{k_2^L} + \pi_3 \delta^{k_3^L}}. \tag{4.7}$$

The salience ranking in Definition 1 depends in essence on the ratio of lottery prices p_b and p_l, as well as on the preferred follow-up subscription if state 2 realizes and L_b was chosen initially. We distinguish between four cases.

Case 1 and Case 3

This paragraph comprises two cases viewed equally by salience theory. Case 1 requires the price relation $\frac{p_l}{p_b} - \frac{p_b}{p_l} < 1$ and follow-up choice L_l, given L_b was selected initially. Case 3 links the price ratio $\frac{p_l}{p_b} < \sqrt{2}$ with follow-up choice L_b, if L_b was initially chosen. Given these conditions, equation (4.4) generates in both cases the same salience ranking of states for the subscription sets $\{L_b, L_l\}$ and $\{L_{b'}, L_l\}$

$$\text{Case 1} = \begin{cases} V(L_b - L_l)\colon & \underbrace{\sigma(-p_b, -p_l)}_{\text{state 1}} = \underbrace{\sigma(-p_b, -p_l)}_{\text{state 3}} < \underbrace{\sigma(-p_b - p_l, -p_l)}_{\text{state 2}} \\ V(L_{b'} - L_l)\colon & \underbrace{\sigma(-p_b, -p_l)}_{\text{state 3}} < \underbrace{\sigma(-p_b - p_l, -p_l)}_{\text{state 1}} = \underbrace{\sigma(-p_b - p_l, -p_l)}_{\text{state 2}} \end{cases}$$

$$\text{Case 3} = \begin{cases} V(L_b - L_l)\colon & \underbrace{\sigma(-p_b, -p_l)}_{\text{state 1}} = \underbrace{\sigma(-p_b, -p_l)}_{\text{state 3}} < \underbrace{\sigma(-2p_b, -p_l)}_{\text{state 2}} \\ V(L_{b'} - L_l)\colon & \underbrace{\sigma(-p_b, -p_l)}_{\text{state 3}} < \underbrace{\sigma(-2p_b, -p_l)}_{\text{state 1}} = \underbrace{\sigma(-2p_b, -p_l)}_{\text{state 2}}, \end{cases}$$

with the corresponding salience ranks

$$k_s^L = \begin{cases} k_1^{L_b} = 2; k_1^{L_b'} = 1 \\ k_2^{L_b} = 1; k_2^{L_b'} = 1 \\ k_3^{L_b} = 2; k_3^{L_b'} = 2. \end{cases} \tag{4.8}$$

Assuming linear utility[59], the value of L_b over L_l ($L_{b'}$ over L_l) computes in Case 1 by

$$\begin{aligned} V_{ST}(L_b|L_l) =& \frac{\pi_1\delta^2(p_l - p_b) + \pi_2\delta^1(-p_b) + \pi_3\delta^2(p_l - p_b)}{\pi_1\delta^2 + \pi_2\delta^1 + \pi_3\delta^2} \\ =& \frac{(\pi_1 + \pi_3)\delta^2(p_l - p_b) + \pi_2\delta^1(-p_b)}{(\pi_1 + \pi_3)\delta^2 + \pi_2\delta^1} \end{aligned} \tag{4.9}$$

$$V_{ST}(L_{b'}|L_l) = \frac{(\pi_1 + \pi_2)\delta^1(-p_b) + \pi_3\delta^2(p_l - p_b)}{(\pi_1 + \pi_2)\delta^1 + \pi_3\delta^2},$$

[59]See Bordalo et al. (2012) for proof considering non-linear utility functions.

and in Case 3 by

$$V_{ST}(L_b|L_l) = \frac{(\pi_1+\pi_3)\delta^2(p_l-p_b)+\pi_2\delta^1(p_l-2p_b)}{(\pi_1+\pi_3)\delta^2+\pi_2\delta^1}$$

$$V_{ST}(L_{b'}|L_l) = \frac{(\pi_1+\pi_2)\delta^1(-p_b)+\pi_3\delta^2(p_l-p_b)}{(\pi_1+\pi_2)\delta^1+\pi_3\delta^2}.$$

Appendix 4.B.2 formally proves that $V_{ST}(L_b - L_l) > V_{ST}(L_{b'} - L_l)$ holds for any $\pi_d > 0$ in Case 1 and Case 3. Furthermore, from equation (4.9) we can easily see that state-splitting does not affect salience theory.

Case 2 and Case 4

The two remaining cases cover the corresponding, opposite conditions. Case 2 refers to the price relation $\frac{p_l}{p_b} - \frac{p_b}{p_l} > 1$ with follow-up preference L_l, if L_b was chosen initially. Case 4 links the price relation $\frac{p_l}{p_b} > \sqrt{2}$ with follow-up choice L_b. By ordering

$$\text{Case 2} = \begin{cases} (L_b, L_l)\colon & \underbrace{\sigma(-p_b-p_l,-p_l)}_{\text{state 2}} < \underbrace{\sigma(-p_b,-p_l)}_{\text{state 1}} = \underbrace{\sigma(-p_b,-p_l)}_{\text{state 3}} \\ (L_{b'}, L_l)\colon & \underbrace{\sigma(-p_b-p_l,-p_l)}_{\text{state 1}} = \underbrace{\sigma(-p_b-p_l,-p_l)}_{\text{state 2}} < \underbrace{\sigma(-p_b,-p_l)}_{\text{state 3}} \end{cases}$$

$$\text{Case 4} = \begin{cases} (L_b, L_l)\colon & \underbrace{\sigma(-2p_b,-p_l)}_{\text{state 2}} < \underbrace{\sigma(-p_b,-p_l)}_{\text{state 1}} = \underbrace{\sigma(-p_b,-p_l)}_{\text{state 3}} \\ (L_{b'}, L_l)\colon & \underbrace{\sigma(-2p_b,-p_l)}_{\text{state 1}} = \underbrace{\sigma(-2p_b,-p_l)}_{\text{state 2}} < \underbrace{\sigma(-p_b,-p_l)}_{\text{state 3}} \end{cases}$$

we have

$$k_s^L = \begin{cases} k_1^{L_b} = 1; k_1^{L'_b} = 2 \\ k_2^{L_b} = 2; k_2^{L'_b} = 2 \\ k_3^{L_b} = 1; k_3^{L'_b} = 1. \end{cases} \tag{4.10}$$

The general proofs in Appendix 4.B.2 demonstrate that salience theory cannot accommodate the PH effect. In consequence

$$V_{ST}(L_b - L_l) > V_{ST}(L_{b'} - L_l)$$

holds for any p_b, p_l and any $\delta \in (0,1]$, given $\pi_d > 0$. Instead, let us suppose the reason for this type of PR lies in a stimulus not grasped in the model. The following paragraph proposes an extension of salience theory in this regard.

4.3.6 Pigeonholing: a novel dimension of salience

In an attempt to model the stimulus of the short subscription d_b, we generalize the foundations of salience theory for a third dimension. A DM suffering from pigeonholing deviates from equation (4.7) by overweighting states of maximum regret (largest payoff difference) in growing d_b. By formally defining the salience of a decision's rational dimension through the length of the short subscription in a choice set, we appeal to Bordalo et al.'s (2012) request of advancing the salience conjecture to further variables beyond pure payoff differences:

Definition 3. The salience $\delta(d_b)$ of the acquisition utility[60] in assessing a lottery set $\{L_b, L_l\}$ is a continuous and bounded function that satisfies two conditions.

1. Ordering: If for the length $d_b, d_{b'}$ of any short option, we have $d_b > d_{b'} > 0$, then

$$\delta(d_b) < \delta(d_{b'}).$$

2. Diminishing sensitivity: If $d_b > d_{b'} > 0$, then for any $\epsilon > 0$

$$\delta(d_b + \epsilon) - \delta(d_{b'} + \epsilon) < \delta(d_b) - \delta(d_{b'}).$$

To implement Definition 3, consider a salience function of the type

$$\delta(d_b) = \left(\frac{1}{d_b}\right)^c, \quad \text{where } c > 0. \tag{4.11}$$

In standard salience theory, parameter δ represents a constant estimated ~ 0.7 for most problems (Bordalo et al., 2012). This invariability satisfies the dominance principle of salience theory. Since the examined PR results from monotonicity violation, standard salience theory cannot explain the PH effect.

In a generalization of salience theory, δ indicates the degree to which salience of the acquisition utility distorts state probabilities in the evaluation of a subscription pair. From equation (4.5) and Definition 2 follows: the lower δ, the greater the

[60] Acquisition utility refers to rational utility, see Section 4.1.2.

probability distortion in favor of maximum regret states. Note that equation (4.11) accounts for the special role of $d_b = 1$ by $\delta(1) = 1$. L_b is then preferred over L_l, if

$$V(L_b - L_l) = \sum_{s \in S} \pi_s \omega_s^{L_b} \cdot (x_s^{L_b} - x_s^{L_l})$$
$$= \frac{\pi_1 \delta^{k_1^L}(x_1^{L_b} - x_1^{L_l}) + \pi_2 \delta^{k_2^L}(x_2^{L_b} - x_2^{L_l}) + \pi_3 \delta^{k_3^L}(x_3^{L_b} - x_3^{L_l})}{\pi_1 \delta^{k_1^L} + \pi_2 \delta^{k_2^L} + \pi_3 \delta^{k_3^L}} > 0.$$

For any pairwise comparison of a subscription triple $\{L_b, L_{b'}, L_l\}$ with $d_{b'} < d_b < d_l$, a DM commits pigeonholing, if her evaluation satisfies:

$$\begin{aligned} V(L_b - L_{b'}) &> 0 \\ V(L_b - L_l) &< V(L_{b'} - L_l) \end{aligned} \tag{4.12}$$

Case 1 $\left(\frac{p_l}{p_b} - \frac{p_b}{p_l} < 1\right.$**, follow-up** $\left.L_l\right)$ **and Case 3** $\left(\frac{p_l}{p_b} < \sqrt{2}, L_b\right)$

Given the state salience ordering in (4.8), from equation (4.12) follows

$$(1-\pi_d)^{d_b-1} - (1 - \frac{d_b}{d_{b'}})(1-\pi_d)^{d_l-2}((1-\pi_d)^{d_b} - (1-\pi_d)^{d_l}) > 1$$

which holds for low enough values of π_d and $d_b >> d_{b'}$. See App 4.B.3 for proof.

Case 2 $\left(\frac{p_l}{p_b} - \frac{p_b}{p_l} > 1, L_l\right)$ **and Case 4** $\left(\frac{p_l}{p_b} > \sqrt{2}, L_b\right)$

Based on the ordering in (4.10), equation (4.12) gives

$$\frac{1}{d_{b'}}(1 - (1-\pi_d)^{d_b-1}) + (1-\pi_d)^{d_l-2}(\frac{1}{d_b} - \frac{1}{d_{b'}})((1-\pi_d)^{d_l} - (1-\pi_d)^{d_b}) < 0$$

which holds only for the impossible cases $\pi_d < 0$ and $\pi_d > 1$. (See Appendix 4.B.3 for proof.) PH can exclusively occur for price ratios $\frac{p_l}{p_b} - \frac{p_b}{p_l} < 1$ combined with follow-up preference L_l, or $\frac{p_l}{p_b} < \sqrt{2}$ with follow-up preference L_b, and never otherwise. Moreover, we can infer from the requirement of low enough π_d-values, that the PR only materializes if $E[V(L_b)]$ exceeds $E[V(L_l)]$. In line with standard salience theory, the extended model is independent of absolute price levels.

Attraction for the short subscription

Based on the core idea of PH – the short subscription's influential role in a choice set evaluation – let us define the attraction $\eta(L_b|L_l) = \eta(d_b, d_l, p_b, p_l)$ for the short subscription L_b relative to L_l, expressed in the share of preferences. We can then formulate the following theorem.

Hypothesis: Given the preference relation $L_{b'} \preceq L_b \preceq L_l$ with prices $p_{b'} = p_b < p_l$ and duration $d_{b'} < d_b < d_l$, we have $\frac{p_l}{p_b} - \frac{p_b}{p_l} < 1$, respectively $\frac{p_l}{p_b} < \sqrt{2}$, then applies

$$\eta(L_b|L_l) < \eta(L_{b'}|L_l).$$

4.3.7 Note on regret theory

Interpretation in light of regret theory may prove helpful to comprehend this conjecture (Bell, 1982; Loomes and Sugden, 1982). Let us label states yielding negative payoffs relative to the long subscription *regret states*. Compared to the rational benchmark of zero probability distortion, regret states gain salience for $\frac{p_l}{p_b} - \frac{p_b}{p_l} < 1$, respectively $\frac{p_l}{p_b} < \sqrt{2}$, but lose salience for $\frac{p_l}{p_b} - \frac{p_b}{p_l} > 1$, respectively $\frac{p_l}{p_b} > \sqrt{2}$. Low price ratios provide more leverage for the notion of regret than high price ratios.

Due to $\delta = (\frac{1}{d_b})^c$, salience of regret states grows in d_b if $\frac{p_l}{p_b} - \frac{p_b}{p_l} < 1$, respectively $\frac{p_l}{p_b} < \sqrt{2}$ (see the salience relations derived in Section 4.3.5). In relative terms, a week-long subscription yields a higher chance of closely missing out than a day pass (Gilbert et al., 2004). Missing out by one day in the latter implies a *regret ratio* of 1:1. Supposedly, this does not feel as close as a miss than the same absolute deviation of one day given a weekly subscription (seven days), which produces a much higher ratio of 7:1. Even missing out in one of the next six days would spawn a larger regret ratio. The elevated regret risk may contribute to the PR. Feedback from participants supports considerations of non-linear salience effects in regret states with rising d_b. In the following, we test the predictive value of the formulated PH hypothesis along varying context to control for identification effects.

4.4 Online experiments

We tested the replicability of the results from the classroom study in a series of online experiments in August 2020 (Table 4.1). Before running the sessions coded in *oTree* (Chen et al., 2016) and conducted via *Prolific*, a crowdfunding platform tailored explicitly to research purposes (Palan and Schitter, 2018), the hypothesis and experimental design was pre-registered and approved by the AEA registry for randomized controlled trials. The sample sizes derive from a power analysis of the classroom study results (Appendix 4.C).

Table 4.1: Questionnaire information from 185 participants.

	Total	Pilot 1	Pilot 2	Study 2	Study 3	Study 4a	Study 4b	Study 5a	Study 5b
Total participants	343/344	22/25	23/18	30/37	51/51	62/62	30/32	61/60	64/59
Attrition rate (%)	-	0.0/0.0	0.0/0.0	0.0/10.8	0.0/2.0	1.6/4.8	6.7/3.1	1.6/3.3	0.0/3.4
Comprehension test failed (%)	-	-	-	13.3/15.2	7.8/6.0	1.6/11.9	14.3/25.8	21.7/13.8	18.8/21.1
Consistency check failed (%)	-	-	-	-	9.8/38.0	-	17.9/3.2	8.3/6.9	4.7/12.3
Female (%)	37.6/42.0	72.7/56.0	91.3/44.4	60.0/40.0	31.4/56.0	30.0/47.5	28.6/19.3	40.0/34.5	34.4/40.3
Age (yrs)	25.4/25.4	21.2/21.4	22.0/21.9	28.7/27.1	24.4/24.4	24.8/24.9	24.6/24.7	27.2/28.3	27.1/26.8

4.4.1 Design

A session welcomes participants on the instructions page. It briefly describes their task of choosing between two subscriptions as if they face this decision in real life. Proceeding to the main stage, subjects state their preferences over the set of choice problems (constructs), including control questions for comprehension and consistency. After that, the program elicits the hypothetical probability of reusing such a subscription service within the duration of the long subscription, termed as *re-use likelihood* in the following. Some sessions feature a non-incentivized cognitive reflection test (CRT), slightly modified from the original version by Frederick (2005) to account for familiarity effects (Stieger and Reips, 2016). The experiment concludes with a questionnaire on demographic information. An average session lasts between 6 and 15 min. Subjects earn between 1 and 3 £.

4.4.2 Analysis

The analytical section centers on the treatment effect related to the PH hypothesis. Appendix 4.F onward provides more information on the treatment difference concerning general subscription preferences and stated re-use likelihood. Since many questions feature one-week options as short subscriptions in the treatment condition, the notation refers to it as *week* treatment compared to the control condition, labeled as *single*. The question sets used in the individual studies can be provided on request.

PH hypothesis

Contrary to the paper-and-pen Study 1, the online experiments do not bear significant evidence for the PH effect in the pooled data of 599 independent observations

(App. Table 4.6).[61] Besides various sub-samples analyzed below, the data reports statistically significant PRs in 4 of 21 questions. More importantly, the result on the aggregate level is remedied by two striking observations, discussed below.

Probability interaction term

The most considerable effect on preferences exerts re-use likelihood. Across all contexts and studies except Study 4a, there is overwhelming evidence (p<.001) of a negative relationship between re-use likelihood and preference for the shorter subscription.[62,63] It seems reasonable to favor a short subscription if future re-use is unlikely.

Regressing preferences for the short option on re-use likelihood in a difference-in-difference approach gives insight into behavioral drivers (Appendix Table 4.6). The sign of the interaction term for re-use likelihood and single treatment follows the PH prediction attributing PRs to scenarios of low re-use probability (see Section 4.3.6). A negative sign implies that subjects stating low re-use likelihood are more likely to choose the short subscription in the single than in the week treatment. This relationship promotes the PH effect. Across Studies 2-5, 26 out of 33 constructs display a negative sign of the interaction term. Seven constructs find mild evidence on the p<.1 level. Three find evidence on the p<.05 level. One construct reports strong evidence on the p<.01 level, while none of the seven constructs with a positive interaction term bear statistical significance in one-sided tests. The estimated re-use likelihoods of specific demographics explain PRs in a variety of sub-samples, described in the following.

Study 2

Study 2 is a direct validation check of the classroom experiments. It varies in two further aspects besides the distinct form (online) of data collection, i.e., questions are presented in English and prices in £.

Re-use likelihoods for ride-sharing and express delivery, products specifically targeted at younger consumer groups, decreases with age (p<.05). Yet, the Rockefeller question reports a tendency of older subjects stating higher re-use likelihood in the control than in the treatment group (Table 4.11). This relation possibly causes the PR observed in this sub-sample. De facto, the data notes an age driven PH effect in the fragment of 30+ years old subjects (one-sided t-test, $t(17)$=-2.4, p=.014). That

[61] All probit regressions cluster on the subject level.

[62] Distinct phrasing of the likelihood elicitation question appears influential (see Section 4.4.1). Note that this variation neither affects the sign nor the size of the interaction term.

[63] See Appendix 4.F onward for probit regressions on the study level.

is to say, 40% in the treatment group of this sub-sample chose the short subscription vs. 89% in the control.[64]

As illustrated in Figure 4.1, this sample is much more diverse in terms of age, nationality, and education than the pilot study. In particular, age seems to affect re-use likelihood and thereby may indirectly alter results. Except for the Rockefeller question, the subscription options yield a fairly large price gap, which may impair the occurrence of PRs (see Section 4.3.6). Overall high re-use likelihood, which does not provide the theoretical ground for the PH hypothesis (see Section 4.3.6) could further explain the null results.

Study 3

Accounting for such probability dependency, Study 3 tests additional choice contexts (accommodation-sharing, flood insurance, lock-down insurance). To control for age effects, we restrict participation to the range of 21-30 years.[65] The relative price gap between the short and the long subscription is reduced in conformation with the proposition in Section 4.3.6. Study 3 and subsequent studies pose choice problems to subjects in growing order of price levels.

Table 4.2: Study 3, one-sided t-tests - share of short subscription across questions

	All	Accom	Ride	Ride Age<25	Rockef.	Flood	Covid	Covid Age>25	Covid Edu>3	Covid Prob<15
Week (treatment)	.281	.511	.150	.111	.468	.043	.234	.067	.219	.375
		(.074)	(.052)	(.061)	(.074)	(.030)	(.062)	(.067)	(.074)	(.125)
Single (control)	.340	.468	.213	.250	.617	.191	.213	.250	.400	.636
		(.074)	(.060)	(.083)	(.072)	(.058)	(.061)	(.112)	(.112)	(.152)
degr. of freedom	468	92	92	53	92	92	92	29	50	25
p-value	0.084	0.658	0.213	0.094	0.075	0.012	0.597	0.088	0.083	0.098

Clustered on subject level and corrected for failed comprehension questions. Standard errors in parentheses.

Study 3 finds overall mild evidence of the PH effect in one-sided hypothesis tests (probit: p=.059, Appendix Table 4.12; t-test: $t(468)$=-1.4, p=.084, Table 4.2). Except for the context of accommodation-sharing, the treatment effect works across all constructs in the direction predicted by the PH hypothesis.[66] Table 4.2 reports mild evidence of an age-driven PH effect in the ride-sharing context, $t(53)$=-1.3, p=.094.

[64] Robustness tests of the PH hypothesis are performed one-sided throughout the paper.

[65] Comparing standard deviations in age between studies confirms much less variation in Study 3. Although conducted in English, further requisites applied are Italian fluency and Italy as country of residence for better comparability to Study 1. In none of the two treatment groups, subjects returned their submission, which points to no apparent selection effects.

[66] See App. Tables 4.12 and 4.14 for signs of control dummy in probit, resp. linear regressions.

In contrast to the PH effect observed for older participants (30+ years) in the Rockefeller context of Study 2, here, younger participants (<25 yrs) tend to fall for the "subscription trap". 11% in this sub-sample prefer the short subscription in the treatment group vs. 25% in the control. Moreover, Appendix Table 4.12 reports a mild PH effect for the Rockefeller context, p=.075. Note that price parameters have been raised roughly four times compared to Studies 1 and 2.

Similar to Study 2, Appendix Table 4.12 also denotes a strong negative impact of re-use likelihood on preferences for the short subscription (p<.01), except for the flood construct. In line with the model predictions, it states a significant negative relationship between re-use likelihood and PRs (p=.012, one-sided).

COVID lock-down insurance

Study 3 extends the question set for the context of insurance.

Problem 4 Your government wants to learn from the current COVID-19 pandemic and protect the economy against future lock-downs through state insurance. The state insurance pays 10,000 € during every lock-down to insured businesses, including a retrospective payment for the recent lock-down as instant aid. If you were a business owner, what would you pick?

Treatment group (47 subjects)

A: 2 yrs insurance against pandemic-related lock-downs incl. recent one for 3,000 € [23.4%]
B: 10 yrs insurance against pandemic-related lock-downs incl. recent one for 3,500 € [76.6%]

Control group (47 subjects)

A': Single retrospective insurance against recent one for 3,000 € [21.3%]
B: 10 yrs insurance against pandemic-related lock-downs incl. recent one for 3,500 € [78.7%]

Whereas the probit analysis in Appendix Table 4.12 (column 11) finds no general evidence for the PH hypothesis in the lock-down insurance context, it identifies two sample characteristics, having a stimulating impact nonetheless. Column 12 finds a mild, positive age effect in a one-sided t-test for the sub-sample of 25+ years, $t(29)$=-1.4, p=.088. Yet, it remains unclear why older subjects are more likely to fall for the "insurance trap" as none of the characteristics correlate significantly with the estimated risk of experiencing a second lock-down in their respective home country. Appendix Table 4.12 reports supportive evidence (p<.05) in favor of the theoretical precondition of the PR: low re-use likelihood. Table 4.2 states mild evidence for the PH effect (p<.1) in the sub-sample of participants estimating the odds of a further lock-down less than 15%.

Flood insurance

The second insurance question, placed in reverse order in the experiment, presents the subsequent problem to subjects.

Problem 5 As you may know: Venice, the Italian lagoon city, was struck desperately by a historic flood last November. Imagine the Venetian administration created a fund to support property owners in the current situation and to provide insurance against damage from potential floods in the future. The fund insurance pays 5,000 € during every critical flood, including a retrospective payment for the November flood. An expert committee estimates the probability of further critical floods to occur in any given year to be 1% = 0.01. If you owned property in Venice that has been seriously damaged during the November flooding, what would you choose?

Control group (47 subjects)

A: 6 months insurance against flood-related damage incl. the Nov. flood for 1,500 €	[4.3%]
B: 60 months insurance against flood-related damage incl. the Nov. flood for 1,700 €	[95.7%]

Treatment group (47 subjects)

A': Single retrospective insurance against November flooding for 1,500 €	[19.1%]
B: 60 months insurance against flood-related damage incl. the Nov flood for 1,700 €	[80.9%]

Probit analysis in App. Table 4.12 (column 9) finds supportive evidence for the PH effect, p=.012. The dummy for female exercises a significant, positive effect (p<.01) on the estimated flood risk (App. Table 4.13). This relationship, also detected in Study 5, is in line with the common finding of more pronounced risk aversion among women (e.g., Borghans et al., 2009).[67,68]

Providing objective information on flood risks intends to limit the impact of subjective likelihood estimates on subscription preferences. Still, we find evidence of a PH effect, predominantly attributed to subjects stating low risk estimates (p=.020).

A probability bias?

Not only is it that subjective probability formation affects preferences. A one-sided hypothesis test reveals that subjects in the week condition evaluate the risk of future floods significantly lower than their peers in the single condition. Following the idea behind the availability heuristic (Tversky and Kahneman, 1973), pigeonholing pre-

[67] Correlation analysis states a small positive effect between women and risk aversion elicited on a 9-point Likert scale (r=-.15)

[68] An analysis of other demographic factors in Appendix Section 4.D surfaces no effect from education, explicitly rejecting the conjecture that better-educated individuals are less prone to fall for a subscription/insurance trap.

dicts the opposite. Namely that a sufficiently short length of the brief alternative aids rational evaluation of the expected re-use likelihood, which is typically overrated. Regardless, if likelihood estimates suffer from a treatment effect even though an anchor probability is provided (Problem 5), there is reason to suspect an impact also in constructs featuring no focus point. Yet, aggregate data show no significant difference in likelihood estimates between treatments (Appendix Tables 4.7-4.9).

Choice-matching preference reversals

Besides extending the research scope for the insurance context, Study 3 introduces the price-matching task as a further hypothesis test. The revised construct shows only the price of the long option, and subjects state the price for the short alternative that would make them indifferent between both (WTP). The PH hypothesis predicts lower WTP-values in the treatment group than in the control. A one-sided t-test ratifies that despite the reference price held constant between treatments (8 weeks of free delivery for 5.99 €), subjects in the control group indicate a significant higher WTP (3.68 €) for a single delivery than their peers in the treatment group for one week of free delivery (2.73 €, $t(92)=-2.3$, $p=.011$).

Table 4.3: Study 3 and 4a, median test - WTP in € across questions

	Express Study 3	Music-streaming	Accom-sharing	Ride-sharing	Skate rental	Express delivery	Rocke-feller	Flood insurance	Lock-down insurance
Week treatment	2.99	0.10	0.50	0.50	0.75	1.20	8.99	190	800
SD	2.10	0.30	0.58	0.48	1.61	1.74	12.14	206.3	475.8
Single treatment	3.99	0.25	1.00	1.00	3.00	3.50	20	500	1000
SD	1.90	0.31	0.83	1.46	1.87	2.09	15.42	499.6	1093.4
p-value	.074	.003	.008	.008	.005	.001	.009	.001	.023
continuity corrected	.107	.006	.013	.013	.009	.002	.015	.001	.034

Note: Standard deviation corrected for outliers.

Study 4a

Based on the strong evidence of a PH effect in the matching task, Study 4a applies this preference elicitation method to a wider range of contexts, including skate rental (Problem 6) and music-streaming. We use a non-parametric two-sample median test to falsify the null hypothesis that the two treatment samples are drawn from populations with the same median.[69] Table 4.3 reports one-sided *p*-values documenting a strong PH effect on the $p<.01$ level in all but the lock-down construct, for which evidence is only significant on the $p<.05$ level. A complementary Wilcoxon rank-sum test for unmatched data finds similar evidence.

[69] For two samples, the chi^2 test statistic is computed with and without continuity correction.

Problem 6 Your uncle from across the country travels with his family to the coast for a summer vacation. They are stopping at your place. Your uncle and his wife would like to see a popular art gallery in town. In the meantime, you are taking their 6-year old daughter for a walk. Passing by the skating rink downtown, your little niece begs you to skate with her. As children get in for free, you can choose between two alternatives to rent inline skates for yourself. You expect only a small chance of 5% that you will go skating again over the next half year.

Treatment condition (62 subjects)

A: 1 week unlimited skating for ___ [median 0.75 €]
B: 20 weeks unlimited skating for 4.50 €

Control condition (62 subjects)

A': One-time skating for ___ [median 3.00 €]

B: 20 weeks unlimited skating for 4.50 €

In line with the previous studies, the best PR predictor seems to be low re-use likelihood. Across all but the lock-down context, the sign of the interaction term between re-use likelihood and single treatment follows the prediction by the PH hypothesis. For ride-sharing and skate-rental, respectively accommodation-sharing and express delivery, one-sided tests in Appendix Table 4.15 state significant (p<.05), respectively mild evidence (p<.1) of a PH effect. Note that all other studies elicit subjects' re-use likelihood conditional on their choice. Study 4a does not particularly mention this. Significantly lower estimates in Appendix Tables 4.8 and 4.9 point out that people evaluate utilization probabilities more conservative if likelihood elicitation is less attached to the choice problem. When re-use likelihood assessment is related to the chosen subscription, possibly creating "lock-in costs" effects, it seems subjects are more inclined to rate chances of follow-up use higher.

Study 4b

Being a direct variation and robustness check of Study 4a, Study 4b features the same set of eight questions but returns to the choice method for preference elicitation. In the accommodation-sharing and skating constructs, the short option of the corresponding treatment condition has been doubled by one week. As a result, the PH effect vanishes almost completely. Appendix Table 4.17 reports evidence of PRs merely in the music-streaming context (p<.05). Although still present in the majority of cases, the PH effect is not as prominent among subjects with low re-use likelihood as in previous studies. One-sided tests find mild evidence only for accommodation-sharing and skate rental (p<.1). Note that like the flood insurance construct, Problem 7 also specifies a likelihood anchor (5%). Nonetheless, the signs of the interaction terms between re-use likelihood and single treatment are negative in both cases. The skating context shows even mild evidence. Moreover, re-use likelihood estimates are lower in the single condition than in the week treatment, yet mildly significant only in the flood construct.

Study 5: testing the PH model

Study 5 tests the prediction by the PH hypothesis regarding the relative price gap between alternative subscriptions (insurances). According to the proposition in Section 4.3.6, a PR can occur given $\frac{p_l}{p_b} - \frac{p_b}{p_l} < 1$.[70] Table 4.4 finds evidence against this conjecture. It reports significant PH effects in the Rockefeller context when price levels do not meet this requirement (Study 5a) but not when they do (Study 5b). There is also evidence in favor of the model prediction. Table 4.4 states a mildly significant PH effect in the Venice context when $\frac{p_l}{p_b} - \frac{p_b}{p_l} < 1$ holds, whereas not when price levels reverse the inequality.

Table 4.4: Study 5a+b, one-sided t-tests - share of short subscription across questions

	All a	All b	Music a	Music b	Ride a	Ride b	Rockef. a	Rockef. b	Flood a	Flood b
Week treatment	.522	.423	.534	.306	.586	.387	.828	.887	.138	.113
			(.066)	(.059)	(.065)	(.062)	(.050)	(.041)	(.046)	(.041)
Single treatment	.565	.403	.482	.222	.603	.370	.948	.815	.224	.204
			(.066)	(.066)	(.065)	(.057)	(.029)	(.053)	(.055)	(.055)
degrees of freedom	462	462	114	114	114	114	114	114	114	114
p-value	.166	.820	.291	.236	.426	.972	.020	.862	.116	.091

Note: Clustered on subject level and corrected for failed comprehension questions. Standard errors in parentheses.

In line with Study 2 and 3, Study 5 also measures cognitive ability via a CRT, but exhibits no significant dependencies between the CRT performance and stated preferences (see Appendix 4.D for more details). In summary, there is no supportive evidence that the PH effect is predominantly driven by low-CRT or naïve individuals and that subjects performing better in the CRT tend to lessen it. Aggregate data shows further no evident correlation between educational level and CRT performances, and rejects the conjecture that better-educated individuals are less prone to fall for the so-called subscription or insurance trap.

[70] We believe it is reasonable to test the more conservative case $\frac{p_l}{p_b} - \frac{p_b}{p_l} < 1$ before $\frac{p_l}{p_b} < \sqrt{2}$.

4.5 Discussion

The framing of decisions depends on the language of presentation, on the context of choice, and on the nature of the display

— Tversky and Kahneman (1986)

Normative approaches to frame human decision-making date back to the ancient Greek philosophers (e.g., Socrates, Plato, Aristotle) and existed most likely even before. Despite enjoying the appeal of rationality and the justification of logical reasoning, the critique of lacking empirical validity has coexisted ever since. Von Neumann and Morgenstern (1944) expected utility theory (EUT) - mother of all rational, thus prescriptive models - set the gold standard of the classical *homo oeconomicus* assumption. Starting with Allais (1953), scholars have showcased numerous empirical examples demonstrating violations of the axiomatic foundations of the prevailing EUT model. Simon (1957) seminal paper marks the uprise of descriptive models that advocate *bounded rationality* to close the gap between theoretically predicted and empirically observed behavior. The urge for a unifying explanatory model culminated eventually in the manifestation of prospect theory (Tversky and Kahneman, 1979, 1992).This section presents a reflection on the PH effect in the context of well-known PRs in the literature. First, we describe these behavioral anomalies and relate them to the pigeonholing phenomenon. The second part summarizes our findings in this paper enhanced by a semantic analysis of verbalized strategies by the participants and provides implications for applied use as well as an outlook on follow-up research.

4.5.1 Preference reversals

As noted above, the axiomatic postulates of rational choice theory have been questioned in terms of their external validity one after the other. In the spirit of Savage (1954) *sure thing principle*, the Allais paradox profaned the *cancellation* or *common consequence axiom*, i.e., the principle neglecting any influence on preferences from equally adding or subtracting probability of some event to all alternatives in a given choice set:

Problem 7 Imagine you had to choose between the following two lotteries:

A: 5M € with 33% and 0 € with 67%
B: 1M € with 34% and 0 € with 66%

A′: 5M € with 33%, 1M € with 66%, and 0 € with 1%
B′: 1M € with 100%

Switching preferences from A over B to favoring B′ over A′ is the observed norm, although the only difference between the latter and the former two lotteries lies in a 66% probability increase of winning 1M €. Ellsberg (1961) similarly corroded the principles of rational choice theory by demonstrating that uncertainty affects human decision-making in irrational manner:

Problem 8 Imagine an urn consisting of 30 red, and 60 black and yellow marbles with the respective numbers of black and yellow marbles are unknown:

A: 100 € if you draw a red marble, 0 € otherwise
B: 100 € if you draw a black marble, 0 € otherwise

A′: 100 € if you draw a red or yellow marble, 0 € otherwise
B′: 100 € if you draw a black or yellow marble, 0 € otherwise

Most people veer from choosing lottery A over B to preferring B′ over A′. Under the premises of rational choice, this type of switching implies inconsistent preferences. Preferring A over B means essentially assuming fewer than 30 black marbles in the urn. By merely adding extra winning probability (i.e., drawing a yellow marble) equally to both lotteries A′ and B′, a rational DM should not switch preferences. The pigeonholing paradox goes even one step further by adding extra probability to only one of the alternatives. Thereby, it counterintuitively depreciates its attractiveness in the eye of a judge. Problems 7 and 8 show the evolutionary, human dislike for risk, respectively ambiguity. Single-use (A′) in Problem 1 provides an unambiguous value and, thus, may be assessed differently than the 6-month entrance (A) which gained comparatively free utility but added some form of risk (cf. Tversky and Wakker, 1995).

Tversky (1969) challenged EUT's *transitivity axiom* by observing violations in stated preference orders. Contrasting pairwise five different gambles, he showed that many subjects do not follow an absolute preference order but prefer one gamble over the other depending on the pair configuration:[71] Tversky reported the following preference relations for the greater part of his sample: $a \succsim b$, $b \succsim c$, $c \succsim d$, $d \succsim$

[71]The original paper presented probabilities in the form of pie charts.

Problem 9

Gamble	Probability of winning	Payoff in $
a	7/24	5.00
b	8/24	4.75
c	9/24	4.50
d	10/24	4.25
e	11/24	4.00

e, but $e \gtrsim a$. The standard explanation for this cyclical preference order are similarity effects (Rubinstein, 1988). When comparing similar probabilities in the interior value domain (i.e., excluding 0 and 1), people tend to choose the riskier alternative. When probabilities are ostensibly different, as in the case of choosing between a and e, individuals behave risk-averse by preferring e over a. Like the contrast of distinct (similar) probabilities evokes (no) risk aversion in people, the length of the short subscription likely switches on and off certain criteria in decision-making, such as marginal evaluation or regret.

Among others, the described paradoxes in Problem 7-9 paved the way for the notion of non-linear probability weighting, later formalized in prospect theory (see Section 4.3.3). Beyond that, Tversky headed a further stream, which centers on the role of the descriptive context in choice problems:

Problem 10 Imagine your country is preparing for the outbreak of an Asian disease expected to kill 600 people. The government can choose between two programs to combat the disease:

A: If this program is adopted, 200 people will be saved. [72%]
B: If this program is adopted, there is a 1/3 probability that 600 people will be saved, and a 2/3 probability that no people will be saved. [28%]

A′: If this program is adopted, 400 people will die. [22%]
B′: If this program is adopted, there is a 1/3 probability that nobody will die, and a 2/3 probability that 600 people will die. [78%]

A *framing effect* emerges whenever a DM switches preferences, depending on whether choices are framed positively as in A vs. B, or negatively as in A′ vs. B′. Tversky and Kahneman (1986) found that most people choose A in the positive setting whereas B′ in the loss frame. This violates the *completeness principle* of EUT, stating that if a choice A is preferred to B, B(=B′) cannot be preferred to A(=A′). Put differently, framing determines the perspective on a matter.

[71] The issue of external validity of hypothetical questions is addressed in the discussion.

Birnbaum et al. (1992) asked subjects to indicate the minimum (maximum) price for a traded good if they had to advise a friend being the respective seller (buyer). Alone by manipulating the subjects' point of view, recommended sale prices far exceeded recommended purchase prices. The selling perspective arguably evokes an *endowment effect* (Knetsch, 1989) in subjects typically caused by loss aversion, which leads to an increased evaluation. The Asian disease question in Problem 10 represents the canonical example to illustrate how framing triggers loss aversion in people. Yet, the construct of framing can be taken one step further by annulling the *dominance criterion* in rational choice theory (Tversky and Kahneman, 1986):

Problem 11 Consider the following two lotteries, described by the percentage of marbles of different colors in each urn and the amount of money you win or lose depending on the color of a randomly drawn marble. Which lottery do you prefer:

A:	90% white	6% red	1% green	1% blue	2% yellow
	$0	win $45	win $30	lose $15	lose $15
B:	90% white	6% red	1% green	1% blue	2% yellow
	$0	win $45	win $45	lose $10	lose $15
A′:	90% white	6% red	1% green	3% yellow	-
	$0	win $45	win $30	lose $15	-
B′:	90% white	7% red	1% green	2% yellow	-
	$0	win $45	lose $10	lose $15	-

B clearly dominates A in the first lottery set. Instead, in the latter choice between A′ and B′, a lottery comparison seems fairly more complicated when marble ratios do not coincide. Most subjects ignore the fact that the primed lotteries are concealed versions of the original lotteries. Likewise, one can regard a week-long or even a 6-month subscription as a masked version of single-use given re-use likelihood is negligible. While the concealment in Problem 11 leads to violation of dominance within the same lottery (within-subject), Problem 1 denotes a case of transitivity violation across lotteries (between-subject). Framing in the classical sense cannot explain the present paradox since the descriptive context does not vary between treatments. Nevertheless, it seems evident that the short alternative alters a DM's perspective, similar to Birnbaum et al.'s (1992) manipulation above.

To be precise, the notion of framing refers to the violation of two different principles in standard choice theory - *description invariance* (Problems 10 and 11) and *procedure invariance*. The literature considers them related but disparate (e.g., Tversky et al., 1988). Whereas description invariance adverts the common framing effect, procedure invariance involves differences stemming from the applied evaluation scale. Lichtenstein and Slovic (1971) were the first to document so-called *choice-judgment* PRs. Typically, low-payoff/high-probability gambles (e.g., \$12 with 90%, \$0 otherwise) are preferred to high-payoff/low-probability gambles (\$96 with 15%, \$0 otherwise). Yet, asked for their minimum selling price, subjects often state higher amounts for the latter. It is argued that in a mere choice evaluation, the lower risk of the high-probability option appeals to common risk aversion in people, while a price evaluation directs at the price attribute of gambles and therefore makes the high-payoff lottery seem superior (compatibility principle; Slovic et al., 1990).

Hsee (1996) presents a different type of PR, subject to the same evaluation scale but to a different *evaluation mode* (Problem 12). Asked for their WTP, subjects indicate higher amounts for dictionary A than for dictionary B in separate evaluation (between-subject), vice versa in joint evaluation (within-subject).

Problem 12

	Dictionary A	Dictionary B
Year of publication	1993	1993
Number of entries	10,000	20,000
Any defects?	No, it's like new.	Yes, the cover is torn; otherwise it's like new.

The author proposes the *evaluability hypothesis*, according to which joint-separate evaluation PRs occur when a good features at least one attribute that is hard to assess independently and one that allows easy, independent evaluation. In absence of any reference point, the number of entries represents a rather difficult criterion to assess the value of a dictionary. Instead, judging a dictionary's value on obvious defects such as the condition of its cover seems a much easier task. What obvious defects are in the case of dictionaries may be marginal valuation in the field of subscriptions and insurances. Whereas the straightforward value of single-use arguably facilitates rational assessment, a 6-month subscription complicates evaluation through the incorporation of risk. It therefore opens room for other criteria that are easy to assess.

Findings in Irwin et al. (1993) support the notion of pigeonholing along similar lines. Subjects indicate their WTP for various goods, including air quality in Denver and improvement in a VCR. Air quality is considered more relevant in joint evaluation but reaches lower WTP-values in separate evaluation. The single treatment in this paper compares to their joint evaluation condition in which expected utility, in rational terms, is the dominant decision attribute. Instead, the weekly or even monthly framing evokes an assessment similar to the separate evaluation in which the most relevant criterion is blurred in favor of rationally less important attributes.

Joint-separate evaluation PRs are closely linked to another type of context-dependent PRs. The *decoy* or *menu effect* postulates that an ex-ante difficult choice between two undominated goods can be manipulated towards either direction by adding a dominated good to the choice menu (e.g., Television C in Problem 13). In line with the evaluability hypothesis, it ascribes the presence of a dominated alternative (decoy) the effect of letting its dominant counterpart (Television B) appear superior to all other alternatives.

Problem 13

	Television A	Television B	*Television C*
Screen size	30 inch	40 inch	*35 inch*
Price	$800	$1,200	*$1,300*

The following quasi joint-separate evaluation PR and pure form of dominance violation represents another puzzling observation in Birnbaum et al. (1992). Asked for their WTP in a between-subject design, individuals express a higher value for the dominated lottery that pays $96 with 95% chance and $0 with 5%, than for the dominant lottery that pays the same upside of $96 with 95% but a higher downside of $24 in the remaining 5%. The authors attribute underweighting of zero outcomes to this phenomenon. Similar to the pigeonholing PR, adding a greater downside value to the lottery at no costs reduces its attraction. Birnbaum (1974) configural weight model, analyzed in Sections 4.3.4 and 4.B.1, accommodates both the *zero-outcome effect* and the before mentioned price discrepancy resulting from distinct perspectives of a DM.

On a different note but no less insightful to the discussion, Mellers et al. (1992) brought forward that the degree of uncertainty about another person's likableness is reflected in the variance of corresponding likableness ratings. According to this relationship, lower variance in WTP-values for the week treatment across all constructs in Study 4 can be a signal for an unconsciously more guided thought process when single-use is absent in the choice menu. It could be that evaluation formation in the treatment condition is governed by the benchmark of the long subscription, adopting the simple heuristic of linear down-scaling to make the long subscription comparable to the short one. Greater variance among WTP-values in the single treatment points towards a more heterogeneous assessment process depending on personal factors[72] that, compared to the treatment condition, are not overwritten.

[72] Controlling for re-use estimates, the difference in SD between treatments falls only marginally.

Correlation analysis finds support for this idea in seven of the nine examined constructs in Studies 3 and 4.

4.5.2 Semantic evidence

The common theme in the reviewed PRs, except for zero underweighting, is varying prominence between two decision aspects. Instead, this paper attributes the root of the PH phenomenon to a third dimension. The length of the shorter alternative determines the evaluation scale endogenously. An exploratory semantic analysis of participants' described strategies mirrors this assumption.

Feedback from the week treatment frequently mentions the price-value aspect, presumably referring to the theoretical market value of a subscription which is inseparably connected to biased likelihood perceptions: *"I tried to look at the price-value ratio of each offer. In most cases, the difference in cost was so low (compared to what's offered), so I believed it was worth the additional cost, as I may have used the service again."* Similarly, others overlook a potential utility advantage of the short option due to marginal price differences and biased likelihood estimates: *"... difference between the long and short term costs were so minimal that I simply felt it would be reasonable to pay a little more and enjoy the services for longer in the remote event I might need them."* And still others show signs of regret aversion: *"... for this kind of things it's always 'better be safe than sorry.' For the other questions I picked the options that lasted longer because the price increase wasn't very much ..."* Though, the second part of the same comment proves that probability biases likely vary in context: *"... only exception being the New York question where I thought it would be basically impossible for me to go back to NYC in the next 6 months."* One participant epitomizes in a textbook-like manner the dual-process framework of decision-making (Frederick, 2005; Li et al., 2017a), according to which individual cognitive ability reveals whether intuitive thoughts can be suppressed and finally overwritten by logical solutions: *"I have computed the price divided by every single week to compare the advantage, later I have thought whether I would actually use the service again."*[73]

On the other hand, comments from the single treatment often contain terms such as "likelihood" and "rational value", supposedly putting absolute utility gains into context: *"Price and likelihood of using the service more than one time"*. The signal word "never" captures the rational assessment of re-use likelihood: *"Because most of the time I would have never used the service again."* In the same vein, this one-word com-

[73]Translated from Italian: "Ho calcolato il prezzo diviso per ogni settimana e ho calcolato il vantaggio, successivamente ho pensato se quel servizio lo avrei utilizzato ancora oppure no."

ment highlights the prescribed treatment effect: *"Rationality."* Yet, in some cases, extreme events such as natural disasters (e.g., floods) are possibly too strong a trigger for risk aversion despite the mitigating single-use effect: *"I choose the short subscription because to need that service again is not so probable. Only exceptions are for very small variation in price (1 euro or less) or for situations where it's better to be protected in case of natural disasters that may happen regularly."*

The latter comment also illustrates that marginal price differences can work in both ways when designing a pigeonholing trap. On the one hand, differences ought to be small enough to render the short alternative in the week treatment trivial. On the other hand, they need to prompt saving incentives to choose the single-use option in the control. The window for a PH effect is narrow if provoked exogenously and the manifestation highly subjective. This reflects in the much more robust evidence of the price-matching task when price parameters are set endogenously by the subjects.

4.5.3 Implications for marketers and policymakers

Depending on the desired outcome, it can be advantageous to model a choice set promoting or preventing the PH effect. It is common practice in marketing to lure consumers into subscription traps through specifically designed product menus, similar to those herein (see real-market examples in Appendix 4.E).[74]Comparing subscription providers of movies and series finds that only a few offer single-use options. An illustrative example represents the niche-movie platform Dafilms (2020). Before the COVID-19 pandemic, it has featured a one-movie option alongside a yearly subscription. It is mere speculation to link the subsequent replacement of the one-movie option with a monthly subscription to profit maximization considering the ballooning demand during the outbreak (Appendix 4.E, Fig 4.1 Ofcom.org, 2020). Even though this incidence can be explained with the PH hypothesis, we do not impute the platform to have consciously exploited the pigeonholing effect.

When the market does not address single-users (market failure), policy interventions could countersteer PH effects by imposing single-use offers on providers. Yet, awareness of this cognitive bias may also benefit regulators as a tool to nudge people into socially desirable, long-term services (e.g., basic health insurance).

[74]In fact, the 2017 price menu of BlaBlaCar led to the idea behind the pigeonholing effect.

4.5.4 Concluding remarks

The pigeonholing effect conjectures that the availability of a sufficiently brief option in a two-choice menu spurs evaluation in terms of rational utility, whereas analogous absence induces less rational measures in a judge. Hinging on the concept of transaction utility (Thaler, 1985) and availability (Tversky and Kahneman, 1973), it provides an intuitively plausible and economically meaningful argument. In the context of subscriptions and insurances, this can come about preferring a dominated prospect over its dominant complement in separate evaluation against a third option. Studies 1 to 5 investigate this phenomenon with varying results. While Studies 2 and 4b find no evidence of a PH effect, the classroom experiments of Study 1 and the matching questions in Study 4a show significant violation of monotonicity.

These findings raise various questions. First and foremost: which conditions favor the behavioral paradox? As with all monotonicity and transitivity violations, the PH effect holds only for a specific parameter range and is very context-sensitive.[75] The parameter influence on the PH effect surfaces particularly in Studies 3 and 4. Despite posing the same price-matching question with identical subscription parameters, the medians of subjects' indicated WTP vary tremendously between studies (Table 4.3). Moreover, it is worthwhile noting that the matching questions robustly evoke the PH effect but less so the basic choice method. This finding emphasizes the procedure and parameter sensitivity of the cognitive bias studied in this paper (see, e.g., Tversky et al., 1988).

It is the purpose of this work to shape the understanding of its psychological root. The model in Section 4.3 undertakes a first analytical approach to determine promoting factors. In line with the considerations underlying regret theory (Bell, 1982; Loomes and Sugden, 1982), also indirectly incorporated in salience theory (Bordalo et al., 2012), it seems reasonable that the PH effect is limited to cases that deal with marginal utility differences between alternatives. These differences can result from minimal price discrepancies, and more importantly, from a low probability of using a subscription again. While all studies identify low re-use likelihood as the main driver of the PH effect,[76] Study 5, a robustness check for parameter sensitivity, finds no evidence of a relationship between relative price differences and

[75] One participant, although selecting the short option in all but the painting question, commented that he only deviated in this question because he loves paintings so much.

[76] Nevertheless, Study 5 states low re-use likelihood to be neither a necessary nor a sufficient condition. The flood construct shows signs of a PH effect both in Study 5a (marginal evidence) and Study 5b (mild evidence). Yet, the corresponding interacted probability/treatment term exhibits mild statistical significance in the direction predicted by the PH hypothesis in Study 5a, but points, though not significantly, the opposite direction in Study 5b.

the PH phenomenon, calling for validation in future research.

Also, other robustness tests of the PH hypothesis are plausible. Related to the literature on decoy effects, one could study the impact of phantom choices. Adding a crossed out single-use subscription to a choice menu without single-use should weaken the PH effect. Further variations could brief subjects on bounded rationality or permit communication among subjects and test for reduced susceptibility to pigeonholing. Li et al. (2017b) show that asking subjects to put themselves first in the shoes of others made them reach more rational decisions in a choice problem.

Similarly interesting is whether a weekly subscription presented as single-use plus free use over the following week affects preferences differently. What is known as state-splitting does not alter the subscription assessment by prospect theory (Tversky and Kahneman, 1992) or salience theory (Bordalo et al., 2012), though the herein proposed model shows how state-splitting can cause preference reversals. Future work needs to refine a more specific decision model to increase its external validity. Furthermore, the current analysis confines to hypothetical questions in accord with its advocates in the evergreen debate, whether they elicit preferences for risky prospects equally good as incentivized questions (see, e.g., Tversky and Kahneman, 1992). This claim can undergo a stress test in future research.

Appendix 4.A Case study for $d_b = 2, d_l = 3$

Problem 1 Imagine you are visiting New York City over the Christmas break. As last item on your agenda before flying back home, you have saved the panorama view from the Rockefeller Center. At the entrance, you have to register with your ID card and choose between two alternatives.

L_b: $d_b = 2$ days free entrance for p_b

L_l: $d_l = 3$ days free entrance for p_l

$$\textit{State probabilities}\ \ \pi_s = \begin{cases} \pi_1 = \pi_d & \textit{(next visit on day 2)} \\ \pi_2 = (1-\pi_d)\cdot\pi_d & \textit{(next visit on day 3)} \\ \pi_3 = (1-\pi_d-(1-\pi_d)\cdot\pi_d), & \textit{(next visit after day 3)} \end{cases}$$

Without specifying the utility of using the service, independent state payoffs for each lottery cannot be determined. One solution represents the computation of relative state payoffs. Since L_l remains present in both subscription sets, it seems reasonable setting the long subscription L_l as the reference point (zero utility) for a relative payoff computation. In line with the premise in salience and prospect theory stating that judges "focus on payoffs, rather than on absolute wealth levels, when evaluating risky prospects" (Bordalo et al., 2012, p. 1246), we denote relative payoffs by r_s^L to evaluate subscription $L \in \{L_b, L_l\}$ in state $s \in \{1,2,3\}$.

$$r_s^L = \begin{cases} r_1^{L_b} = p_l - p_b - p_l\pi_d + p_l(1-\pi_d)\pi_d + p_l(\pi_d)^2 + ... \ ; & r_1^{L_l} = 0 \\ r_2^{L_b} = -p_b + p_l(1-\pi_d)\pi_d + p_l\pi_d + ... \ ; & r_2^{L_l} = 0 \\ r_3^{L_b} = p_l - p_b; & r_3^{L_l} = 0 \end{cases} \tag{4.13}$$

Note that relative payoffs in equation (4.13) are conditioned on the assumption that L_l is purchased in a follow-up choice of the same problem if renewal is necessary. In state 1 and relative to L_l, the short subscription L_b pays the difference of subscription prices $(p_l - p_b)$ minus the expected costs $(-p_l\pi_d)$ of having to take out the long subscription in two days if service use will occur then, plus the expected cost advantage of not having to purchase L_l if the service is used the next time on day 4 ($p_l(1-\pi_d)\pi_d$) or the next time on day 5 ($p_l\pi_d^2$) and so on. A long subscription L_l taken out on day 4 or day 5 would leverage added-value influencing relative payoffs in the opposite direction. One can easily grasp that each state realization other than state 3 leads to follow-up subscriptions generating an infinitely large payoff structure. For small π_d we see that the generic model in equation (4.1) is a suitable approximation of the complete payoff structure. Relative payoffs in equation (4.14) reflect the opposite case, i.e., L_b is preferred in a follow-up choice.

$$r_s^L = \begin{cases} r_1^{L_b} = p_l - p_b - p_b\pi_d + p_b(1-\pi_d)\pi_d + p_b(\pi_d)^2 + ... \ ; & r_1^{L_l} = 0 \\ r_2^{L_b} = -p_b + p_b(1-\pi_d)\pi_d + p_b\pi_d + ... \ ; & r_2^{L_l} = 0 \\ r_3^{L_b} = p_l - p_b; & r_3^{L_l} = 0 \end{cases} \tag{4.14}$$

Appendix 4.B Proofs

4.B.1 Configural weight model

Birnbaum et al. (1992) use the below expression for three-outcome gambles

$$U_V(x_1, \pi_1, x_2, \pi_2, x_3, \pi_3\text{=1-}\pi_1\text{-}\pi_2) = \frac{Au(x_1) + Bu(x_2) + Cu(x_3)}{A + B + C},$$

where $C = (1 - a_v)\big(1 - S_{x_1}(1 - \pi_3)\big)$. Note that the transformation of payoffs, defined in Section 4.3.1, to the configural weight model requires the assumption of linear utility. By normalizing $x_1^{L_b'} - x_1^{L_l} = -p_b = 0$, we get $x_1^{L_b} - x_1^{L_l} = p_l - p_b = p_l$, and $x_3^{L_b} - x_3^{L_l} = p_l - p_b = p_l$.[77] Plugging the derived

$$\text{outcomes} = \begin{cases} x_1 = p_l; & x'_1 = 0 \\ x_2 = 0; & x'_2 = 0 \\ x_3 = p_l; & x'_3 = p_l \end{cases}$$

into the model, we have a pigeonholing PR if

$$\begin{aligned} U_V(0, \pi_1, 0, \pi_2, x'_3, \pi_3) =& \frac{A_0 u(0) + B_0 u(0) + C_0 u(x'_3)}{A_0 + B_0 + C_0} \\ > U_V(x_1, \pi_1, 0, \pi_2, x_3, \pi_3) =& \frac{A_{x_1} u(x_1) + B_{x_1} u(0) + C_{x_1} u(x_3)}{A_{x_1} + B_{x_1} + C_{x_1}} \end{aligned}$$

$$\frac{C_0}{A_0 + B_0 + C_0} > \frac{A_{x_1} + C_{x_1}}{A_{x_1} + B_{x_1} + C_{x_1}}$$

$$\begin{aligned} & \frac{(1-a_v)(1-0.74(\pi_1+\pi_2)-0.14)}{a_v(0.74\pi_1+0.14)+(1-a_v)(1-0.74(1-\pi_2)-0.14)+(1-a_v)(1-0.74(\pi_1+\pi_2)-0.14)} \\ = & \frac{0.86-0.74\pi_1-0.74\pi_2-0.86a_v+0.74a_v\pi_1+0.74a_v\pi_2}{1.48a_v\pi_1+0.98-0.74\pi_1-0.84a_v} \\ > & \frac{a_v(0.59\pi_1+0.29)+(1-a_v)(1-0.59(\pi_1+\pi_2)-0.29)}{a_v(0.59\pi_1+0.29)+(1-a_v)(1-0.59(1-\pi_2)-0.29)+(1-a_v)(1-0.59(\pi_1+\pi_2)-0.29)} \\ = & \frac{1.18a_v\pi_1+0.71-0.59\pi_1-0.59\pi_2-0.42a_v+0.59a_v\pi_2}{1.18a_v\pi_1+0.83-0.59\pi_1-0.54a_v} \end{aligned}$$

$$\begin{aligned} & 2.536a_v\pi_1+0.7138-1.121\pi_1-1.1782a_v-1.3098a_v\pi_1^2+0.4366\pi_1^2-1.3098a_v\pi_1\pi_2-0.6142\pi_2 \\ & +0.4366\pi_1\pi_2-1.4144a_v^2\pi_1+0.4644a_v^2+0.8732a_v^2\pi_1^2+0.8732a_v^2\pi_1\pi_2+1.0138a_v\pi_2-0.3996a_v^2\pi_2 \\ > & 1.7464a_v^2\pi_1^2+3.0136a_v\pi_1-1.7464a_v\pi_1^2-1.6128a_v^2\pi_1+0.6958-1.1036\pi_1-1.008a_v+0.4366\pi_1^2 \\ & -1.3098a_v\pi_1\pi_2-0.5782\pi_2+0.4366\pi_1\pi_2+1.0738a_v\pi_2+0.3528a_v^2+0.8732a_v^2\pi_1\pi_2-0.4956a_v^2\pi_2 \end{aligned}$$

$$\begin{aligned} & -0.4776a_v\pi_1+0.018-0.018\pi_1-0.1702a_v+0.4366a_v\pi_1^2-0.036\pi_2 \\ & +0.1984a_v^2\pi_1+0.1116a_v^2-0.8732a_v^2\pi_1^2-0.06a_v\pi_2+0.096a_v^2\pi_2 > 0 \end{aligned}$$

$$\begin{aligned} & a_v^2(-0.8732\pi_1^2+0.1984\pi_1+0.096\pi_2+0.1116) \\ & +a_v(0.4366\pi_1^2-0.4776\pi_1-0.06\pi_2-0.1702)+0.018(-\pi_1-2\pi_2+1) > 0 \end{aligned}$$

[77] Compared to salience theory (see Section 4.3.5), computing absolute differences between gambles in the congifural weight model through zero-normalization does not afford discriminatory analysis of follow-up preferences, L_b or L_l, for our context.

4.B.2 Salience theory

$\frac{p_l}{p_b} - \frac{p_b}{p_l} < 1$, follow-up subscription L_l

$$
\begin{aligned}
V(L_b - L_l) &= \frac{(1-(1-\pi_d)^{d_b-1} + (1-\pi_d)^{d_l-1})\delta^2(p_l - p_b) + ((1-\pi_d)^{d_b-1} - (1-\pi_d)^{d_l-1})\delta^1(-p_b)}{(1-(1-\pi_d)^{d_b-1} + (1-\pi_d)^{d_l-1})\delta^2 + ((1-\pi_d)^{d_b-1} - (1-\pi_d)^{d_l-1})\delta^1} \\
&= \frac{(p_l - p_b)\delta - (1-\pi_d)^{d_b-1}((p_l - p_b)\delta + p_b) + (1-\pi_d)^{d_l-1}((p_l - p_b)\delta + p_b)}{\delta + (1-\delta)(1-\pi_d)^{d_b-1} - (1-\delta)(1-\pi_d)^{d_l-1}} \\
V(L_{b'} - L_l) &= \frac{(1-(1-\pi_d)^{d_b-1})\delta^1(-p_b) + ((1-\pi_d)^{d_b-1} - (1-\pi_d)^{d_l-1})\delta^1(-p_b) + (1-\pi_d)^{d_l-1}\delta^2(p_l - p_b)}{(1-(1-\pi_d)^{d_b-1})\delta^1 + ((1-\pi_d)^{d_b-1} - (1-\pi_d)^{d_l-1})\delta^1 + (1-\pi_d)^{d_l-1}\delta^2} \\
&= \frac{-p_b + (1-\pi_d)^{d_l-1}((p_l - p_b)\delta + p_b)}{1-(1-\delta)(1-\pi_d)^{d_l-1}}
\end{aligned}
$$

$$
\begin{aligned}
V(L_b - L_l) &> V(L_{b'} - L_l) \\
\frac{(p_l - p_b)\delta - (1-\pi_d)^{d_b-1}((p_l - p_b)\delta + p_b) + (1-\pi_d)^{d_l-1}((p_l - p_b)\delta + p_b)}{\delta + (1-\delta)(1-\pi_d)^{d_b-1} - (1-\delta)(1-\pi_d)^{d_l-1}} &> \frac{-p_b + (1-\pi_d)^{d_l-1}((p_l - p_b)\delta + p_b)}{1-(1-\delta)(1-\pi_d)^{d_l-1}}
\end{aligned}
$$

$$
\begin{aligned}
& \underbrace{(p_l - p_b)\delta} - \underbrace{(1-\pi_d)^{d_b-1}((p_l - p_b)\delta + p_b)} + \underbrace{(1-\pi_d)^{d_l-1}((p_l - p_b)\delta + p_b)} - \underbrace{(p_l - p_b)\delta(1-\delta)(1-\pi_d)^{d_l-1}} \\
& + \underbrace{((p_l - p_b)\delta + p_b)(1-\delta)(1-\pi_d)^{d_l+d_b-2}} - \underbrace{((p_l - p_b)\delta + p_b)(1-\delta)(1-\pi_d)^{2d_l-2}} \\
> \quad & - \underbrace{p_b\delta} - \underbrace{p_b(1-\delta)(1-\pi_d)^{d_b-1}} + \underbrace{p_b(1-\delta)(1-\pi_d)^{d_l-1}} + \underbrace{\delta(1-\pi_d)^{d_l-1}((p_l - p_b)\delta + p_b)} \\
& + \underbrace{(1-\pi_d)^{d_l+d_b-2}((p_l - p_b)\delta + p_b)(1-\delta)} - \underbrace{(1-\pi_d)^{2d_l-2}((p_l - p_b)\delta + p_b)(1-\delta)}
\end{aligned}
$$

$$
\begin{aligned}
p_l\delta - (1-\pi_d)^{d_b-1}p_l\delta &> 0 \\
1 &> (1-\pi_d)^{d_b-1} \\
1 &> 1-\pi_d \\
\pi_d &> 0
\end{aligned}
$$

$\frac{p_l}{p_b} - \frac{p_b}{p_l} > 1$, follow-up subscription L_l

$$
\begin{aligned}
V(L_b - L_l) &= \frac{(1-(1-\pi_d)^{d_b-1})\delta^1(p_l - p_b) + ((1-\pi_d)^{d_b-1} - (1-\pi_d)^{d_l-1})\delta^2(-p_b) + (1-\pi_d)^{d_l-1}\delta^1(p_l - p_b)}{(1-(1-\pi_d)^{d_b-1})\delta^1 + ((1-\pi_d)^{d_b-1} - (1-\pi_d)^{d_l-1})\delta^2 + (1-\pi_d)^{d_l-1}\delta^1} \\
&= \frac{(p_l - p_b) - (1-\pi_d)^{d_b-1}(p_l - p_b + p_b\delta) + (1-\pi_d)^{d_l-1}(p_l - p_b + p_b\delta)}{1-(1-\pi_d)^{d_b-1}(1-\delta) + (1-\pi_d)^{d_l-1}(1-\delta)} \\
V(L_{b'} - L_l) &= \frac{(1-(1-\pi_d)^{d_b-1})\delta^2(-p_b) + ((1-\pi_d)^{d_b-1} - (1-\pi_d)^{d_l-1})\delta^2(-p_b) + (1-\pi_d)^{d_l-1}\delta^1(p_l - p_b)}{(1-(1-\pi_d)^{d_b-1})\delta^2 + ((1-\pi_d)^{d_{b'}-1} - (1-\pi_d)^{d_l-1})\delta^2 + (1-\pi_d)^{d_l-1}\delta^1} \\
&= \frac{-p_b\delta + (1-\pi_d)^{d_l-1}(p_b\delta + p_l - p_b)}{\delta + (1-\pi_d)^{d_l-1}(1-\delta)}
\end{aligned}
$$

$$
\begin{aligned}
V(L_b - L_l) &> V(L_{b'} - L_l) \\
\frac{(p_l - p_b) - (1-\pi_d)^{d_b-1}(p_l - p_b + p_b\delta) + (1-\pi_d)^{d_l-1}(p_l - p_b + p_b\delta)}{1-(1-\pi_d)^{d_b-1}(1-\delta) + (1-\pi_d)^{d_l-1}(1-\delta)} &> \frac{-p_b\delta + (1-\pi_d)^{d_l-1}(p_b\delta + p_l - p_b)}{\delta + (1-\pi_d)^{d_l-1}(1-\delta)}
\end{aligned}
$$

$$
\begin{aligned}
&(p_l - p_b)\delta + (p_l - p_b)(1-\pi_d)^{d_l-1}(1-\delta) - (1-\pi_d)^{d_b-1}(p_l - p_b + p_b\delta)\delta \\
&- (1-\pi_d)^{d_l+d_b-2}(p_l - p_b + p_b\delta)(1-\delta) + (1-\pi_d)^{d_l-1}(p_l - p_b + p_b\delta)\delta + (1-\pi_d)^{2d_l-2}(p_l - p_b + p_b\delta)(1-\delta) \\
> \quad &- p_b\delta + (1-\pi_d)^{d_b-1}(1-\delta)p_b\delta - (1-\pi_d)^{d_l-1}(1-\delta)p_b\delta + (1-\pi_d)^{d_l-1}(p_b\delta + p_l - p_b) \\
&- (1-\pi_d)^{d_l+d_b-2}(p_b\delta + p_l - p_b)(1-\delta) + (1-\pi_d)^{2d_l-2}(p_b\delta + p_l - p_b)(1-\delta)
\end{aligned}
$$

$$
\begin{aligned}
p_l\delta - (1-\pi_d)^{d_b-1}p_l\delta &> 0 \\
1 &> (1-\pi_d)^{d_b-1} \\
1 &> 1-\pi_d \\
\pi_d &> 0
\end{aligned}
$$

$\frac{p_l}{p_b} < \sqrt{2}$, follow-up subscription L_b

$$
\begin{aligned}
V(L_b - L_l) &= \frac{(1-(1-\pi_d)^{d_b-1} + (1-\pi_d)^{d_l-1})\delta(p_l - p_b) + ((1-\pi_d)^{d_b-1} - (1-\pi_d)^{d_l-1})(-2p_b + p_l)}{(1-(1-\pi_d)^{d_b-1} + (1-\pi_d)^{d_l-1})\delta + ((1-\pi_d)^{d_b-1} - (1-\pi_d)^{d_l-1})} \\
&= \frac{(p_l - p_b)\delta - (1-\pi_d)^{d_b-1}((p_l - p_b)\delta + 2p_b - p_l) + (1-\pi_d)^{d_l-1}((p_l - p_b)\delta + 2p_b - p_l)}{\delta + (1-\delta)(1-\pi_d)^{d_b-1} - (1-\delta)(1-\pi_d)^{d_l-1}} \\
V(L_{b'} - L_l) &= \frac{(1-(1-\pi_d)^{d_l-1})(p_l - 2p_b) + (1-\pi_d)^{d_l-1}\delta(p_l - p_b)}{(1-(1-\pi_d)^{d_l-1}) + (1-\pi_d)^{d_l-1}\delta} \\
&= \frac{p_l - 2p_b - (1-\pi_d)^{d_l-1}(p_l - 2p_b - \delta(p_l - p_b))}{1-(1-\delta)(1-\pi_d)^{d_l-1}}
\end{aligned}
$$

$$
V(L_b - L_l) > V(L_{b'} - L_l)
$$

$$
\begin{aligned}
&\frac{(p_l - p_b)\delta - (1-\pi_d)^{d_b-1}((p_l - p_b)\delta + 2p_b - p_l) + (1-\pi_d)^{d_l-1}((p_l - p_b)\delta + 2p_b - p_l)}{\delta + (1-\delta)(1-\pi_d)^{d_b-1} - (1-\delta)(1-\pi_d)^{d_l-1}} \\
> \quad &\frac{p_l - 2p_b - (1-\pi_d)^{d_l-1}(p_l - 2p_b - \delta(p_l - p_b))}{1-(1-\delta)(1-\pi_d)^{d_l-1}}
\end{aligned}
$$

$$
\begin{aligned}
&(p_l - p_b)\delta - (1-\pi_d)^{d_b-1}((p_l - p_b)\delta + 2p_b - p_l) + (1-\pi_d)^{d_l-1}((p_l - p_b)\delta + 2p_b - p_l) - (p_l - p_b)\delta(1-\delta)(1-\pi_d)^{d_l-1} \\
&+ ((p_l - p_b)\delta + 2p_b - p_l)(1-\delta)(1-\pi_d)^{d_l+d_b-2} - ((p_l - p_b)\delta + 2p_b - p_l)(1-\delta)(1-\pi_d)^{2d_l-2} \\
> \quad &(p_l - 2p_b)\delta - (2p_b - p_l)(1-\delta)(1-\pi_d)^{d_b-1} + (2p_b - p_l)(1-\delta)(1-\pi_d)^{d_l-1} + \delta(1-\pi_d)^{d_l-1}((p_l - p_b)\delta + 2p_b - p_l) \\
&+ (1-\pi_d)^{d_l+d_b-2}((p_l - p_b)\delta + 2p_b - p_l)(1-\delta) - (1-\pi_d)^{2d_l-2}((p_l - p_b)\delta + 2p_b - p_l)(1-\delta)
\end{aligned}
$$

$$
\begin{aligned}
p_b\delta - (1-\pi_d)^{d_b-1}p_b\delta &> 0 \\
1 &> (1-\pi_d)^{d_b-1} \\
1 &> 1-\pi_d \\
\pi_d &> 0
\end{aligned}
$$

$\frac{p_l}{p_b} > \sqrt{2}$, follow-up subscription L_b

$$
\begin{aligned}
V(L_b - L_l) &= \frac{(1-(1-\pi_d)^{d_b-1} + (1-\pi_d)^{d_l-1})(p_l - p_b) + ((1-\pi_d)^{d_b-1} - (1-\pi_d)^{d_l-1})\delta(p_l - 2p_b)}{(1-(1-\pi_d)^{d_b-1} + (1-\pi_d)^{d_l-1}) + ((1-\pi_d)^{d_b-1} - (1-\pi_d)^{d_l-1})\delta} \\
&= \frac{(p_l - p_b) - (1-\pi_d)^{d_b-1}(p_l - p_b + (2p_b - p_l)\delta) + (1-\pi_d)^{d_l-1}(p_l - p_b + (2p_b - p_l)\delta)}{1-(1-\pi_d)^{d_b-1}(1-\delta) + (1-\pi_d)^{d_l-1}(1-\delta)} \\
V(L_{b'} - L_l) &= \frac{(1-(1-\pi_d)^{d_l-1})\delta(p_l - 2p_b) + (1-\pi_d)^{d_l-1}(p_l - p_b)}{(1-(1-\pi_d)^{d_l-1})\delta + (1-\pi_d)^{d_l-1}} \\
&= \frac{(-2p_b + p_l)\delta + (1-\pi_d)^{d_l-1}((2p_b - p_l)\delta + p_l - p_b)}{\delta + (1-\pi_d)^{d_l-1}(1-\delta)}
\end{aligned}
$$

$$
V(L_b - L_l) > V(L_{b'} - L_l)
$$

$$
\begin{aligned}
& \frac{(p_l - p_b) - (1-\pi_d)^{d_b-1}(p_l - p_b + (2p_b - p_l)\delta) + (1-\pi_d)^{d_l-1}(p_l - p_b + (2p_b - p_l)\delta)}{1-(1-\pi_d)^{d_b-1}(1-\delta) + (1-\pi_d)^{d_l-1}(1-\delta)} \\
> \quad & \frac{(p_l - 2p_b)\delta + (1-\pi_d)^{d_l-1}((2p_b - p_l)\delta + p_l - p_b)}{\delta + (1-\pi_d)^{d_l-1}(1-\delta)}
\end{aligned}
$$

$$
\begin{aligned}
& \underbrace{(p_l - p_b)\delta}_{} + \underbrace{(p_l - p_b)(1-\pi_d)^{d_l-1}(1-\delta)}_{} - \underbrace{(1-\pi_d)^{d_b-1}(p_l - p_b + (2p_b - p_l)\delta)\delta}_{} \\
& - \underbrace{(1-\pi_d)^{d_l+d_b-2}(p_l - p_b + (2p_b - p_l)\delta)(1-\delta)}_{} + \underbrace{(1-\pi_d)^{d_l-1}(p_l - p_b + (2p_b - p_l)\delta)\delta}_{} + \underbrace{(1-\pi_d)^{2d_l-2}(p_l - p_b + (2p_b - p_l)\delta)(1-\delta)}_{} \\
> & - \underbrace{(2p_b - p_l)\delta}_{} + \underbrace{(1-\pi_d)^{d_b-1}(1-\delta)(2p_b - p_l)\delta}_{} - \underbrace{(1-\pi_d)^{d_l-1}(1-\delta)(2p_b - p_l)\delta}_{} + \underbrace{(1-\pi_d)^{d_l-1}((2p_b - p_l)\delta + p_l - p_b)}_{} \\
& - \underbrace{(1-\pi_d)^{d_l+d_b-2}((2p_b - p_l)\delta + p_l - p_b)(1-\delta)}_{} + \underbrace{(1-\pi_d)^{2d_l-2}((2p_b - p_l)\delta + p_l - p_b)(1-\delta)}_{}
\end{aligned}
$$

$$
\begin{aligned}
p_b\delta - (1-\pi_d)^{d_b-1}p_b\delta &> 0 \\
1 &> (1-\pi_d)^{d_b-1} \\
1 &> 1 - \pi_d \\
\pi_d &> 0
\end{aligned}
$$

4.B.3 Pigeonholing

$\frac{p_l}{p_b} < 2$, follow-up subscription L_l

$$
\begin{aligned}
V(L_b - L_l) &= \frac{(1-(1-\pi_d)^{d_b-1})(\frac{1}{d_b})^2(p_l - p_b) + ((1-\pi_d)^{d_b-1} - (1-\pi_d)^{d_l-1})(\frac{1}{d_b})^1(-p_b) + (1-\pi_d)^{d_l-1}(\frac{1}{d_b})^2(p_l - p_b)}{(1-(1-\pi_d)^{d_b-1})(\frac{1}{d_b})^2 + ((1-\pi_d)^{d_b-1} - (1-\pi_d)^{d_l-1})(\frac{1}{d_b})^1 + (1-\pi_d)^{d_l-1}(\frac{1}{d_b})^2} \\
&= \frac{(p_l - p_b)\frac{1}{d_b} - (1-\pi_d)^{d_b-1}((p_l - p_b)\frac{1}{d_b} + p_b) + (1-\pi_d)^{d_l-1}((p_l - p_b)\frac{1}{d_b} + p_b)}{\frac{1}{d_b} + (1-\frac{1}{d_b})(1-\pi_d)^{d_b-1} - (1-\frac{1}{d_b})(1-\pi_d)^{d_l-1}} \\
V(L_{b'} - L_l) &= \frac{(1-(1-\pi_d)^{d_b-1})(\frac{1}{d_{b'}})^1(-p_b) + ((1-\pi_d)^{d_b-1} - (1-\pi_d)^{d_l-1})(\frac{1}{d_{b'}})^1(-p_b) + (1-\pi_d)^{d_l-1}(\frac{1}{d_{b'}})^2(p_l - p_b)}{(1-(1-\pi_d)^{d_b-1})(\frac{1}{d_{b'}})^1 + ((1-\pi_d)^{d_b-1} - (1-\pi_d)^{d_l-1})(\frac{1}{d_{b'}})^1 + (1-\pi_d)^{d_l-1}(\frac{1}{d_{b'}})^2} \\
&= \frac{-p_b + (1-\pi_d)^{d_l-1}((p_l - p_b)\frac{1}{d_{b'}} + p_b)}{1 - (1-\frac{1}{d_{b'}})(1-\pi_d)^{d_l-1}}
\end{aligned}
$$

$$V(L_b - L_l) < V(L_{b'} - L_l)$$

$$\frac{(p_l - p_b)\frac{1}{d_b} - (1-\pi_d)^{d_b-1}((p_l - p_b)\frac{1}{d_b} + p_b) + (1-\pi_d)^{d_l-1}((p_l - p_b)\frac{1}{d_b} + p_b)}{\frac{1}{d_b} + (1-\frac{1}{d_b})(1-\pi_d)^{d_b-1} - (1-\frac{1}{d_b})(1-\pi_d)^{d_l-1}} < \frac{-p_b + (1-\pi_d)^{d_l-1}((p_l - p_b)\frac{1}{d_{b'}} + p_b)}{1 - (1-\frac{1}{d_{b'}})(1-\pi_d)^{d_l-1}}$$

$$\begin{aligned}
&\underbrace{(p_l - p_b)\frac{1}{d_b}} - \underbrace{(1-\pi_d)^{d_b-1}((p_l - p_b)\frac{1}{d_b} + p_b)} + \underbrace{(1-\pi_d)^{d_l-1}((p_l - p_b)\frac{1}{d_b} + p_b)} \\
&- \underbrace{(p_l - p_b)\frac{1}{d_b}(1-\frac{1}{d_{b'}})(1-\pi_d)^{d_l-1}} + \underbrace{((p_l - p_b)\frac{1}{d_b} + p_b)(1-\frac{1}{d_{b'}})(1-\pi_d)^{d_l+d_b-2}} - \underbrace{((p_l - p_b)\frac{1}{d_b} + p_b)(1-\frac{1}{d_{b'}})(1-\pi_d)^{2d_l-2}} \\
< &- \underbrace{p_b\frac{1}{d_b}} - \underbrace{p_b(1-\frac{1}{d_b})(1-\pi_d)^{d_b-1}} + \underbrace{p_b(1-\frac{1}{d_b})(1-\pi_d)^{d_l-1}} + \underbrace{\frac{1}{d_b}(1-\pi_d)^{d_l-1}((p_l - p_b)\frac{1}{d_{b'}} + p_b)} \\
&+ \underbrace{(1-\pi_d)^{d_l+d_b-2}((p_l - p_b)\frac{1}{d_{b'}} + p_b)(1-\frac{1}{d_b})} - \underbrace{(1-\pi_d)^{2d_l-2}((p_l - p_b)\frac{1}{d_{b'}} + p_b)(1-\frac{1}{d_b})}
\end{aligned}$$

$$\begin{aligned}
\frac{p_l}{d_b} - (1-\pi_d)^{d_b-1}\frac{p_l}{d_b} + p_l(\frac{1}{d_b} - \frac{1}{d_{b'}})(1-\pi_d)^{d_l+d_b-2} - p_l(\frac{1}{d_b} - \frac{1}{d_{b'}})(1-\pi_d)^{2d_l-2} &< 0 \\
1 - (1-\pi_d)^{d_b-1} + (1-\frac{d_b}{d_{b'}})(1-\pi_d)^{d_l+d_b-2} - (1-\frac{d_b}{d_{b'}})(1-\pi_d)^{2d_l-2} &< 0 \\
1 - (1-\pi_d)^{d_b-1} + (1-\frac{d_b}{d_{b'}})(1-\pi_d)^{d_l-2}((1-\pi_d)^{d_b} - (1-\pi_d)^{d_l}) &< 0
\end{aligned}$$

$\frac{p_l}{p_b} > 2$, follow-up subscription L_l

$$\begin{aligned}
V(L_b - L_l) &= \frac{(1-(1-\pi_d)^{d_b-1})(\frac{1}{d_b})^1(p_l - p_b) + ((1-\pi_d)^{d_b-1} - (1-\pi_d)^{d_l-1})(\frac{1}{d_b})^2(-p_b) + (1-\pi_d)^{d_l-1}(\frac{1}{d_b})^1(p_l - p_b)}{(1-(1-\pi_d)^{d_b-1})(\frac{1}{d_b})^1 + ((1-\pi_d)^{d_b-1} - (1-\pi_d)^{d_l-1})(\frac{1}{d_b})^2 + (1-\pi_d)^{d_l-1}(\frac{1}{d_b})^1} \\
&= \frac{(p_l - p_b) - (1-\pi_d)^{d_b-1}(p_l - p_b + p_b\frac{1}{d_b}) + (1-\pi_d)^{d_l-1}(p_l - p_b + p_b\frac{1}{d_b})}{1 - (1-\pi_d)^{d_b-1}(1-\frac{1}{d_b}) + (1-\pi_d)^{d_l-1}(1-\frac{1}{d_b})} \\
V(L_{b'} - L_l) &= \frac{(1-(1-\pi_d)^{d_b-1})(\frac{1}{d_{b'}})^2(-p_b) + ((1-\pi_d)^{d_b-1} - (1-\pi_d)^{d_l-1})(\frac{1}{d_{b'}})^2(-p_b) + (1-\pi_d)^{d_l-1}(\frac{1}{d_{b'}})^1(p_l - p_b)}{(1-(1-\pi_d)^{d_b-1})(\frac{1}{d_{b'}})^2 + ((1-\pi_d)^{d_{b'}-1} - (1-\pi_d)^{d_l-1})(\frac{1}{d_{b'}})^2 + (1-\pi_d)^{d_l-1}(\frac{1}{d_{b'}})^1} \\
&= \frac{-p_b\frac{1}{d_{b'}} + (1-\pi_d)^{d_l-1}(p_b\frac{1}{d_{b'}} + p_l - p_b)}{\frac{1}{d_{b'}} + (1-\pi_d)^{d_l-1}(1-\frac{1}{d_{b'}})}
\end{aligned}$$

$$V(L_b - L_l) < V(L_{b'} - L_l)$$

$$\frac{(p_l - p_b) - (1-\pi_d)^{d_b-1}(p_l - p_b + p_b\frac{1}{d_b}) + (1-\pi_d)^{d_l-1}(p_l - p_b + p_b\frac{1}{d_b})}{1 - (1-\pi_d)^{d_b-1}(1-\frac{1}{d_b}) + (1-\pi_d)^{d_l-1}(1-\frac{1}{d_b})} < \frac{-p_b\frac{1}{d_{b'}} + (1-\pi_d)^{d_l-1}(p_b\frac{1}{d_{b'}} + p_l - p_b)}{\frac{1}{d_{b'}} + (1-\pi_d)^{d_l-1}(1-\frac{1}{d_{b'}})}$$

$$
\begin{aligned}
& (p_l - p_b)\frac{1}{d_{b'}} + (p_l - p_b)(1-\pi_d)^{d_l-1}(1-\frac{1}{d_{b'}}) - (1-\pi_d)^{d_b-1}(p_l - p_b + p_b\frac{1}{d_b})\frac{1}{d_{b'}} \\
& - (1-\pi_d)^{d_l+d_b-2}(p_l - p_b + p_b\frac{1}{d_b})(1-\frac{1}{d_{b'}}) + (1-\pi_d)^{d_l-1}(p_l - p_b + p_b\frac{1}{d_b})\frac{1}{d_{b'}} + (1-\pi_d)^{2d_l-2}(p_l - p_b + p_b\frac{1}{d_b})(1-\frac{1}{d_{b'}}) \\
< \; & - p_b\frac{1}{d_{b'}} + (p_b\frac{1}{d_{b'}})(1-\pi_d)^{d_b-1}(1-\frac{1}{d_b}) - (p_b\frac{1}{d_{b'}})(1-\pi_d)^{d_l-1}(1-\frac{1}{d_b}) + (1-\pi_d)^{d_l-1}(p_b\frac{1}{d_{b'}} + p_l - p_b)) \\
& - (1-\pi_d)^{d_l+d_b-2}(p_b\frac{1}{d_{b'}} + p_l - p_b)(1-\frac{1}{d_b}) + (1-\pi_d)^{2d_l-2}(p_b\frac{1}{d_{b'}} + p_l - p_b)(1-\frac{1}{d_b})
\end{aligned}
$$

$$
\begin{aligned}
p_l\frac{1}{d_{b'}} - (1-\pi_d)^{d_b-1}\frac{p_l}{d_{b'}} + (1-\pi_d)^{d_l+d_b-2}p_l(\frac{1}{d_{b'}} - \frac{1}{d_b}) + (1-\pi_d)^{2d_l-2}p_l(\frac{1}{d_b} - \frac{1}{d_{b'}}) &< 0 \\
p_l\frac{1}{d_{b'}} - (1-\pi_d)^{d_b-1}\frac{p_l}{d_{b'}} + (1-\pi_d)^{d_l-2}p_l(\frac{1}{d_b} - \frac{1}{d_{b'}})((1-\pi_d)^{d_l} - (1-\pi_d)^{d_b}) &< 0 \\
\frac{1}{d_{b'}}(1-(1-\pi_d)^{d_b-1}) + (1-\pi_d)^{d_l-2}(\frac{1}{d_b} - \frac{1}{d_{b'}})((1-\pi_d)^{d_l} - (1-\pi_d)^{d_b}) &< 0
\end{aligned}
$$

$\frac{p_l}{p_b} < \frac{3}{2}$, follow-up subscription L_b

$$
\begin{aligned}
V(L_b - L_l) &= \frac{(1-(1-\pi_d)^{d_b-1} + (1-\pi_d)^{d_l-1})(\frac{1}{d_b})(p_l - p_b) + ((1-\pi_d)^{d_b-1} - (1-\pi_d)^{d_l-1})(-2p_b + p_l)}{(1-(1-\pi_d)^{d_b-1} + (1-\pi_d)^{d_l-1})(\frac{1}{d_b}) + ((1-\pi_d)^{d_b-1} - (1-\pi_d)^{d_l-1})} \\
&= \frac{(p_l - p_b)\frac{1}{d_b} - (1-\pi_d)^{d_b-1}((p_l - p_b)\frac{1}{d_b} + (2p_b - p_l)) + (1-\pi_d)^{d_l-1}((p_l - p_b)\frac{1}{d_b} + (2p_b - p_l))}{\frac{1}{d_b} + (1-\frac{1}{d_b})(1-\pi_d)^{d_b-1} - (1-\frac{1}{d_b})(1-\pi_d)^{d_l-1}} \\
V(L_{b'} - L_l) &= \frac{(1-(1-\pi_d)^{d_l-1})(-2p_b + p_l) + (1-\pi_d)^{d_l-1}(\frac{1}{d_{b'}})(p_l - p_b)}{(1-(1-\pi_d)^{d_l-1}) + (1-\pi_d)^{d_l-1}(\frac{1}{d_{b'}})} \\
&= \frac{-2p_b + p_l + (1-\pi_d)^{d_l-1}((p_l - p_b)\frac{1}{d_{b'}} + 2p_b - p_l)}{1-(1-\frac{1}{d_{b'}})(1-\pi_d)^{d_l-1}}
\end{aligned}
$$

$$
V(L_b - L_l) < V(L_{b'} - L_l)
$$

$$
\frac{\frac{p_l - p_b}{d_b} - (1-\pi_d)^{d_b-1}(\frac{p_l - p_b}{d_b} + 2p_b - p_l) + (1-\pi_d)^{d_l-1}(\frac{p_l - p_b}{d_b} + 2p_b - p_l)}{\frac{1}{d_b} + (1-\frac{1}{d_b})(1-\pi_d)^{d_b-1} - (1-\frac{1}{d_b})(1-\pi_d)^{d_l-1}} < \frac{p_l - 2_b + (1-\pi_d)^{d_l-1}(\frac{p_l - p_b}{d_{b'}} + 2p_b - p_l)}{1-(1-\frac{1}{d_{b'}})(1-\pi_d)^{d_l-1}}
$$

$$
\begin{aligned}
& (p_l - p_b)\frac{1}{d_b} - (1-\pi_d)^{d_b-1}((p_l - p_b)\frac{1}{d_b} + 2p_b - p_l) + (1-\pi_d)^{d_l-1}((p_l - p_b)\frac{1}{d_b} + 2p_b - p_l) \\
& - \frac{p_l - p_b}{d_b}(1-\frac{1}{d_{b'}})(1-\pi_d)^{d_l-1} + (\frac{p_l - p_b}{d_b} + 2p_b - p_l)(1-\frac{1}{d_{b'}})(1-\pi_d)^{d_l+d_b-2} - (\frac{p_l - p_b}{d_b} + 2p_b - p_l)(1-\frac{1}{d_{b'}})(1-\pi_d)^{2d_l-2} \\
< \; & (p_l - 2p_b)\frac{1}{d_b} - (2p_b - p_l)(1-\frac{1}{d_b})(1-\pi_d)^{d_b-1} + (2p_b - p_l)(1-\frac{1}{d_b})(1-\pi_d)^{d_l-1} + \frac{1}{d_b}(1-\pi_d)^{d_l-1}((p_l - p_b)\frac{1}{d_{b'}} + 2p_b - p_l) \\
& + (1-\pi_d)^{d_l+d_b-2}((p_l - p_b)\frac{1}{d_{b'}} + 2p_b - p_l)(1-\frac{1}{d_b}) - (1-\pi_d)^{2d_l-2}((p_l - p_b)\frac{1}{d_{b'}} + 2p_b - p_l)(1-\frac{1}{d_b})
\end{aligned}
$$

$$
\begin{aligned}
\frac{p_b}{d_b} - (1-\pi_d)^{d_b-1}\frac{p_b}{d_b} + p_b(\frac{1}{d_b} - \frac{1}{d_{b'}})(1-\pi_d)^{d_l+d_b-2} - p_b(\frac{1}{d_b} - \frac{1}{d_{b'}})(1-\pi_d)^{2d_l-2} &< 0 \\
1 - (1-\pi_d)^{d_b-1} + (1 - \frac{d_b}{d_{b'}})(1-\pi_d)^{d_l+d_b-2} - (1 - \frac{d_b}{d_{b'}})(1-\pi_d)^{2d_l-2} &< 0 \\
1 - (1-\pi_d)^{d_b-1} + (1 - \frac{d_b}{d_{b'}})(1-\pi_d)^{d_l-2}((1-\pi_d)^{d_b} - (1-\pi_d)^{d_l}) &< 0
\end{aligned}
$$

$\frac{p_l}{p_b} > \frac{3}{2}$, follow-up subscription L_b

$$
\begin{aligned}
V(L_b - L_l) =& \frac{(1-(1-\pi_d)^{d_b-1} + (1-\pi_d)^{d_l-1})(\frac{1}{d_b})^1(p_l - p_b) + ((1-\pi_d)^{d_b-1} - (1-\pi_d)^{d_l-1})(\frac{1}{d_b})^2(-2p_b + p_l)}{(1-(1-\pi_d)^{d_b-1})(\frac{1}{d_b})^1 + ((1-\pi_d)^{d_b-1} - (1-\pi_d)^{d_l-1})(\frac{1}{d_b})^2 + (1-\pi_d)^{d_l-1}(\frac{1}{d_b})^1} \\
=& \frac{(p_l - p_b) - (1-\pi_d)^{d_b-1}(p_l - p_b + (2p_b - p_l)\frac{1}{d_b}) + (1-\pi_d)^{d_l-1}(p_l - p_b + (2p_b - p_l)\frac{1}{d_b})}{1 - (1-\pi_d)^{d_b-1}(1 - \frac{1}{d_b}) + (1-\pi_d)^{d_l-1}(1 - \frac{1}{d_b})} \\
V(L_{b'} - L_l) =& \frac{(1-(1-\pi_d)^{d_l-1})(\frac{1}{d_{b'}})^2(-2p_b + p_l) + (1-\pi_d)^{d_l-1}(\frac{1}{d_{b'}})^1(p_l - p_b)}{(1-(1-\pi_d)^{d_b-1})(\frac{1}{d_{b'}})^2 + ((1-\pi_d)^{d_{b'}-1} - (1-\pi_d)^{d_l-1})(\frac{1}{d_{b'}})^2 + (1-\pi_d)^{d_l-1}(\frac{1}{d_{b'}})^1} \\
=& \frac{-(2p_b - p_l)\frac{1}{d_{b'}} + (1-\pi_d)^{d_l-1}((2p_b - p_l)\frac{1}{d_{b'}} + p_l - p_b)}{\frac{1}{d_{b'}} + (1-\pi_d)^{d_l-1}(1 - \frac{1}{d_{b'}})}
\end{aligned}
$$

$$
V(L_b - L_l) < V(L_{b'} - L_l)
$$

$$
\frac{(p_l - p_b) - (1-\pi_d)^{d_b-1}(p_l - p_b + \frac{2p_b - p_l}{d_b}) + (1-\pi_d)^{d_l-1}(p_l - p_b + \frac{2p_b - p_l}{d_b})}{1 - (1-\pi_d)^{d_b-1}(1 - \frac{1}{d_b}) + (1-\pi_d)^{d_l-1}(1 - \frac{1}{d_b})} < \frac{\frac{p_l - 2p_b}{d_{b'}} + (1-\pi_d)^{d_l-1}(\frac{2p_b - p_l}{d_{b'}} + p_l - p_b)}{\frac{1}{d_{b'}} + (1-\pi_d)^{d_l-1}(1 - \frac{1}{d_{b'}})}
$$

$$
\begin{aligned}
&(p_l - p_b)\frac{1}{d_{b'}} + (p_l - p_b)(1-\pi_d)^{d_l-1}(1 - \frac{1}{d_{b'}}) - (1-\pi_d)^{d_b-1}(p_l - p_b + (2p_b - p_l)\frac{1}{d_b})\frac{1}{d_{b'}} \\
&- (1-\pi_d)^{d_l+d_b-2}(p_l - p_b + \frac{2p_b - p_l}{d_b})(1 - \frac{1}{d_{b'}}) + (1-\pi_d)^{d_l-1}(p_l - p_b + \frac{2p_b - p_l}{d_b})\frac{1}{d_{b'}} + (1-\pi_d)^{2d_l-2}(p_l - p_b + \frac{2p_b - p_l}{d_b})(1 - \frac{1}{d_{b'}}) \\
< &- \frac{2p_b - p_l}{d_{b'}} + (\frac{2p_b - p_l}{d_{b'}})(1-\pi_d)^{d_b-1}(1 - \frac{1}{d_b}) - ((2p_b - p_l)\frac{1}{d_{b'}})(1-\pi_d)^{d_l-1}(1 - \frac{1}{d_b}) + (1-\pi_d)^{d_l-1}((2p_b - p_l)\frac{1}{d_{b'}} + p_l - p_b)) \\
&- (1-\pi_d)^{d_l+d_b-2}((2p_b - p_l)\frac{1}{d_{b'}} + p_l - p_b)(1 - \frac{1}{d_b}) + (1-\pi_d)^{2d_l-2}((2p_b - p_l)\frac{1}{d_{b'}} + p_l - p_b)(1 - \frac{1}{d_b})
\end{aligned}
$$

$$
\begin{aligned}
p_b\frac{1}{d_{b'}} - (1-\pi_d)^{d_b-1}\frac{p_b}{d_{b'}} + (1-\pi_d)^{d_l+d_b-2}p_b(\frac{1}{d_{b'}} - \frac{1}{d_b}) + (1-\pi_d)^{2d_l-2}p_b(\frac{1}{d_b} - \frac{1}{d_{b'}}) &< 0 \\
p_b\frac{1}{d_{b'}} - (1-\pi_d)^{d_b-1}\frac{p_b}{d_{b'}} + (1-\pi_d)^{d_l-2}p_b(\frac{1}{d_b} - \frac{1}{d_{b'}})((1-\pi_d)^{d_l} - (1-\pi_d)^{d_b}) &< 0 \\
\frac{1}{d_{b'}}(1 - (1-\pi_d)^{d_b-1}) + (1-\pi_d)^{d_l-2}(\frac{1}{d_b} - \frac{1}{d_{b'}})((1-\pi_d)^{d_l} - (1-\pi_d)^{d_b}) &< 0
\end{aligned}
$$

Appendix 4.C Power analysis

A between-subject design seems reasonable in the present experiment to minimize experimenter demand effects. We carried out a two-sample power analysis of the pilot data, following the conduct of (Moffatt, 2015, Table 4.5), to calculate the approximated sample size for the online sessions. Following the standard in economic literature, we set α=0.05 and β=0.2 for a test power of π=0.8. Although the effect of pigeonholing is statistically significant in subgroup 1 of problem 1, according to the power analysis, the sample size needs to double up to comply with the target test power. Subgroup 2 provides overwhelming evidence of pigeonholing. The power analysis states a sufficient sample size. The PR observed in subgroup 2 of problem 2 bears strong statistical significance; the sample size is close to the target value. The pigeonholing effect in subgroup 1 of problem 3 is almost marginally significant. The required sample size needs to triple up to satisfy a test power of 0.8.

Table 4.5: Power analysis (α=0.05, β=0.2) and two-sample hypothesis testing

	Problem 1 (Rockefeller)	Problem 2 (Ride-sharing)	Problem 3 (Delivery service)
p_b	5.00$	3.00€	4.00€; 6.00€
p_l	5.50$	7.95€; 4.99€	5.99€; 9.95€
d_b	2 w; 6 m	1 w	1 w
d_l	3 m; 18 m	10 w	8 w
n control group	25; 18	25; 18	25; 18
n treatment group	22; 23	22; 23	22; 23
t-test statistic	-1.90*; -2.77***	0.52; -2.32**	-1.54; 0.36
required n per group	44; 18	484; 25	64; 139

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Appendix 4.D Demographic analysis

Study 2 finds that participants holding higher academic degrees state higher re-use likelihood, particularly in the Rockefeller question, p=.025. Under the assumption that probabilities generally suffer from overestimation in these contexts, this finding opposes the idea that better-educated individuals think more rational and are less manipulable than people who enjoyed less education (Janssen et al., 2019). Study 3 also finds that estimates of better educated people are more biased by the treatment effect in the Rockefeller context (p<.01, Appendix Table 4.13). The stimulating PH effect from education appears to be present in other contexts in Study 3, too. Accordingly, a one-sided t-test for participants having received further education after high school finds mild evidence of a PH effect, $t(50)$=-1.4, p=.083 (Table 4.2). Yet on the aggregate level, Appendix Table 4.6 can determine neither a significant mitigating nor promoting influence on the PH effect by better-educated participants. Similarly, Appendix Table 4.7 finds no significant difference in re-use likelihood across treatments between higher and lower educated subjects.

A question of cognitive ability?

The literature considers the CRT to be a potent predictor for performance in heuristics-and-biases tasks (Toplak et al., 2011). In an attempt to proxy cognitive abilities and logical thinking, this paragraph relates subjects' CRT scores obtained in the experiment to their stated preferences.

Besides examining implications of the CRT score for correct answers, hereinafter referred to as *rational* CRT score, we also test for effects by the *naïve* CRT score. The naïve CRT score counts incorrect answers in the test, prompted by intuitive thinking but not suppressed and overwritten by a more sophisticated deliberation process (Frederick, 2005). Reasonably, the regression terms of the rational and the naïve CRT score show opposite signs, even in sessions combining multiple CRT questions. Only 4 of 19 constructs falsify the null hypothesis that the treatment does not affect high-CRT performers' re-use likelihood evaluation (Appendix Table 4.11 onward). Also, the interacted naïve-CRT/treatment term finds significant evidence of a treatment effect in the re-use likelihood of "naïve" participants. One could therefore assume, that choice menus containing single-use aids the rational assessment of the re-use likelihood, leading to lower estimates in low-probability scenarios. Out of 19 constructs examined, this hypothesis finds support in six, while two constructs exhibit a significant effect in the opposite direction. Pooled data from all online sessions reveals a slight positive bias between women and higher education (r=.15). There is further a negative relationship between women and the rational CRT score (r=-.1), yet not so for the naïve CRT score (r<.1). Generally performing worse in the CRT (r=-.14), older participants rise more often to the naïve bait (r=.12).

Appendix 4.E Empirical evidence

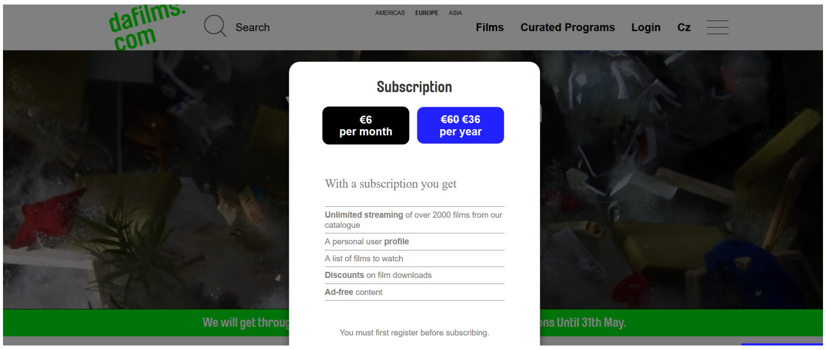


Figure 4.1: Platform for documentary and experimental films, Dafilms (2020)

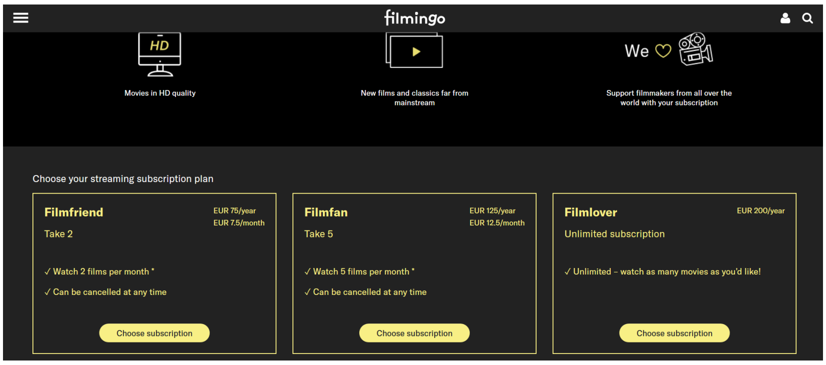


Figure 4.2: Online movie platform Filmingo (2020)

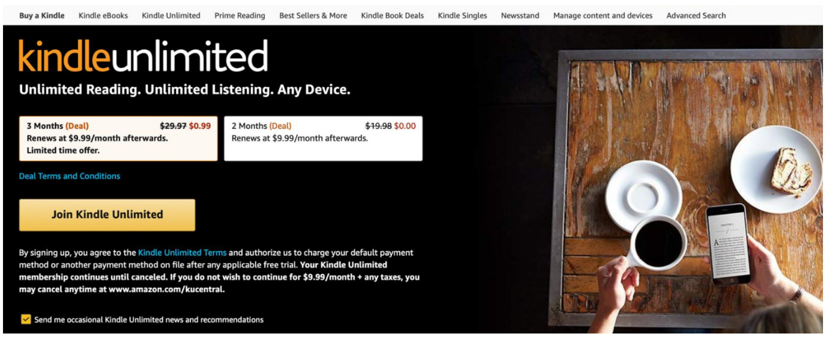


Figure 4.3: E-book provider Kindle (2020)

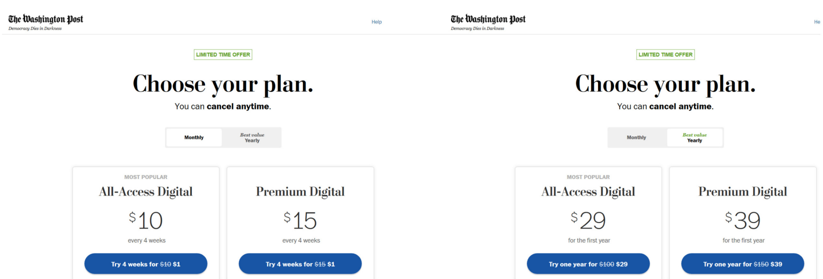


Figure 4.4: Newspaper WashingtonPost (2020)

Appendix 4.F Regression tables

All data

Table 4.6: All studies. Probit regressions, dep. variable: preference for short subscription

	(1) All	(2) All × S	(3) All-Pilot	(4) × S	(5) All+CRT	(6) × S	(7) All+Risk	(8) × S
Dependent variable	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b
Single treatment (S)	0.021 (0.057)	0.035 (0.244)	-0.081 (0.072)	-0.090 (0.391)	-0.167** (0.085)	-0.358 (0.468)	-0.070 (0.099)	-0.254 (0.678)
Age	0.014*** (0.005)	0.014** (0.007)	0.016** (0.007)	0.011 (0.009)	0.013* (0.007)	0.007 (0.010)	0.018** (0.008)	0.019** (0.009)
Female	0.074 (0.057)	0.111 (0.082)	0.237*** (0.074)	0.278*** (0.097)	0.250*** (0.091)	0.234** (0.111)	0.259** (0.111)	0.457*** (0.148)
Probability			-0.026*** (0.002)	-0.023*** (0.002)	-0.029*** (0.002)	-0.028*** (0.002)	-0.030*** (0.003)	-0.031*** (0.003)
Education			-0.021 (0.031)	-0.015 (0.043)	-0.051 (0.038)	-0.030 (0.049)	-0.020 (0.042)	-0.064 (0.053)
CRT-naïve					-0.057 (0.088)	0.021 (0.105)	0.097 (0.110)	0.139 (0.139)
CRT					0.016 (0.083)	-0.044 (0.105)	-0.182* (0.101)	-0.204 (0.141)
Risk							-0.035 (0.033)	-0.018 (0.038)
Single × Age		0.001 (0.009)		0.010 (0.013)		0.013 (0.014)		0.002 (0.016)
Single × Female		-0.076 (0.114)		-0.066 (0.149)		0.035 (0.183)		-0.367* (0.218)
Single × Probability				-0.006* (0.003)		-0.003 (0.004)		0.001 (0.006)
Single × Education				-0.017 (0.063)		-0.039 (0.075)		0.074 (0.082)
Single × CRT-naïve						-0.178 (0.175)		-0.109 (0.214)
Single × CRT						0.136 (0.167)		0.060 (0.203)
Single × Risk								-0.027 (0.069)
Constant	-0.570*** (0.127)	-0.578*** (0.181)	0.211 (0.195)	0.207 (0.216)	0.583** (0.230)	0.654*** (0.234)	0.496 (0.330)	0.560 (0.385)
Observations	2308	2308	2033	2033	1538	1538	920	920
Log likelihood	-1573.255	-1572.998	-1063.427	-1059.126	-773.790	-772.106	-454.998	-452.999
Pseudo R-squared	0.005	0.005	0.235	0.238	0.272	0.273	0.286	0.289

Note: Standard errors in parentheses. Clustered on subject level. * $p < .1$, ** $p < .05$, *** $p < .01$.

Table 4.7: All studies. Tobit regressions, dep. variable: stated re-use likelihood

	(1) All-Pilot	(2) × S	(3) All+CRT	(4) × S	(5) All+Risk	(6) × S
Dependent variable	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}
Single	0.247	4.408	-0.177	1.332	-5.657*	-11.894
	(2.039)	(9.607)	(2.192)	(10.614)	(3.357)	(19.021)
Female	2.858	1.635	0.603	0.115	3.952	3.463
	(2.086)	(2.976)	(2.268)	(3.326)	(3.481)	(5.229)
Age	-0.111	-0.028	-0.144	-0.013	-0.012	-0.032
	(0.166)	(0.238)	(0.165)	(0.233)	(0.254)	(0.360)
Education	-0.381	-0.303	0.193	-0.288	-0.010	-1.040
	(0.791)	(1.099)	(0.874)	(1.257)	(1.392)	(1.986)
CRT-naïve			5.997***	6.895**	6.260*	7.499
			(2.259)	(3.196)	(3.578)	(5.217)
CRT			-7.191***	-7.680**	-10.256***	-11.258**
			(2.200)	(3.169)	(3.400)	(4.990)
Risk					0.946	1.618
					(1.132)	(1.487)
Single × Age		-0.169		-0.257		0.073
		(0.331)		(0.330)		(0.508)
Single × Female		2.264		1.014		1.355
		(4.170)		(4.546)		(7.031)
Single × Education		-0.151		0.928		2.158
		(1.577)		(1.742)		(2.812)
Single × CRT-naïve				-1.853		-1.739
				(4.534)		(7.096)
Single × CRT				1.071		1.568
				(4.396)		(6.794)
Single × Risk						-1.610
						(2.303)
Constant	29.254***	27.169***	28.873***	28.134***	27.953***	31.062**
	(4.939)	(6.658)	(5.437)	(7.285)	(9.561)	(12.751)
Observations	3020	3020	2426	2426	920	920
Log likelihood	-12571.704	-12571.154	-9796.071	-9795.313	-3807.107	-3805.694
Pseudo R-squared	0.000	0.000	0.001	0.001	0.003	0.003

Note: Censored between 0 and 100. Standard errors in parentheses. Clustered on subject level. Corrected for comprehension control. * $p < .1$, ** $p < .05$, *** $p < .01$.

Table 4.8: Mean comparison of likelihood estimates in % (treatment/control)

	Total	Study 2	Study 3	Study 4a	Study 4b	Study 5a	Study 5b
All	29.7/29.8	29.8/38.0**	37.5/41.9**	21.4/21.9	31.9/30.5	30.7/23.6***	33.3/32.6
Music-streaming	25.4/25.3	-	-	7.0/7.8	30.3/28.6	29.2/29.8	37.1/37.1
Accom-sharing	29.2/30.0	-	35.6/33.6	27.7/27.0	20.8/30.1*	-	-
Ride-sharing	36.5/36.6	38.8/46.7**	42.1/43.4	19.4/20.4	37.6/30.7	40.7/32.7*	45.9/48.7
Skate rental	11.8/14.4	-	-	4.1/8.9**	28.7/24.8	-	-
Express delivery	44.6/43.8	35.7/49.1*	48.8/50.6	45.6/37.8*	44.3/38.6	-	-
Rockefeller	13.4/13.5	21.8/18.1	21.0/23.9	3.9/6.1	9.5/16.5	14.1/7.5**	13.3/13.7
Flood	34.2/30.4	-	38.9/51.1**	20.8/18.9	39.1/29.7	38.9/24.4***	36.8/30.5
Covid	41.4/47.8**	-	38.3/48.9**	42.4/48.5	44.9/44.8	-	-

Note: Clustered on subject level, one-sided: * $p < .1$, ** $p < .05$, *** $p < .01$.

Table 4.9: Median comparison of likelihood estimates in % (treatment/control)

	Total	Study 2	Study 3	Study 4a	Study 4b	Study 5a	Study 5b
All	15/15	20/25	30/40	5/10	20/20	15/10*	20/20
Music-streaming	5/5	-	-	0/1*	10/10	10/10	27.5/25
Accom-sharing	20/20	-	30/30	20/15	12.5/20	-	-
Ride-sharing	25/20	22.5/50	45/45	10/10	40/20*	35/20	45/50
Skate rental	1/2.5	-	-	0/1***	20/5	-	-
Express delivery	50/40	25/50	50/50	50/30*	35/33	-	-
Rockefeller	1/1	9/10	5/6.5	0/0	2/2	5/0**	5/1
Flood	20/20	-	25/57.5	10/10	40/25	40/10*	30/20
Covid	35/50*	-	30/50*	30/50	50/50	-	-

Note: Clustered on subject level and corrected for continuity; one-sided: * $p < .1$, ** $p < .05$, *** $p < .01$.

Pilot

Table 4.10: Pilot study. Probit regressions, dep. variable: preference for short subscription

	Pilot 1								Pilot 2							
	(1) All	(2) × S	(3) Ride	(4) × S	(5) Rock	(6) × S	(7) Exp	(8) × S	(9) All	(10) × S	(11) Ride	(12) × S	(13) Rock	(14) × S	(15) Exp	(16) × S
Dep. variable	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b
Single (S)	0.384*	-17.244***	-0.099	0.317	0.735*	-25.812**	0.654	2.240	0.609***	1.547	1.694***	9.179	1.418***	-3.360	-1.185**	-1.868
	(0.226)	(5.098)	(0.387)	(0.651)	(0.399)	(12.727)	(0.404)	(11.348)	(0.212)	(3.703)	(0.574)	(10.163)	(0.543)	(9.372)	(0.603)	(9.358)
Age	-0.097	-0.735***	0.099		-0.029	-0.894*	-0.409*	-0.341	0.147**	0.167	0.342**	0.543	0.238	0.102	-0.039	-0.107
	(0.102)	(0.234)	(0.219)		(0.234)	(0.497)	(0.238)	(0.459)	(0.069)	(0.147)	(0.158)	(0.361)	(0.168)	(0.282)	(0.143)	(0.396)
Female	0.194	0.555	0.725*	1.105*	0.019	0.486	-0.188	-0.367	0.136	-0.482	1.353**	0.925	0.534	-0.324	-1.435**	-1.233*
	(0.227)	(0.399)	(0.396)	(0.629)	(0.417)	(0.672)	(0.407)	(0.657)	(0.191)	(0.619)	(0.671)	(1.921)	(0.610)	(1.453)	(0.679)	(0.722)
S × Age		0.843***				1.268**		-0.083		-0.072		-0.360		0.179		0.040
		(0.245)				(0.604)		(0.538)		(0.150)		(0.403)		(0.392)		(0.432)
S × Female		-0.440		-0.625		-0.603		0.284		0.800		0.557		1.010		0.000
		(0.462)		(0.811)		(0.898)		(0.840)		(0.652)		(2.046)		(1.597)		(.)
Constant	1.836	15.127***	-2.278	-0.431	0.712	18.778*	8.039	6.727	-3.883**	-3.757	-9.787**	-13.836	-6.386	-2.595	2.143	3.378
	(2.127)	(4.861)	(4.670)	(0.529)	(4.974)	(10.450)	(5.035)	(9.627)	(1.550)	(3.620)	(3.895)	(9.311)	(3.990)	(7.256)	(3.444)	(8.625)
Observations	141	141	47	47	47	47	47	47	123	123	41	41	41	41	41	39
Log likelihood	-95.457	-91.164	-29.715	-29.520	-27.660	-25.050	-27.207	-27.143	-77.955	-76.842	-19.859	-18.664	-23.458	-23.260	-24.540	-24.172
Pseudo R^2	0.019	0.063	0.063	0.069	0.060	0.149	0.097	0.099	0.052	0.066	0.225	0.271	0.166	0.173	0.118	0.070

Note: Standard errors in parentheses. Clustered on subject level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Study 2

Table 4.11: Study 2. Tobit regressions, dep. variable: stated re-use likelihood

	(1) All	(2) × S	(3) Rockefeller	(4) × S	(5) Express	(6) × S	(7) Ride-sharing	(8) × S
Dependent variable	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}
Single	2.598	-69.933*	-16.808*	-59.703	12.406	-57.827	9.268	-74.961
	(6.990)	(37.278)	(9.933)	(53.382)	(10.403)	(55.620)	(9.063)	(48.323)
Age	-0.945**	-1.164***	-0.764	-0.628	-1.109**	-1.407*	-0.966**	-1.369**
	(0.397)	(0.377)	(0.519)	(0.663)	(0.532)	(0.717)	(0.472)	(0.631)
Female	-9.740	-19.884**	-19.959**	-26.523*	3.677	-7.382	-14.186	-23.790*
	(6.614)	(9.489)	(9.698)	(13.855)	(10.145)	(14.945)	(8.823)	(12.899)
Education	-2.496	-7.907***	-1.934	-9.687**	-5.047	-8.287*	-0.529	-5.311
	(2.250)	(2.489)	(3.253)	(4.785)	(3.441)	(4.927)	(2.998)	(4.236)
CRT-naïve	4.469	9.148*	3.887	5.153	0.235	5.657	9.483**	15.838**
	(3.223)	(5.148)	(4.269)	(6.384)	(4.568)	(7.168)	(4.045)	(6.306)
CRT	-5.006	-10.701**	-4.032	-4.490	-0.433	-8.482	-10.747**	-17.742***
	(3.283)	(5.321)	(4.338)	(6.633)	(4.611)	(7.365)	(4.025)	(6.454)
Time	-0.009	-0.019	-0.006	-0.025	-0.013	-0.017	-0.008	-0.013
	(0.009)	(0.015)	(0.012)	(0.017)	(0.013)	(0.019)	(0.011)	(0.016)
Single × Age		0.127		-1.866		0.562		0.759
		(0.780)		(1.292)		(1.084)		(0.948)
Single × Female		12.124		3.194		15.443		11.346
		(12.599)		(19.435)		(20.241)		(17.364)
Single × Education		10.042**		17.601**		5.703		9.350
		(4.101)		(7.574)		(7.145)		(6.145)
Single × CRT-naïve		-6.930		-0.403		-8.862		-10.390
		(6.492)		(8.403)		(9.280)		(8.074)
Single × CRT		8.307		-1.449		13.020		11.249
		(6.645)		(8.583)		(9.397)		(8.148)
Single × Time		0.004		0.032		-0.009		-0.012
		(0.017)		(0.025)		(0.028)		(0.024)
Constant	86.484***	140.615***	67.752***	114.929**	95.985***	144.848***	97.851***	153.891***
	(19.321)	(28.503)	(24.596)	(46.103)	(26.654)	(47.639)	(23.327)	(41.649)
Observations	162	162	54	54	54	54	54	54
Log likelihood	-683.041	-679.562	-179.468	-174.758	-242.032	-240.533	-238.501	-236.021
Pseudo R-squared	0.013	0.018	0.024	0.049	0.020	0.026	0.033	0.043

Note: Censored between 0 and 100. Standard errors in parentheses. Clustered on subject level. Corrected for comprehension control. * $p <$ .1, ** $p < .05$, *** $p < .01$.

Study 3

Table 4.12: Study 3: Probit regressions, dep. variable: preference for short subscription

	(1) All	(2) × S	(3) Accom	(4) × S	(5) Ride	(6) × S	(7) Rockef.	(8) × S	(9) Flood	(10) × S	(11) Covid	(12) × S
Dep. variable	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b
Single (S)	0.219	-2.002	-0.048	0.832	0.271	0.807	0.427	-1.753	0.764*	-3.351	0.172	-9.527**
	(0.140)	(1.468)	(0.290)	(2.834)	(0.401)	(4.534)	(0.321)	(3.218)	(0.405)	(4.546)	(0.337)	(4.510)
Probability	-0.020***	-0.013***										
	(0.003)	(0.004)										
Age	0.042	-0.007	0.078	0.087	0.117	0.132	0.056	0.022	0.039	-0.063	-0.058	-0.230*
	(0.026)	(0.041)	(0.056)	(0.082)	(0.087)	(0.140)	(0.063)	(0.091)	(0.071)	(0.164)	(0.070)	(0.120)
Female	0.275**	0.452**	-0.172	-0.210	0.667	0.620	0.708**	0.944*	0.431		0.016	0.654
	(0.137)	(0.211)	(0.291)	(0.433)	(0.406)	(0.639)	(0.331)	(0.513)	(0.407)		(0.339)	(0.520)
Education	0.038	-0.035	0.194*	0.178	-0.017	0.181	-0.180	-0.314*	-0.210	-0.290	0.231*	-0.149
	(0.050)	(0.082)	(0.106)	(0.160)	(0.147)	(0.256)	(0.120)	(0.190)	(0.146)	(0.353)	(0.124)	(0.191)
Time	0.000	0.000	-0.001	-0.000	-0.000	-0.001	0.001*	0.001*	0.000	-0.000	0.000	0.001
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.000)	(0.001)
Prob Accom			-0.014***	-0.010								
			(0.005)	(0.007)								
Prob Ride					-0.037***	-0.043***						
					(0.009)	(0.016)						
Prob Rock							-0.026***	-0.020**				
							(0.006)	(0.009)				
Prob Flood									-0.007	0.017		
									(0.006)	(0.013)		
Prob Covid											-0.018***	-0.006
											(0.006)	(0.007)
S × Prob		-0.014**										
		(0.006)										
S × Age		0.099*		-0.009		0.002		0.061		0.190		0.329*
		(0.056)		(0.114)		(0.183)		(0.127)		(0.188)		(0.176)
S × Female		-0.272		0.166		0.034		-0.398				-1.208
		(0.269)		(0.597)		(0.853)		(0.689)				(0.844)
S × Edu		0.102		-0.016		-0.284		0.217		0.156		0.806***
		(0.105)		(0.223)		(0.329)		(0.272)		(0.393)		(0.307)
S × Time		-0.000		-0.001		0.001		0.000		0.001		-0.001
		(0.000)		(0.001)		(0.001)		(0.001)		(0.002)		(0.001)
S×Prob Accom				-0.008								
				(0.010)								
S×Prob Ride						0.007						
						(0.020)						
S×Prob Rock								-0.009				
								(0.012)				
S×Prob Flood										-0.035**		
										(0.015)		
S×Prob Covid												-0.030*
												(0.015)
Constant	-1.292*	-0.136	-1.852	-2.253	-2.968	-3.747	-1.070	0.024	-1.743	0.176	0.110	4.876*
	(0.682)	(0.980)	(1.379)	(2.011)	(2.111)	(3.464)	(1.533)	(2.188)	(1.820)	(3.881)	(1.652)	(2.737)
Obs.	470	470	94	94	94	94	94	94	94	94	94	94
Log l.hood	-245.014	-238.940	-58.110	-57.542	-29.020	-28.426	-45.541	-44.621	-29.051	-24.823	-41.555	-33.537
Pseudo R^2	0.159	0.180	0.108	0.117	0.347	0.360	0.297	0.312	0.144	0.268	0.168	0.328

Note: Censored between 0 and 100. Standard errors in parentheses. Clustered on subject level. Corrected for comprehension control. * $p < .1$, ** $p < .05$, *** $p < .01$.

Table 4.13: Study 3. Tobit regressions, dep. variable: stated re-use likelihood

	(1) All	(2) × S	(3) Accom	(4) × S	(5) Ride	(6) × S	(7) Express	(8) × S	(9) Rockef.	(10) × S	(11) Flood	(12) × S	(13) Covid	(14) × S	(15)	(16)
Dep. var.	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}		
Single (S)	2.351		-2.002		1.538		-2.225		2.619		5.233		9.157			
	(4.445)		(6.355)		(6.777)		(7.079)		(8.673)		(7.552)		(7.379)			
Age	1.581**	1.227	1.498	0.281	2.362*	0.574	-0.035	0.359	1.745	2.210	2.000	1.235	1.890	2.784		
	(0.747)	(1.006)	(1.208)	(1.767)	(1.293)	(1.901)	(1.346)	(1.973)	(1.654)	(2.331)	(1.435)	(2.126)	(1.402)	(2.041)		
Female	1.194		-3.637		-7.672		11.446		-12.291		15.746**		1.530			
	(4.333)		(6.419)		(6.834)		(7.124)		(8.793)		(7.622)		(7.442)			
Education	-0.121	2.512	-1.373	0.344	2.138	4.779	-0.442	-3.592	-2.928	6.543	1.361	4.695	0.309	2.784		
	(1.573)	(2.098)	(2.323)	(3.486)	(2.479)	(3.743)	(2.596)	(3.921)	(3.169)	(4.580)	(2.754)	(4.190)	(2.698)	(4.028)		
Time	-0.014**	-0.016**	-0.006	-0.010	-0.022**	-0.026**	-0.008	-0.001	-0.015	-0.030*	-0.016	-0.021	-0.017	-0.010		
	(0.006)	(0.008)	(0.009)	(0.012)	(0.009)	(0.013)	(0.010)	(0.013)	(0.012)	(0.015)	(0.011)	(0.014)	(0.010)	(0.014)		
Single		6.203		-64.635		-74.845		23.347		74.117		-15.540		100.973		
		(37.784)		(60.109)		(64.551)		(66.920)		(79.362)		(72.232)		(69.379)		
Female		2.732		-9.930		-8.530		23.447**		-8.968		15.084		3.969		
		(5.229)		(9.470)		(10.115)		(10.535)		(12.424)		(11.415)		(10.894)		
S × Age		0.732		2.707		3.679		-1.365		-0.338		1.700		-2.059		
		(1.451)		(2.414)		(2.587)		(2.685)		(3.181)		(2.897)		(2.783)		
S × Fem		-2.427		12.376		2.482		-22.065		-5.969		1.595		-3.064		
		(8.468)		(12.843)		(13.674)		(14.235)		(16.815)		(15.400)		(14.769)		
S × Edu		-4.861		-3.129		-4.747		5.478		-16.525***		-5.937		-5.621		
		(3.166)		(4.709)		(5.029)		(5.260)		(6.201)		(5.640)		(5.428)		
S × Time		0.001		0.008		0.011		-0.011		0.019		0.008		-0.027		
		(0.014)		(0.019)		(0.020)		(0.021)		(0.024)		(0.022)		(0.022)		
Constant	7.386	4.795	8.537	34.772	-7.461	27.104	52.284	48.530	-1.742	-47.597	-10.309	-3.537	3.350	-35.017		
	(19.641)	(25.698)	(30.050)	(43.336)	(32.256)	(46.839)	(33.553)	(48.494)	(41.179)	(57.598)	(35.773)	(52.205)	(34.949)	(50.156)		
Obs.	564	564	94	94	94	94	94	94	94	94	94	94	94	94		
Log l.hood	-2577.403	-2574.947	-427.321	-425.802	-433.986	-432.428	-434.086	-432.224	-362.024	-358.087	-443.622	-442.900	-434.333	-432.311		
Pseudo R^2	0.004	0.004	0.003	0.006	0.010	0.014	0.004	0.008	0.008	0.019	0.012	0.014	0.007	0.012		

Note: Censored between 0 and 100. Standard errors in parentheses. Clustered on subject level. Corrected for comprehension control. * $p < .1$, ** $p < .05$, *** $p < .01$.

Table 4.14: Study 3. Linear regressions, dep. variable: WTP express delivery

	(1) Express delivery	(2) Express delivery × S
Dependent variable	*WTP*	*WTP*
Single	1.290***	0.128
	(0.424)	(4.113)
Probability	-0.002	0.007
	(0.007)	(0.010)
Age	-0.049	-0.069
	(0.081)	(0.119)
Female	-0.794*	-1.058
	(0.432)	(0.667)
Education	0.128	-0.240
	(0.155)	(0.236)
Time	0.001	0.002**
	(0.001)	(0.001)
Single × Probability		-0.019
		(0.013)
Single × Age		-0.000
		(0.162)
Single × Female		0.430
		(0.881)
Single × Education		0.644**
		(0.317)
Single × Time		-0.001
		(0.001)
Constant	3.046	4.323
	(2.040)	(2.951)
Observations	94	94
Log likelihood	-191.857	-187.709

Note: Standard errors in parentheses. Corrected for comprehension control. * $p < .1$, ** $p < .05$, *** $p < .01$.

Table 4.15: Study 4a. Linear regressions, dep. variable: WTP

	(1) Music	(2) × S	(3) Accom	(4) × S	(5) Ride	(6) × S	(7) Skate	(8) × S	(9) Exp	(10) × S	(11) Rockef.	(12) × S	(13) Flood	(14) × S	(15) Covid	(16) × S
Single	0.299	-0.440	0.784**	-0.673	0.363	2.320*	1.181***	1.161	1.401***	-0.119	8.074***	-7.540	201.558**	752.184	-8.270	811.819
	(0.195)	(1.048)	(0.300)	(1.579)	(0.245)	(1.297)	(0.360)	(1.877)	(0.451)	(2.529)	(2.666)	(13.744)	(87.984)	(465.274)	(406.068)	(2233.968)
Age	0.049***	0.029	0.109***	0.051	0.044*	0.084***	0.064**	0.095**	0.087**	0.107*	0.337	0.668**	16.670**	34.714***	13.687	18.181
	(0.018)	(0.026)	(0.027)	(0.038)	(0.022)	(0.031)	(0.032)	(0.046)	(0.041)	(0.058)	(0.240)	(0.336)	(7.932)	(11.195)	(36.645)	(53.397)
Female	0.017	-0.122	0.397	0.125	0.661**	0.337	0.814**	0.586	0.934*	0.686	5.304*	0.005	-22.747	-45.657	155.441	650.638
	(0.206)	(0.310)	(0.316)	(0.456)	(0.260)	(0.386)	(0.379)	(0.556)	(0.476)	(0.701)	(2.806)	(4.072)	(92.723)	(134.656)	(428.970)	(643.453)
Education	-0.072	-0.055	-0.182	-0.050	-0.207**	-0.259*	-0.097	-0.218	-0.039	-0.312	0.719	-1.372	-13.551	-5.372	7.153	5.902
	(0.078)	(0.113)	(0.119)	(0.168)	(0.097)	(0.139)	(0.139)	(0.202)	(0.178)	(0.255)	(1.058)	(1.487)	(35.013)	(49.759)	(160.522)	(236.394)
CRT-naïve	0.057	0.106	0.079	0.004	0.018	0.066	-0.034	-0.176	0.168	-0.160	-0.359	-2.086	-53.320	-32.265	35.022	3.864
	(0.091)	(0.127)	(0.138)	(0.182)	(0.113)	(0.149)	(0.162)	(0.218)	(0.207)	(0.280)	(1.235)	(1.606)	(42.058)	(54.690)	(187.929)	(256.339)
CRT	-0.107	-0.108	-0.055	0.026	-0.006	-0.016	0.066	0.148	0.009	0.220	1.299	1.915	77.781*	49.166	-12.234	22.252
	(0.085)	(0.122)	(0.130)	(0.177)	(0.107)	(0.144)	(0.153)	(0.212)	(0.196)	(0.272)	(1.166)	(1.564)	(39.526)	(53.027)	(176.819)	(248.974)
Prob. Music	-0.009*	-0.002														
	(0.005)	(0.007)														
Prob. Accom			-0.002	0.005												
			(0.005)	(0.007)												
Prob. Ride-sharing					0.004	0.012*										
					(0.005)	(0.007)										
Prob. Skate							0.010	0.045**								
							(0.011)	(0.021)								
Prob. Express									-0.007	0.003						
									(0.007)	(0.010)						
Prob. Rockefeller											0.042	0.210				
											(0.088)	(0.143)				
Prob. Flood													3.636**	4.667**		
													(1.808)	(2.203)		
Prob. Covid															6.430	1.358
															(6.431)	(8.432)
Single × Age		0.043		0.107**		-0.079*		-0.042		-0.045		-0.570		-34.372**		-14.821
		(0.037)		(0.054)		(0.044)		(0.064)		(0.082)		(0.470)		(15.655)		(74.782)
Single × Female		0.265		0.551		0.672		0.563		0.863		11.274**		-2.112		-1034.574
		(0.430)		(0.636)		(0.528)		(0.777)		(0.970)		(5.657)		(192.190)		(895.231)
Single × Education		-0.047		-0.235		0.104		0.244		0.624*		4.661**		-10.774		-84.424
		(0.164)		(0.241)		(0.196)		(0.286)		(0.364)		(2.110)		(70.586)		(336.243)
Single × CRT-naïve		-0.089		0.265		-0.044		0.223		0.532		3.239		-11.943		25.394
		(0.195)		(0.284)		(0.238)		(0.339)		(0.431)		(2.527)		(89.216)		(416.618)
Single × CRT		-0.037		-0.249		-0.071		-0.152		-0.301		-0.959		25.049		-23.018
		(0.181)		(0.264)		(0.222)		(0.315)		(0.402)		(2.389)		(82.139)		(379.040)
Single × Prob. Music		-0.013														
		(0.011)														
Single × Prob. Accom				-0.014												
				(0.010)												
Single × Prob. Ride						-0.021**										
						(0.010)										
Single × Prob. Skate								-0.047*								
								(0.025)								
Single × Prob. Exp										-0.022						
										(0.014)						
Single × Prob. Rockef.												-0.221				
												(0.182)				
Single × Prob. Flood														-2.683		
														(3.897)		
Single × Prob. Covid																13.147
																(13.771)
Constant	0.114	0.443	-0.398	0.157	0.849	-0.111	-0.076	-0.468	0.678	0.876	-4.570	-1.756	-385.838*	-664.482**	314.598	210.203
	(0.492)	(0.642)	(0.770)	(0.964)	(0.637)	(0.801)	(0.892)	(1.173)	(1.254)	(1.630)	(6.736)	(8.440)	(223.443)	(282.302)	(1047.351)	(1378.723)
Observations	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111

Note: Standard errors in parentheses. Corrected for comprehension control. * $p < .1$, ** $p < .05$, *** $p < .01$.

Table 4.16: Study 4a. Tobit regressions, dep. variable: stated re-use likelihood

	(1) Music	(2) × S	(3) Accom	(4) × S	(5) Ride	(6) × S	(7) Skate	(8) × S	(9) Exp	(10) × S	(11) Rock.	(12) × S	(13) Flood	(14) × S	(15) Covid	(16) × S
Dep. var.	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}
Single	8.593	45.311	2.187	27.004	6.746	9.163	18.992***	38.226	-6.646	-22.918	4.416	63.200	0.059	22.629	3.052	-0.105
	(7.910)	(39.527)	(6.564)	(34.224)	(6.185)	(31.136)	(6.103)	(31.151)	(6.848)	(35.664)	(8.666)	(45.852)	(5.476)	(28.183)	(6.294)	(32.443)
Age	1.070	0.865	-0.311	0.540	-0.862	-0.965	-0.325	-0.508	-1.083*	-0.545	-0.041	0.161	0.181	0.044	0.392	0.269
	(0.660)	(0.921)	(0.591)	(0.829)	(0.551)	(0.756)	(0.506)	(0.723)	(0.623)	(0.878)	(0.728)	(0.953)	(0.487)	(0.683)	(0.566)	(0.797)
Female	-3.580	-4.291	1.520	-1.035	-10.227	-24.583**	-17.494***	-23.690**	-9.575	-14.950	-4.388	-0.802	-4.320	7.026	7.403	9.885
	(8.386)	(12.772)	(6.887)	(10.094)	(6.517)	(9.403)	(6.528)	(11.492)	(7.208)	(10.488)	(9.153)	(13.033)	(5.781)	(8.319)	(6.620)	(9.564)
Education	-2.464	-0.713	3.443	1.559	3.018	4.770	2.022	2.738	1.092	0.042	3.794	4.446	3.876*	4.210	1.448	-1.085
	(3.071)	(4.362)	(2.622)	(3.743)	(2.415)	(3.335)	(2.337)	(3.505)	(2.733)	(3.869)	(3.416)	(4.840)	(2.199)	(3.149)	(2.497)	(3.532)
CRT-naïve	-4.886	-8.591*	-0.351	-1.261	2.690	-1.077	-4.075	-1.966	2.853	4.793	1.839	0.035	6.172**	3.815	4.162	-1.697
	(3.667)	(4.934)	(3.009)	(4.016)	(2.839)	(3.705)	(2.835)	(3.941)	(3.173)	(4.207)	(3.959)	(5.269)	(2.564)	(3.352)	(2.888)	(3.778)
CRT	2.763	7.682	-0.378	0.293	-2.663	0.924	1.820	3.554	-2.059	-3.351	-3.863	1.447	-6.107**	-4.309	-3.260	1.454
	(3.389)	(4.756)	(2.838)	(3.918)	(2.657)	(3.563)	(2.604)	(3.881)	(2.983)	(4.102)	(3.672)	(5.016)	(2.396)	(3.283)	(2.734)	(3.727)
Time	-0.021*	-0.019	-0.007	-0.007	-0.003	-0.005	-0.013	-0.005	-0.008	-0.014	-0.016	-0.016	0.003	0.001	-0.003	-0.001
	(0.011)	(0.014)	(0.009)	(0.011)	(0.008)	(0.010)	(0.009)	(0.012)	(0.009)	(0.012)	(0.012)	(0.015)	(0.007)	(0.009)	(0.008)	(0.010)
S × Age		0.311		-1.698		0.255		0.327		-1.021		-0.622		0.214		0.222
		(1.281)		(1.176)		(1.052)		(1.009)		(1.237)		(1.443)		(0.955)		(1.113)
S × Fem		-3.108		4.112		24.281*		7.276		12.482		-11.622		-23.687**		-4.710
		(16.796)		(14.016)		(12.807)		(14.088)		(14.612)		(17.986)		(11.588)		(13.337)
S × Edu		-4.902		3.775		-3.634		-1.770		3.535		-3.597		-2.471		3.645
		(6.080)		(5.339)		(4.709)		(4.728)		(5.569)		(6.692)		(4.406)		(5.039)
S × Time		-0.003		-0.000		0.014		-0.016		0.017		-0.004		-0.000		-0.003
		(0.022)		(0.019)		(0.017)		(0.017)		(0.019)		(0.024)		(0.015)		(0.018)
S × CRT-n.		9.668		1.237		9.527*		-4.046		-5.436		5.953		7.173		13.060**
		(7.360)		(6.243)		(5.666)		(5.676)		(6.554)		(8.069)		(5.236)		(5.866)
S × CRT		-11.033		-1.024		-8.874*		-3.413		2.853		-12.342*		-4.601		-9.284*
		(6.739)		(5.799)		(5.219)		(5.267)		(6.085)		(7.321)		(4.843)		(5.541)
Constant	-14.351	-27.229	22.927	11.045	31.723**	25.902	-1.969	-10.418	80.131***	81.258***	-17.741	-34.861	3.830	0.811	32.585**	36.808*
	(19.923)	(25.243)	(16.639)	(21.264)	(15.718)	(19.325)	(15.392)	(20.515)	(17.312)	(22.081)	(22.604)	(27.966)	(13.983)	(17.649)	(15.869)	(20.025)
Obs.	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111	111
Log l.hood	-258.368	-256.354	-469.745	-468.555	-436.940	-431.929	-259.079	-257.865	-503.495	-502.287	-187.913	-185.850	-447.129	-444.711	-518.580	-516.799
Pseudo R^2	0.014	0.022	0.003	0.006	0.008	0.019	0.036	0.040	0.009	0.011	0.015	0.026	0.010	0.015	0.006	0.009

Note: Censored between 0 and 100. Standard errors in parentheses. Corrected for comprehension control. * $p < .1$, ** $p < .05$, *** $p < .01$.

Table 4.17: Study 4b. Probit regressions, dep. variable: preference for short subscription

	(1) All	(2) All	(3) Music	(4) × S	(5) Accom	(6) × S	(7) Ride	(8) × S	(9) Skate	(10) × S	(11) Exp	(12) × S	(13) Rockef.	(14) × S	(15) Flood	(16) × S	(17) Covid	(18) × S
Dep. var.	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b
Single (S)	-0.388**	0.242	1.681**	15.844*	-0.738*	0.705	-0.463	0.364	-0.915**	-2.952	-2.726**	15.673	-0.697	-4.636	0.080	3.449	0.173	-11.188
	(0.168)	(1.029)	(0.709)	(9.267)	(0.444)	(2.659)	(0.395)	(2.387)	(0.454)	(2.737)	(1.159)	(4272.9)	(0.496)	(4.343)	(0.482)	(3.633)	(0.485)	(4160.2)
Probability	-0.029***	-0.024***																
	(0.004)	(0.005)																
Age	0.025**	0.031**	0.161**	0.494	0.032	0.004	0.036	-0.006	0.047	0.046	0.044		0.070	0.294	0.001	0.185	-0.079*	-2.762
	(0.013)	(0.014)	(0.071)	(0.889)	(0.042)	(0.080)	(0.039)	(0.070)	(0.045)	(0.082)	(0.065)		(0.060)	(0.380)	(0.047)	(0.169)	(0.048)	(333.06)
Female	0.172	0.002	1.645**	5.837**	0.431	-0.489	0.339	0.111	0.158	-0.059	-2.189**	-21.221	0.261	0.753	0.542	1.739	1.053	15.942
	(0.204)	(0.195)	(0.828)	(2.385)	(0.579)	(0.901)	(0.464)	(0.620)	(0.576)	(0.714)	(1.113)	(3104.2)	(0.595)	(0.773)	(0.538)	(1.288)	(0.656)	(2179.3)
Education	-0.184**	-0.157**	-0.582**	0.894	-0.336*	-0.231	-0.203	-0.092	-0.203	-0.272	-0.196	5.892	-0.567**	-2.734	-0.067	-0.805	-0.199	6.145
	(0.073)	(0.068)	(0.278)	(4.267)	(0.195)	(0.396)	(0.171)	(0.328)	(0.188)	(0.373)	(0.274)	(800.48)	(0.235)	(1.957)	(0.209)	(0.756)	(0.208)	(1086.5)
CRT-naïve	-0.043	-0.066	0.031	2.802	-0.305	-0.521	-0.031	-0.265	-0.233	-0.183	0.480	3.878	-0.181	-0.667*	0.010	-0.169	0.232	-1.780
	(0.073)	(0.104)	(0.260)	(2.013)	(0.204)	(0.388)	(0.175)	(0.283)	(0.202)	(0.359)	(0.322)	(4049.4)	(0.223)	(0.392)	(0.223)	(0.398)	(0.262)	(2.714)
CRT	0.059	0.044	0.064	-2.029	0.395**	0.863*	0.023	0.142	0.261	0.204	-0.242	-3.783	0.180	1.451	-0.105	-0.142	-0.083	3.866
	(0.070)	(0.089)	(0.232)	(1.285)	(0.197)	(0.503)	(0.160)	(0.241)	(0.186)	(0.318)	(0.258)	(1537.0)	(0.216)	(1.236)	(0.210)	(0.397)	(0.227)	(919.49)
Time	0.000	0.000	0.001**	0.005	-0.000	0.000	-0.000	-0.000	0.000	-0.000	0.001	-0.001	0.000	0.002	0.001	0.001	0.000	0.033
	(0.000)	(0.000)	(0.001)	(0.003)	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)	(6.668)	(0.001)	(0.004)	(0.000)	(0.001)	(0.000)	(4.062)
Prob. Music			-0.069***	-0.612														
			(0.021)	(0.534)														
Prob. Acc					-0.029***	0.000												
					(0.010)	(0.021)												
Prob. Ride							-0.020**	-0.010										
							(0.009)	(0.011)										
Prob. Skate									-0.032***	-0.026**								
									(0.009)	(0.012)								
Prob. Exp											-0.059***	-0.252						
											(0.018)	(41.594)						
Prob. Rock.													-0.041***	-0.083				
													(0.012)	(0.057)				
Prob. Flood															-0.017*	-0.022		
															(0.010)	(0.024)		
Prob. Covid																	-0.036***	-0.561
																	(0.011)	(64.064)
S × Prob		-0.013*																
		(0.007)																
S × Age		-0.013		-0.404		0.061		0.048		0.008				-0.223		-0.232		2.704
		(0.024)		(0.896)		(0.102)		(0.090)		(0.105)				(0.386)		(0.181)		(333.060)
S × Female		0.554				1.272		0.666		1.147		56.605				-1.958		-14.973
		(0.449)				(1.386)		(1.024)		(1.494)		(.)				(1.585)		(2179.3)
S × Edu		-0.070		-1.958		-0.243		-0.182		0.038		-10.010		2.267		0.789		-6.370
		(0.139)		(4.359)		(0.486)		(0.404)		(0.455)		(800.48)		(1.952)		(0.806)		(1086.5)
S × CRT-n		0.004		-3.804*		0.278		0.393		-0.439		-4.168		0.452		0.489		2.194
		(0.158)		(2.179)		(0.489)		(0.385)		(0.498)		(4049.4)		(0.445)		(0.532)		(2.734)
S × CRT		0.067		3.069**		-0.546		-0.204		0.449		7.217		-1.278		-0.144		-4.281
		(0.150)		(1.525)		(0.582)		(0.349)		(0.458)		(1537.0)		(1.251)		(0.533)		(919.49)
S × Time		0.000		-0.004		-0.001		-0.000		0.002		0.016		-0.002		0.000		-0.033
		(0.000)		(0.003)		(0.001)		(0.001)		(0.001)		(6.668)		(0.004)		(0.001)		(4.062)
S × Pr. Music				0.495														
				(0.513)														
S × Pr. Acc						-0.038												
						(0.026)												
S × Pr. Ride								-0.048										
								(0.041)										
S × Pr. Skate										-0.038								
										(0.028)								
S × Pr. Exp												-0.625						
												(41.599)						
S × Pr. Rock.														0.047				
														(0.058)				
S × Pr. Flood																-0.005		
																(0.030)		
S × Pr. Covid																		0.521
																		(64.064)
Constant	0.644	0.385	-3.007*	-15.653*	1.412	0.180	0.499	0.427	0.236	1.108	3.898*	5.890	1.962	5.298	-0.991	-2.847	2.762**	14.819
	(0.412)	(0.347)	(1.630)	(8.822)	(1.188)	(1.939)	(1.076)	(1.457)	(1.238)	(1.831)	(2.355)	(4272.9)	(1.338)	(4.051)	(1.379)	(2.541)	(1.399)	(4160.2)
Observations	456	456	57	57	57	57	57	57	57	57	57	57	57	57	57	57	57	57
LL	-229.430	-226.095	-17.427	-8.638	-25.752	-23.915	-31.855	-30.216	-25.707	-23.018	-13.649	-3.155	-20.147	-17.839	-19.772	-17.506	-19.452	-11.472
Pseudo R^2	0.263	0.274	0.542	0.773	0.345	0.391	0.122	0.167	0.349	0.417	0.653	0.920	0.405	0.473	0.145	0.243	0.408	0.651

Note: Censored between 0 and 100. Standard errors in parentheses. Clustered on subject level. Corrected for comprehension control. * $p < .1$, ** $p < .05$, *** $p < .01$.

Table 4.18: Study 4b. Tobit regressions, dep. variable: stated re-use likelihood

	(1) All	(2) × S	(3) Music	(4) × S	(5) Accom	(6) × S	(7) Ride	(8) × S	(9) Skate	(10) × S	(11) Exp	(12) × S	(13) Rock	(14) × S	(15) Flood	(16) × S	(17) Covid	(18) × S
Dep. variable	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}
Single	-1.712	10.354	-2.575	12.148	15.182**	12.867	-8.419	11.216	-8.585	-24.299	-0.676	-29.136	7.453	11.421	-11.678	68.813*	-4.231	19.375
	(4.352)	(26.005)	(9.961)	(48.362)	(6.922)	(34.991)	(7.257)	(34.632)	(9.148)	(47.605)	(9.238)	(46.568)	(8.090)	(43.287)	(8.060)	(40.101)	(8.785)	(45.651)
Age	-0.775*	-0.697	-0.478	-1.777	-0.834	-1.644	-1.090	1.008	-1.688*	-2.982*	-1.942**	-1.905	-1.216	-1.928	0.613	3.358**	0.008	-0.040
	(0.452)	(0.735)	(0.985)	(1.669)	(0.682)	(1.267)	(0.718)	(1.239)	(0.984)	(1.647)	(0.942)	(1.723)	(0.885)	(1.521)	(0.796)	(1.475)	(0.867)	(1.663)
Female	11.141**	5.462	12.177	-6.586	25.760***	17.181	6.006	11.915	14.569	8.575	23.178**	14.658	7.164	1.359	1.199	-3.494	-1.777	-1.141
	(4.485)	(4.408)	(12.103)	(15.062)	(8.604)	(11.173)	(9.049)	(11.125)	(11.236)	(14.370)	(11.413)	(15.038)	(9.802)	(13.066)	(10.011)	(12.804)	(10.865)	(14.543)
Education	-0.592	2.119	5.521	13.963*	-2.273	4.932	-3.305	-8.528	-0.851	8.576	-7.468*	-8.722	1.417	8.434	-3.142	-9.874	6.509	9.488
	(1.960)	(3.889)	(4.457)	(8.097)	(3.077)	(6.026)	(3.246)	(5.819)	(4.258)	(7.701)	(4.187)	(7.993)	(3.650)	(7.057)	(3.600)	(6.779)	(3.920)	(7.843)
CRT-naïve	6.253***	10.856***	12.120**	3.673	7.607**	15.031***	-1.445	5.397	3.662	15.771**	7.197*	15.091**	10.801***	18.140***	4.169	2.165	7.473*	15.060**
	(1.825)	(3.263)	(4.570)	(6.871)	(3.253)	(5.215)	(3.356)	(4.921)	(4.212)	(6.350)	(4.248)	(6.540)	(3.770)	(6.216)	(3.755)	(5.739)	(4.079)	(6.527)
CRT	-5.712***	-9.090***	-9.996**	-3.218	-7.559**	-15.414***	0.089	-5.721	-2.863	-13.059**	-5.975	-10.981*	-9.564***	-14.232**	-4.470	-2.643	-7.143*	-11.475*
	(1.952)	(3.374)	(4.250)	(6.194)	(2.965)	(4.681)	(3.085)	(4.348)	(3.878)	(5.775)	(3.903)	(5.924)	(3.554)	(5.952)	(3.432)	(5.074)	(3.733)	(5.855)
Time	0.003	0.002	-0.004	-0.029	-0.004	0.002	0.004	-0.007	0.008	0.014	-0.006	0.004	0.015*	0.006	0.011	0.011	-0.001	0.012
	(0.005)	(0.008)	(0.010)	(0.020)	(0.007)	(0.014)	(0.008)	(0.014)	(0.010)	(0.018)	(0.010)	(0.018)	(0.008)	(0.015)	(0.008)	(0.016)	(0.009)	(0.018)
Single × Age		-0.366		1.612		0.808		-3.228**		1.465		0.047		0.548		-4.096**		-0.181
		(0.863)		(2.036)		(1.501)		(1.482)		(2.049)		(2.060)		(1.885)		(1.747)		(1.972)
Single × Female		15.103*		29.044		23.804		-6.282		20.040		29.036		11.835		11.617		2.200
		(8.304)		(23.107)		(16.984)		(17.088)		(22.408)		(22.696)		(19.470)		(19.654)		(22.247)
Single × Edu		-4.054		-14.744		-8.941		6.973		-11.563		2.952		-9.588		5.912		-3.827
		(4.356)		(9.552)		(7.043)		(6.889)		(9.274)		(9.400)		(8.300)		(7.988)		(9.180)
Single × CRT-n		-8.415**		13.160		-13.228*		-11.562*		-20.883**		-16.242*		-12.118		2.961		-12.244
		(3.853)		(9.006)		(6.713)		(6.467)		(8.490)		(8.523)		(7.759)		(7.509)		(8.510)
Single × CRT		6.520		-12.522		14.012**		11.551*		18.330**		12.393		7.335		-2.531		6.708
		(3.992)		(8.429)		(6.190)		(5.977)		(8.036)		(8.034)		(7.567)		(6.926)		(7.937)
Single × Time		0.001		0.029		-0.008		0.014		-0.006		-0.016		0.012		-0.004		-0.017
		(0.010)		(0.023)		(0.016)		(0.016)		(0.021)		(0.021)		(0.018)		(0.018)		(0.021)
Constant	56.426***	50.515***	32.047	42.297	56.391***	54.787**	76.128***	62.300**	67.197***	71.708**	130.857***	137.236***	25.391	28.120	37.785*	-2.648	32.525	17.410
	(13.425)	(19.016)	(25.233)	(32.438)	(17.611)	(24.519)	(18.471)	(23.748)	(23.995)	(31.800)	(23.572)	(32.832)	(21.596)	(29.506)	(20.412)	(27.675)	(22.292)	(31.447)
Observations	456	456	57	57	57	57	57	57	57	57	57	57	57	57	57	57	57	57
Log likelihood	-1952.110	-1947.623	-238.417	-234.206	-230.858	-227.675	-256.352	-251.433	-227.147	-224.139	-241.262	-238.864	-175.022	-173.293	-258.932	-255.574	-262.974	-262.275
Pseudo R^2	0.010	0.012	0.022	0.039	0.041	0.055	0.017	0.035	0.020	0.033	0.037	0.047	0.040	0.050	0.012	0.025	0.015	0.017

Note: Censored between 0 and 100. Standard errors in parentheses. Clustered on subject level. Corrected for comprehension control. * $p < .1$, ** $p < .05$, *** $p < .01$.

Study 5

Table 4.19: Study 5a. Probit regressions, dep. variable: preference for short subscription

	(1) All	(2) × S	(3) Music	(4) × S	(5) Ride	(6) × S	(7) Rockef.	(8) × S	(9) Flood	(10) × S
Dependent variable	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b
Single	-0.027	-0.798	-0.177	0.014	-0.333	-2.549	0.563	-1.849	0.060	-0.924
	(0.128)	(0.713)	(0.300)	(1.904)	(0.331)	(2.307)	(0.506)	(3.995)	(0.339)	(1.958)
Probability	-0.030***	-0.034***								
	(0.004)	(0.004)								
Age	0.017	0.017	0.012	0.021	0.025	0.043	0.103*	0.077	0.032	0.043
	(0.011)	(0.012)	(0.022)	(0.034)	(0.023)	(0.041)	(0.062)	(0.083)	(0.024)	(0.034)
Female	0.442***	0.542**	0.802**	0.993*	0.775**	1.331**	1.596**	1.085	-0.630	-0.754
	(0.132)	(0.213)	(0.325)	(0.529)	(0.353)	(0.662)	(0.674)	(0.757)	(0.410)	(0.690)
Education	0.001	-0.077	0.028	-0.026	0.061	-0.240	-0.161	-0.091	-0.067	-0.242
	(0.050)	(0.080)	(0.115)	(0.199)	(0.118)	(0.245)	(0.207)	(0.290)	(0.124)	(0.215)
CRT-naïve	0.053	0.191	0.271	0.011	-0.228	0.746	0.091	-0.718	-0.011	0.626
	(0.145)	(0.213)	(0.312)	(0.516)	(0.330)	(0.645)	(0.477)	(0.707)	(0.333)	(0.607)
CRT	-0.160	-0.257	-0.432	-0.083	0.357	-0.723	-0.290	0.717	-0.194	-0.845
	(0.130)	(0.217)	(0.317)	(0.512)	(0.347)	(0.650)	(0.484)	(0.710)	(0.338)	(0.627)
Risk	-0.055	-0.054	-0.105	-0.127	-0.144	-0.133	0.056	-0.042	-0.030	-0.005
	(0.040)	(0.047)	(0.099)	(0.154)	(0.106)	(0.169)	(0.153)	(0.185)	(0.104)	(0.160)
Prob. Music			-0.034***	-0.037***						
			(0.006)	(0.010)						
Prob. Ride					-0.037***	-0.048***				
					(0.006)	(0.012)				
Prob. Rock							-0.049***	-0.061***		
							(0.014)	(0.021)		
Prob. Flood									-0.035***	-0.026*
									(0.013)	(0.014)
Single × Prob		0.006								
		(0.007)								
Single × Age		-0.000		-0.024		-0.019		-0.050		-0.004
		(0.020)		(0.047)		(0.051)		(0.125)		(0.053)
Single × Female		-0.163		-0.379		-0.626		0.000		0.200
		(0.278)		(0.689)		(0.811)		(.)		(0.891)
Single × Education		0.132		0.126		0.372		0.112		0.235
		(0.100)		(0.251)		(0.291)		(0.612)		(0.277)
Single × CRT-naïve		-0.180		0.576		-1.380*		0.000		-1.152
		(0.284)		(0.682)		(0.792)		(.)		(0.793)
Single × CRT		0.128		-0.692		1.592**		0.000		1.278
		(0.276)		(0.673)		(0.806)		(.)		(0.797)
Single × Risk		0.015		0.037		0.036		0.545		-0.005
		(0.082)		(0.208)		(0.229)		(0.613)		(0.223)
Single × Prob. Music				0.007						
				(0.012)						
Single × Prob. Ride						0.012				
						(0.015)				
Single × Prob. Rock								-0.042		
								(0.069)		
Single × Prob. Flood										-0.075
										(0.051)
Constant	0.540	0.970*	0.901	0.847	1.082	2.590	-0.534	0.177	-0.586	-0.075
	(0.381)	(0.526)	(0.947)	(1.337)	(1.029)	(1.872)	(1.671)	(2.099)	(0.933)	(1.405)
Observations	456	456	114	114	114	114	114	73	114	114
Log likelihood	-225.101	-223.753	-49.180	-47.976	-42.515	-39.681	-20.805	-15.169	-40.500	-36.776
Pseudo R-squared	0.284	0.288	0.377	0.392	0.447	0.484	0.486	0.556	0.235	0.305

Note: Censored between 0 and 100. Standard errors in parentheses. Clustered on subject level. Corrected for comprehension control. * $p < .1$, ** $p < .05$, *** $p < .01$.

Table 4.20: Study 5b. Probit regressions, dep. variable: preference for short subscription

	(1) All	(2) × S	(3) Music	(4) × S	(5) Ride	(6) × S	(7) Rock	(8) × S	(9) Flood	(10) × S
Dependent variable	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b	L_b
Single	-0.119	0.355	0.350	2.758	-0.741**	-5.341*	-0.312	3.485	0.345	-2.561
	(0.150)	(1.283)	(0.307)	(1.714)	(0.342)	(3.101)	(0.349)	(2.776)	(0.320)	(2.321)
Prob	-0.031***	-0.029***								
	(0.004)	(0.004)								
Age	0.016*	0.018	0.026	0.018	0.017	0.032	0.097**	0.184*	0.009	-0.050
	(0.010)	(0.013)	(0.021)	(0.031)	(0.020)	(0.029)	(0.045)	(0.110)	(0.020)	(0.055)
Female	0.068	0.320*	-0.136	-0.021	0.610	1.266**	0.316	1.288	-0.539	-1.122
	(0.172)	(0.191)	(0.328)	(0.489)	(0.376)	(0.592)	(0.410)	(0.952)	(0.365)	(0.814)
Education	-0.012	-0.025	0.041	0.152	0.012	-0.189	-0.036	-0.153	-0.099	-0.014
	(0.068)	(0.071)	(0.121)	(0.187)	(0.132)	(0.184)	(0.131)	(0.236)	(0.126)	(0.226)
CRT-naïve	0.149	0.106	0.529*	0.312	-0.290	-0.571	-0.063	-0.379	0.441	1.262*
	(0.147)	(0.177)	(0.320)	(0.444)	(0.344)	(0.480)	(0.364)	(0.578)	(0.333)	(0.646)
CRT	-0.200	-0.173	-0.593*	-0.309	-0.065	0.166	0.097	0.375	-0.369	-1.206*
	(0.144)	(0.180)	(0.329)	(0.444)	(0.359)	(0.460)	(0.372)	(0.579)	(0.341)	(0.656)
Risk	-0.015	0.002	-0.061	0.053	-0.007	-0.123	-0.043	0.013	-0.020	-0.083
	(0.052)	(0.056)	(0.092)	(0.125)	(0.104)	(0.134)	(0.106)	(0.160)	(0.103)	(0.193)
Prob. Music			-0.033***	-0.033***						
			(0.006)	(0.010)						
Prob. Ride					-0.042***	-0.037***				
					(0.008)	(0.010)				
Prob. Rock							-0.027***	-0.029**		
							(0.007)	(0.013)		
Prob. Flood									-0.019***	-0.029**
									(0.007)	(0.014)
Single × Prob		-0.007								
		(0.010)								
Single × Age		-0.001		0.011		0.043		-0.129		0.079
		(0.020)		(0.043)		(0.059)		(0.121)		(0.061)
Single × Female		-0.548*		-0.299		-1.154		-1.709		0.638
		(0.319)		(0.692)		(0.884)		(1.108)		(0.935)
Single × Education		0.009		-0.214		0.593		0.073		-0.076
		(0.149)		(0.257)		(0.373)		(0.298)		(0.283)
Single × CRT-naïve		0.023		0.476		-0.571		0.389		-1.196
		(0.313)		(0.660)		(0.965)		(0.786)		(0.775)
Single × CRT		-0.022		-0.642		0.302		-0.386		1.236
		(0.294)		(0.678)		(1.009)		(0.797)		(0.791)
Single × Risk		-0.028		-0.263		0.473*		-0.046		0.053
		(0.113)		(0.195)		(0.284)		(0.227)		(0.239)
Single × Prob. Music				-0.005						
				(0.014)						
Single × Prob. Ride						-0.056				
						(0.036)				
Single × Prob. Rock								-0.009		
								(0.017)		
Single × Prob. Flood										0.014
										(0.017)
Constant	0.397	0.190	0.026	-1.002	0.662	1.249	-0.494	-2.570	-0.091	1.979
	(0.547)	(0.544)	(0.845)	(1.109)	(0.966)	(1.182)	(1.190)	(2.276)	(0.949)	(1.988)
Observations	464	464	116	116	116	116	116	116	116	116
Log likelihood	-223.950	-221.413	-48.979	-47.489	-40.538	-35.093	-33.130	-30.218	-41.930	-39.700
Pseudo R-squared	0.288	0.296	0.339	0.359	0.436	0.512	0.315	0.375	0.162	0.207

Note: Censored between 0 and 100. Standard errors in parentheses. Clustered on subject level. Corrected for comprehension control. * $p < .1$, ** $p < .05$, *** $p < .01$.

Table 4.21: Study 5a+5b. Tobit regressions, dep. variable: stated re-use likelihood

	Study 5a										Study 5b									
	(1) All	(2) × S	(3) Music	(4) × S	(5) Ride	(6) × S	(7) Rock	(8) × S	(9) Flood	(10) × S	(11) All	(12) × S	(13) Music	(14) × S	(15) Ride	(16) × S	(17) Rock	(18) × S	(19) Flood	(20) × S
Dep. variable	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}	p_{re}
Single (S)	-10.227**	-30.672	2.125	-89.357	-11.352	9.391	-15.537**	11.448	-15.748***	-39.722	-2.013	5.976	1.378	16.443	1.888	45.390	-6.302	25.061	-5.970	-48.080
	(5.082)	(26.099)	(9.350)	(53.874)	(6.895)	(39.394)	(6.889)	(39.817)	(5.907)	(34.243)	(4.311)	(26.566)	(8.276)	(46.937)	(6.027)	(34.763)	(6.607)	(37.777)	(6.568)	(38.194)
Age	0.181	-0.240	1.022	-0.174	-0.704	-0.633	0.029	-0.354	0.363	0.195	-0.319	-0.142	-0.207	0.027	-0.469	-0.438	-0.740	-0.294	-0.011	0.135
	(0.411)	(0.626)	(0.661)	(0.941)	(0.489)	(0.679)	(0.483)	(0.664)	(0.415)	(0.596)	(0.231)	(0.260)	(0.561)	(0.774)	(0.403)	(0.580)	(0.466)	(0.623)	(0.442)	(0.641)
Female	5.561	-1.957	0.579	-15.154	-0.752	-2.110	5.072	-6.645	17.330***	14.979	2.474	10.252	-5.112	15.029	6.372	7.790	5.866	14.929	3.654	5.173
	(4.712)	(7.030)	(9.759)	(14.351)	(7.188)	(10.269)	(7.116)	(9.916)	(6.175)	(9.110)	(5.097)	(7.425)	(9.138)	(12.652)	(6.627)	(9.512)	(7.301)	(10.125)	(7.214)	(10.355)
Education	-0.514	-0.880	3.029	5.297	-1.872	-5.498	-4.428	-1.880	0.832	-1.210	0.795	-0.656	3.077	-1.527	-0.122	1.329	-2.075	-1.684	1.934	-1.090
	(1.940)	(3.461)	(3.640)	(5.433)	(2.699)	(3.889)	(2.684)	(3.769)	(2.326)	(3.454)	(1.882)	(2.038)	(3.229)	(4.441)	(2.365)	(3.362)	(2.529)	(3.499)	(2.588)	(3.691)
CRT-naïve	2.880	7.705	9.657	19.553	-2.600	5.727	4.845	-1.843	1.500	7.524	9.162**	4.829	16.246*	-5.919	10.801*	10.947	3.736	5.361	5.549	7.842
	(5.327)	(8.119)	(9.758)	(14.438)	(7.174)	(10.051)	(7.227)	(9.755)	(6.149)	(8.850)	(4.355)	(6.381)	(8.554)	(11.641)	(6.199)	(8.701)	(6.722)	(9.246)	(6.738)	(9.482)
CRT	-8.016*	-13.531*	-10.642	-26.284*	-5.034	-9.633	-11.927	-6.934	-6.382	-11.483	-10.849**	-6.507	-19.021**	1.912	-11.868*	-13.680	-6.985	-5.805	-5.580	-7.365
	(4.841)	(7.718)	(9.831)	(14.239)	(7.237)	(9.992)	(7.372)	(9.721)	(6.148)	(8.728)	(4.382)	(6.238)	(8.721)	(11.610)	(6.329)	(8.710)	(6.865)	(9.246)	(6.891)	(9.530)
Risk	-3.082*	-1.840	-1.318	-4.754	-4.284**	1.034	-4.424**	-2.318	-2.434	-1.537	4.565***	4.907***	3.484	4.892	6.860***	9.281***	4.294**	3.054	3.686*	2.373
	(1.726)	(2.198)	(2.891)	(3.953)	(2.117)	(2.788)	(2.123)	(2.695)	(1.805)	(2.482)	(1.281)	(1.618)	(2.524)	(3.164)	(1.836)	(2.379)	(2.032)	(2.544)	(2.004)	(2.610)
S × Age		0.791		2.025		0.017		0.729		0.334		-0.279		-0.226		-0.146		-1.089		-0.119
		(0.765)		(1.304)		(0.951)		(0.953)		(0.833)		(0.443)		(1.092)		(0.805)		(0.952)		(0.888)
S × Female		15.566		25.556		9.159		23.381		6.519		-14.445		-35.707**		-3.490		-18.980		-3.262
		(9.515)		(19.407)		(14.097)		(14.155)		(12.524)		(10.117)		(17.699)		(13.146)		(14.321)		(14.392)
S × Edu		1.007		-3.591		6.031		-3.961		3.926		2.480		8.090		-2.886		-1.839		6.268
		(4.023)		(7.253)		(5.271)		(5.260)		(4.711)		(3.732)		(6.303)		(4.734)		(4.971)		(5.211)
S × CRT-n		-6.433		-22.393		-6.265		18.922		-8.957		6.383		39.676**		0.667		-7.157		-7.230
		(9.870)		(19.803)		(14.277)		(14.853)		(12.514)		(9.043)		(16.856)		(12.499)		(13.471)		(13.640)
S × CRT		8.702		32.086		2.396		-13.810		8.361		-7.270		-39.885**		2.294		0.402		6.144
		(9.351)		(19.435)		(14.016)		(14.517)		(12.250)		(8.767)		(17.019)		(12.638)		(13.568)		(13.832)
S × Risk		-3.671		6.635		-12.453***		-8.096*		-2.130		-0.825		-3.795		-5.649		2.930		3.062
		(3.296)		(5.750)		(4.215)		(4.519)		(3.672)		(2.533)		(4.928)		(3.665)		(4.013)		(4.031)
Constant	39.371***	52.384**	-9.168	41.044	90.142***	84.018***	47.355**	39.875	30.713*	43.935*	18.902	15.122	19.375	13.110	32.263*	13.797	16.192	4.473	12.577	29.758
	(13.846)	(20.987)	(28.126)	(39.146)	(20.841)	(27.974)	(20.787)	(27.014)	(17.594)	(24.458)	(13.236)	(14.297)	(24.545)	(30.055)	(17.834)	(22.688)	(19.253)	(23.454)	(19.475)	(24.876)
Obs	456	456	114	114	114	114	114	114	114	114	464	464	116	116	116	116	116	116	116	116
LL	-1844.939	-1839.249	-442.066	-439.034	-510.802	-504.712	-328.733	-324.445	-511.334	-510.087	-1944.929	-1943.234	-484.120	-479.938	-516.682	-515.171	-359.429	-356.838	-526.725	-525.693
Pseudo R^2	0.006	0.009	0.009	0.016	0.008	0.020	0.020	0.032	0.021	0.023	0.007	0.008	0.009	0.017	0.021	0.024	0.015	0.022	0.006	0.008

Note: Censored between 0 and 100. Standard errors in parentheses. Clustered on subject level. Corrected for comprehension control. * $p < .1$, ** $p < .05$, *** $p < .01$.

Bibliography

Agranov, M., Caplin, A., and Tergiman, C. (2015). Naive play and the process of choice in guessing games. *Journal of the Economic Science Association*, 1(2):146 – 157.

Agranov, M. and Palfrey, T. R. (2020). The effects of income mobility and tax persistence on income redistribution and inequality. *European Economic Review*, 123:103372.

Ahmed, F., Ahmed, N., Pissarides, C., and Stiglitz, J. (2020). Why inequality could spread covid-19. *The Lancet Public Health*.

Aldrich, E. M. and Vargas, K. L. (2020). Experiments in high-frequency trading: comparing two market institutions. *Experimental Economics*, 23(2):322–352.

Alesina, A. and Angeletos, G.-M. (2005). Fairness and redistribution. *American Economic Review*, 95(4):960–980.

Alesina, A. and Giuliano, P. (2011). Preferences for redistribution. In Jess Benhabib, Alberto Bisin, M. J., editor, *Handbook of Social Economics*, volume 1, pages 93–131. Elsevier.

Alesina, A. and La Ferrara, E. (2005). Preferences for redistribution in the land of opportunities. *Journal of Public Economics*, 89(5-6):897–931.

Alesina, A., Stantcheva, S., and Teso, E. (2018). Intergenerational mobility and preferences for redistribution. *American Economic Review*, 108(2):521–54.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine. *Econometrica: Journal of the Econometric Society*, pages 503–546.

Alvaredo, F., Chancel, L., Piketty, T., Saez, E., and Zucman, G. (2018). *World inequality report 2018*. Belknap Press.

Ambühl, M. (2016). Negotiation engineering.

Arnold, T. B. and Emerson, J. W. (2011). Nonparametric goodness-of-fit tests for discrete null distributions. *R Journal*, 3(2).

Arrow, K., Bowles, S., and Durlauf, S. (2000). *Meritocracy and Economic Inequality*. Princeton University Press.

Ashok, V., Kuziemko, I., and Washington, E. (2015). Preferences for redistribution in an era of rising inequality: Some new stylized facts and tentative explanations. *Brookings Papers on Economic Activity (Spring)*, pages 367–405.

Assenza, T., Heemeijer, P., Hommes, C. H., and Massaro, D. (2019). Managing self-organization of expectations through monetary policy: a macro experiment. *Journal of Monetary Economics*.

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory*, 2(3):244–263.

Atkinson, A. B. (2015). *Inequality. What Can Be Done?* Harvard University Press.

Balafoutas, L., Kocher, M. G., Putterman, L., and Sutter, M. (2013). Equality, equity and incentives: An experiment. *European Economic Review*, 60:32–51.

Bao, T., Hennequin, M., Hommes, C., and Massaro, D. (2019). Coordination on bubbles in large-group asset pricing experiments. *Journal of Economic Dynamics and Control*, 36(8):1101–1120.

Bao, T. and Hommes, C. (2015). When speculators meet constructors: Positive and negative feedback in experimental housing markets. *SOM Research Reports*.

Bao, T., Hommes, C., Sonnemans, J., and Tuinstra, J. (2012). Individual expectations, limited rationality and aggregate outcomes. *Journal of Economic Dynamics and Control*, 36(8):1101–1120.

Bartels, L. M. (2018). *Unequal Democracy: The Political Economy of the New Gilded Age*. Princeton University Press.

Bell, D. E. (1982). Regret in decision making under uncertainty. *Operations research*, 30(5):961–981.

Benabou, R. (2000). Unequal societies: Income distribution and the social contract. *American Economic Review*, 90(1):96–129.

Benabou, R. and Ok, E. A. (2001). Social mobility and the demand for redistribution: The POUM hypothesis. *The Quarterly Journal of Economics*, 116(2):447–487.

Benabou, R. and Tirole, J. (2006). Belief in a just world and redistributive politics. *The Quarterly Journal of Economics*, 121(2):699–746.

Benoît, J.-P. and Dubra, J. (2011). Apparent overconfidence. *Econometrica*, 79(5):1591–1625.

Benoît, J.-P., Dubra, J., and Moore, D. A. (2015). Does the better-than-average effect show that people are overconfident?: Two experiments. *Journal of the European Economic Association*, 13(2):293–329.

Bernasconi, M. (2006). Redistributive taxation in democracies: Evidence on people's satisfaction. *European Journal of Political Economy*, 22(4):809–837.

Berninghaus, S. K., Ehrhart, K.-M., and Keser, C. (1999). Continuous-time strategy selection in linear population games. *Experimental Economics*, 2(1):41–57.

Birnbaum, M. H. (1974). The nonadditivity of personality impressions. *Journal of Experimental Psychology*, 102(3):543.

Birnbaum, M. H. (1997). Violations of monotonicity in judgment and decision making.

Birnbaum, M. H., Coffey, G., Mellers, B. A., and Weiss, R. (1992). Utility measurement: Configural-weight theory and the judge's point of view. *Journal of Experimental psychology: human perception and performance*, 18(2):331.

Bleichrodt, H., Doctor, J. N., Gao, Y., Li, C., Meeker, D., and Wakker, P. P. (2019). Resolving rabin's paradox. *Journal of Risk and Uncertainty*, 59(3):239–260.

Bolton, G. E. and Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *American Economic Review*, 90(1):166–193.

Bordalo, P., Coffman, K., Gennaioli, N., and Shleifer, A. (2019). Beliefs about gender. *American Economic Review*, 109(3):739–73.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2012). Salience theory of choice under risk. *The Quarterly journal of economics*, 127(3):1243–1285.

Borghans, L., Heckman, J. J., Golsteyn, B. H., and Meijers, H. (2009). Gender differences in risk aversion and ambiguity aversion. *Journal of the European Economic Association*, 7(2-3):649–658.

Broome, J. (1984). Selecting people randomly. *Ethics*, 95(1):38–55.

Brunnermeier, M. K. and Parker, J. A. (2005). Optimal expectations. *American Economic Review*, 95(4):1092–1118.

Bundesamtfür Statistik (2017). Kennzahlen zur verteilung der einkommen vor und nach staatlichen transfers. https://www.bfs.admin.ch/bfs/de/home/statistiken/querschnittsthemen, accessed January 26, 2019.

Buser, T., Grimalda, G., Putterman, L., and van der Weele, J. (2020). Overconfidence and gender gaps in redistributive preferences: Cross-country experimental evidence. *Journal of Economic Behavior & Organization*, 178:267–286.

Calford, E. and Oprea, R. (2017). Continuity, inertia, and strategic uncertainty: A test of the theory of continuous time games. *Econometrica*, 85(3):915–935.

Cappelen, A. W., Hole, A. D., Sørensen, E. Ø., and Tungodden, B. (2007). The pluralism of fairness ideals: An experimental approach. *American Economic Review*, 97(3):818–827.

Cappelen, A. W., Moene, K. O., Sørensen, E. Ø., and Tungodden, B. (2013). Needs versus entitlements. An international fairness experiment. *Journal of the European Economic Association*, 11(3):574–598.

Case, K. E., Shiller, R. J., and Thompson, A. (2012). What have they been thinking? home buyer behavior in hot and cold markets. Technical report, National Bureau of Economic Research.

Charité, J., Fisman, R., and Kuziemko, I. (2015). Reference points and redistributive preferences: Experimental evidence. *NBER Working Paper No. 21009*.

Charness, G. and Rabin, M. (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics*, 117(3):817–869.

Charness, G., Rustichini, A., and Van de Ven, J. (2018). Self-confidence and strategic behavior. *Experimental Economics*, 21(1):72–98.

Checchi, D. and Filippin, A. (2004). An Experimental Study of the POUM Hypothesis. *Research on Economic Inequality*, 11:115–136.

Chen, D. L., Schonger, M., and Wickens, C. (2016). otree—an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97.

Cheng, S. and Wen, F. (2019). Americans overestimate the intergenerational persistence in income ranks. *Proceedings of the National Academy of Sciences*, 116(28):13909–13914.

Cheung, S. L. and Johnstone, L. (2017). True overconfidence, revealed through actions: An experiment. *IZA Discussion Paper No. 10545*.

Cojocaru, A. (2014). Prospects of upward mobility and preferences for redistribution: Evidence from the life in transition survey. *European Journal of Political Economy*, 34:300–314.

Corneo, G. and Neher, F. (2015). Democratic redistribution and rule of the majority. *European Journal of Political Economy*, 40:96–109.

Cowell, F. A. and Schokkaert, E. (2001). Risk perceptions and distributional judgments. *European Economic Review*, 45(4-6):941–952.

Cruces, G., Perez-Truglia, R., and Tetaz, M. (2013). Biased perceptions of income distribution and preferences for redistribution: Evidence from a survey experiment. *Journal of Public Economics*, 98:100 – 112.

Dafilms (2020). https://dafilms.com.

Davidai, S. and Gilovich, T. (2015). Building a more mobile America — One income quintile at a time. *Perspectives on Psychological Science*, 10(1):60–71.

De Bondt, W. F. and Thaler, R. H. (1995). Financial decision-making in markets and firms: A behavioral perspective. In R. A. Jarrow, V. M. and Ziemba., W. T., editors, *Handbooks in Operations Research and Management Science*, volume 9, pages 385–410. Elsevier.

DellaVigna, S. (2009). Psychology and economics: Evidence from the field. *Journal of Economic literature*, 47(2):315–72.

Diamond, P. A. et al. (1967). Cardinal welfare, individualistic ethics, and interpersonal comparison of utility: Comment. *The Journal of Political Economy*, 75(5):765.

Downs, A. (1957). *An Economic Theory of Democracy*. Harper: New York.

Durante, R., Putterman, L., and van der Weele, J. (2014). Preferences for redistribution and perception of fairness: An experimental study. *Journal of the European Economic Association*, 12(4):1059–1086.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *The quarterly journal of economics*, pages 643–669.

Evans, G., Hommes, C., McGough, B., and Salle, I. (2019). Long-horizon expectations: a lab experiment. *working paper*.

Fang, H., Shapiro, D., and Zillante, A. (2016). An experimental study of alternative campaign finance systems: Transparency, donations, and policy choices. *Economic Inquiry*, 54(1):485–507.

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

Fields, G. S. and Ok, E. A. (1999). The measurement of income mobility: an introduction to the literature. In *Handbook of Income Inequality Measurement*, pages 557–598. Springer.

Filmingo (2020). Subscription plans.

Financial Times (2019). 'fat finger' error sends airline stock on wild ride. https://www.ft.com/content/56188540-a2cf-11e9-974c-ad1c6ab5efd1, accessed December 8, 2020.

Finn, A. and Leibbrandt, M. (2013). Mobility and inequality in the first three waves of nids. *NIDS Discussion Paper*, 120.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178.

Fong, C. (2001). Social preferences, self-interest, and the demand for redistribution. *Journal of Public Economics*, 82(2):225–246.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic perspectives*, 19(4):25–42.

Friedman, D., Huck, S., Oprea, R., and Weidenholzer, S. (2015). From imitation to collusion: Long-run learning in a low-information environment. *Journal of Economic Theory*, 155:185–205.

Frohlich, N. and Oppenheimer, J. A. (1993). *Choosing justice: An experimental approach to ethical theory*, volume 22. Univ of California Press.

Frydman, C. and Mormann, M. M. (2018). The role of salience in choice under risk: An experimental investigation. *Available at SSRN 2778822*.

Garber, P. M. (1989). Tulipmania. *Journal of political Economy*, 97(3):535–560.

Gee, L. K., Migueis, M., and Parsa, S. (2017). Redistributive choices and increasing income inequality: Experimental evidence for income as a signal of deservingness. *Experimental Economics*, 20(4):894–923.

Gilbert, D. T., Morewedge, C. K., Risen, J. L., and Wilson, T. D. (2004). Looking forward to looking backward: The misprediction of regret. *Psychological Science*, 15(5):346–350.

Gilboa, I., Schmeidler, D., and Wakker, P. P. (2002). Utility in case-based decision theory. *Journal of Economic Theory*, 105(2):483–502.

Gill, D. and Prowse, V. (2012). A structural analysis of disappointment aversion in a real effort competition. *American Economic Review*, 102(1):469–503.

Gill, D. and Prowse, V. (2019). Measuring costly effort using the slider task. *Journal of Behavioral and Experimental Finance*, 21:1–9.

Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association*, 1(1):114–125.

Großer, J. and Reuben, E. (2013). Redistribution and market efficiency: An experimental study. *Journal of Public Economics*, 101:39–52.

Haas, G., Ezekiel, M., et al. (1925). What makes hog prices? *US Dept. of Agriculture, Bureau of Agricultural Economics, Division of ...*.

Hanaki, N., Hommes, C., Kopányi, D., and Tuinstra, J. (2019). The effect of framing on the emergence of bubbles and crashes in a learning to forecast experiment.

Hanau, A. (1927). *Die Prognose der Schweinepreise*. Hobbing.

Harms, P. and Zink, S. (2003). Limits to redistribution in a democracy: A survey. *European Journal of Political Economy*, 19(4):651 – 668.

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

Harsanyi, J. C. (1977). *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*. Cambridge: Cambridge University Press.

Hauser, O. P. and Norton, M. I. (2017). (Mis)perceptions of inequality. *Current Opinion in Psychology*, 18:21 – 25.

He, T.-S. (2020). The framing effect of tax–transfer systems. *Journal of the Economic Science Association*, pages 1–13.

Heemeijer, P., Hommes, C., Sonnemans, J., and Tuinstra, J. (2009). Price stability and volatility in markets with positive and negative expectations feedback: An experimental investigation. *Journal of Economic Dynamics and Control*, 33:1052–1072.

Heger, S. A. and Papageorge, N. W. (2018). We should totally open a restaurant: How optimism and overconfidence affect beliefs. *Journal of Economic Psychology*, 67:177–190.

Hennequin, M. (2018). Experiences and expectations in asset markets: an experimental study. Technical report, Working Paper, University of Amsterdam.

Höchtl, W., Sausgruber, R., and Tyran, J.-R. (2012). Inequality aversion and voting on redistribution. *European Economic Review*, 56(7):1406–1421.

Hommes, C., Kopányi-Peuker, A., and Sonnemans, J. (2020). Bubbles, crashes and information contagion in large-group asset market experiments. *Experimental Economics*, 67(3):1–20.

Hommes, C., Sonnemans, J., Tuinstra, J., and van de Velden, H. (2005). Coordination of expectations in asset pricing experiments. *Review of Financial Studies*, 18(3):955–980.

Hommes, C., Sonnemans, J., Tuinstra, J., and Van De Velden, H. (2007). Learning in cobweb experiments. *Macroeconomic Dynamics*, 11(S1):8–33.

Hoshihata, T., Ishikawa, R., Hanaki, N., Akiyama, E., et al. (2017). Flat bubbles in longhorizon experiments: Results from two market conditions. Technical report, Groupe de REcherche en Droit, Economie, Gestion (GREDEG CNRS), University of Nice Sophia Antipolis, 2017.

Hsee, C. K. (1996). The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational behavior and human decision processes*, 67(3):247–257.

Huber, J. and Kirchler, M. (2012). The impact of instructions and procedure on reducing confusion and bubbles in experimental asset markets. *Experimental Economics*, 15(1):89–105.

Irwin, J. R., Slovic, P., Lichtenstein, S., and McClelland, G. H. (1993). Preference reversals and the measurement of environmental values. *Journal of Risk and Uncertainty*, 6(1):5–18.

Janssen, E. M., Meulendijks, W., Mainhard, T., Verkoeijen, P. P., Heijltjes, A. E., van Peppen, L. M., and van Gog, T. (2019). Identifying characteristics associated with higher education teachers' cognitive reflection test performance and their attitudes towards teaching critical thinking. *Teaching and Teacher Education*, 84:139–149.

Jantti, M., Bratsberg, B., Roed, K., Raaum, O., Naylor, R., Osterbacka, E., Bjorklund, A., and Eriksson, T. (2006). American exceptionalism in a new light: a comparison of intergenerational earnings mobility in the nordic countries, the united kingdom and the united states. *IZA discussion paper*.

Jäntti, M. and Jenkins, S. P. (2015). Income mobility. In Atkinson, A. B. and Bourguignon, F., editors, *Handbook of Income Distribution*, volume 2, pages 807–935. Elsevier.

Jiménez-Jiménez, N., Molis, E., and Solano-García, Á. (2018). The effect of initial inequality on meritocracy: A voting experiment on tax redistribution. *Journal of Economic Behavior & Organization*.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Karadja, M., Mollerstrom, J., and Seim, D. (2017). Richer (and holier) than thou? The effect of relative income improvements on demand for redistribution. *Review of Economics and Statistics*, 99(2):201–212.

Kesternich, M., Lange, A., and Sturm, B. (2018). On the performance of rule-based contribution schemes under endowment heterogeneity. *Experimental Economics*, 21(1):180–204.

Kindle (2020). Subscription plans.

Knetsch, J. L. (1989). The endowment effect and evidence of nonreversible indifference curves. *The american Economic review*, 79(5):1277–1284.

Kocher, M. G. and Sutter, M. (2006). Time is money—time pressure, incentives, and the quality of decision-making. *Journal of Economic Behavior & Organization*, 61(3):375–392.

Kondratieff, N. D. (1979). The long waves in economic life. *Review (Fernand Braudel Center)*, pages 519–562.

Kopányi, D., Rabanal, J. P., Rud, O. A., and Tuinstra, J. (2019). Can competition between forecasters stabilize asset prices in learning to forecast experiments? *Journal of Economic Dynamics and Control*, 109:103770.

Kopányi-Peuker, A. and Weber, M. (2019). Experience does not eliminate bubbles: Experimental evidence. *University of St. Gallen, School of Finance Research Paper*, (2018/22).

Köszegi, B. (2006). Ego utility, overconfidence, and task choice. *Journal of the European Economic Association*, 4(4):673–707.

Kraus, M. W. and Tan, J. J. (2015). Americans overestimate social class mobility. *Journal of Experimental Social Psychology*, 58:101–111.

Krawczyk, M. (2010). A glimpse through the veil of ignorance: Equality of opportunity and support for redistribution. *Journal of Public Economics*, 94(1-2):131–141.

Ku, H. and Salmon, T. C. (2013). Procedural fairness and the tolerance for income inequality. *European Economic Review*, 64:111–128.

Kuznets, S. S. (1930). *Secular movement in production and prices: Their nature and their bearing upon cyclical fluctuations*. Houghton Mifflin and company, Boston.

Lahav, Y. (2011). Price patters in experimental asset markets with long horizon. *Journal of Behavioral Finance*, 12:20–28.

Laméris, M. D., Garretsen, H., and Jong-A-Pin, R. (2020). Political ideology and the intragenerational prospect of upward mobility. *European Journal of Political Economy*, 62:101854.

Lefgren, L. J., Sims, D. P., and Stoddard, O. B. (2016). Effort, luck, and voting for redistribution. *Journal of Public Economics*, 143:89–97.

Li, R., Smith, D. V., Clithero, J. A., Venkatraman, V., Carter, R. M., and Huettel, S. A. (2017a). Reason's enemy is not emotion: Engagement of cognitive control networks explains biases in gain/loss framing. *Journal of Neuroscience*, 37(13):3588–3598.

Li, Z., Rohde, K. I., and Wakker, P. P. (2017b). Improving one's choices by putting oneself in others' shoes–an experimental analysis. *Journal of Risk and Uncertainty*, 54(1):1–13.

Lichtenstein, S. and Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of experimental psychology*, 89(1):46.

List, J. A., Sadoff, S., and Wagner, M. (2011). So you want to run an experiment, now what? some simple rules of thumb for optimal experimental design. *Experimental Economics*, 14(4):439.

Logg, J. M., Haran, U., and Moore, D. A. (2018). Is overconfidence a motivated bias? Experimental evidence. *Journal of Experimental Psychology: General*, 147(10):1445.

Loomes, G. and Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The economic journal*, 92(368):805–824.

Lucas Jr, R. E. (1972). Expectations and the neutrality of money. *Journal of economic theory*, 4(2):103–124.

Luttmer, E. F. and Singhal, M. (2011). Culture, context, and the taste for redistribution. *American Economic Journal: Economic Policy*, 3(1):157–79.

Magnani, J., Gorry, A., and Oprea, R. (2016). Time and state dependence in an ss decision experiment. *American Economic Journal: Macroeconomics*, 8(1):285–310.

Mauersberger, F. and Nagel, R. (2018). Chapter 10 - levels of reasoning in keynesian beauty contests: A generative framework. In Hommes, C. and LeBaron, B., editors, *Handbook of Computational Economics*, volume 4 of *Handbook of Computational Economics*, pages 541 – 634. Elsevier.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In Zarembka, P., editor, *Frontiers in Econometrics*, volume 1, pages 105–142. Academic press.

Mellers, B. A., Richards, V., and Birnbaum, M. H. (1992). Distributional theories of impression formation. *Organizational Behavior and Human Decision Processes*, 51(3):313–343.

Meltzer, A. H. and Richard, S. F. (1981). A rational theory of the size of government. *Journal of Political Economy*, 89(5):914–927.

Müller-Lyer, F. C. (1889). Optische urteilstäuschungen. *Archiv für Physiologie*, pages 263–70.

Moffatt, P. G. (2015). *Experimetrics: Econometrics for experimental economics*. Macmillan International Higher Education.

Moore, D. A. and Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review*, 115(2):502.

Moritz, B., Siemsen, E., and Kremer, M. (2014). Judgmental forecasting: Cognitive reflection and decision speed. *Production and Operations Management*, 23(7):1146–1160.

Muth, J. F. (1960). Optimal properties of exponentially weighted forecasts. *Journal of the american statistical association*, 55(290):299–306.

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica: Journal of the Econometric Society*, pages 315–335.

Neunhoeffer, F. (2018). Why do poor people vote for millionaires?

Norton, M. I. and Ariely, D. (2011). Building a better America — One wealth quintile at a time. *Perspectives on Psychological Science*, 6(1):9–12. PMID: 26162108.

Ofcom.org (2020). Lockdown leads to surge in tv screen time and streaming in wales. https://www.ofcom.org.uk/about-ofcom/latest/media/media-releases/2020/lockdown-leads-to-surge-in-tv-screen-time-and-streaming-in-wales, accessed December 08, 2020.

Palan, S. and Schitter, C. (2018). Prolific. ac—a subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, 17:22–27.

Piketty, T. (1995). Social mobility and redistributive politics. *The Quarterly Journal of Economics*, 110(3):551–584.

Piketty, T. (2014). *Capital in the 21st Century*. Harvard University Press.

Rabin, M. (2000). Risk aversion and expected-utility theory: A calibration theorem. *Econometrica*, 68(5):1281–1292.

Rawls, J. (1971). A theory of justice. *Cambridge, Harvard University Press*.

Rieskamp, J. and Hoffrage, U. (2008). Inferences under time pressure: How opportunity costs affect strategy selection. *Acta psychologica*, 127(2):258–276.

Roemer, J. E. (1998). Why the poor do not expropriate the rich: an old argument in new garb. *Journal of Public Economics*, 70(3):399–424.

Rubinstein, A. (1988). Similarity and decision-making under risk (is there a utility theory resolution to the allais paradox?). *Journal of economic theory*, 46(1):145–153.

Ryvkin, D. and Semykina, A. (2017). An experimental study of democracy breakdown, income and inequality. *Experimental Economics*, 20(2):420–447.

Sausgruber, R., Sonntag, A., and Tyran, J.-R. (2019). Disincentives from redistribution: Evidence on a dividend of democracy. *WU International Taxation Research Paper Series No. 2019-05*.

Sausgruber, R. and Tyran, J.-R. (2005). Testing the mill hypothesis of fiscal illusion. *Public choice*, 122(1-2):39–68.

Sausgruber, R. and Tyran, J.-R. (2011). Are we taxing ourselves?: How deliberation and experience shape voting on taxes. *Journal of Public Economics*, 95(1-2):164–176.

Savage, L. J. (1954). The foundations of statistics.

Schildberg-Hörisch, H. (2010). Is the veil of ignorance only a concept about risk? An experiment. *Journal of Public Economics*, 94(11-12):1062–1066.

Schwardmann, P. and Van der Weele, J. (2019). Deception and self-deception. *Nature Human Behaviour*, 3(10):1055–1061.

Sen, A. K. (1970). *Collective Choice and Social Welfare*. San Francisco: Holden-Day.

Simon, H. A. (1957). *Models of man; social and rational.* New York: John Wiley.

Sinn, H.-W. (1996). Social insurance, incentives and risk taking. *International Tax and Public Finance*, 3(3):259–280.

Skidelsky, R. (2016). Answering the queen's question: New approaches to economic challenges. *OECD*.

Slovic, P., Griffin, D., and Tversky, A. (1990). Compatibility effects in judgment and choice.

Smith, V. L., Suchanek, G. L., and Williams, A. W. (1988). Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica: Journal of the Econometric Society*, pages 1119–1151.

Sonnemans, J. and Tuinstra, J. (2010). Positive expectations feedback experiments and number guessing games as models of financial markets. *Journal of Economic Psychology*, 31(6):964–984.

Statista (2020). Largest one-day point losses of the dow jones industrial average index from 1897 to june 2020. https://www.statista.com/statistics/274327/largest-single-day-losses-of-the-dow-jones-index/, accessed December 08, 2020.

Stieger, S. and Reips, U.-D. (2016). A limitation of the cognitive reflection test: familiarity. *PeerJ*, 4:e2395.

Stiglitz, J. E. (2016). Inequality and economic growth. In Jacobs, M. and Mazzucato, M., editors, *Rethinking Capitalism: Economics and Policy for Sustainable and Inclusive Growth*, pages 134–155. Wiley Blackwell.

Stöckl, T., Huber, J., and Kirchler, M. (2010). Bubble measures in experimental asset markets. *Experimental Economics*, 13(3):284–298.

Swan, L. K., Chambers, J. R., Heesacker, M., and Nero, S. S. (2017). How should we measure Americans' perceptions of socio-economic mobility? *Judgment and Decision making*, 12(5):507.

Thaler, R. (1985). Mental accounting and consumer choice. *Marketing science*, 4(3):199–214.

Thaler, R. H. (2018). From cashews to nudges: The evolution of behavioral economics. *American Economic Review*, 108(6):1265–87.

Toplak, M. E., West, R. F., and Stanovich, K. E. (2011). The cognitive reflection test as a predictor of performance on heuristics-and-biases tasks. *Memory & cognition*, 39(7):1275.

Tversky, A. (1969). Intransitivity of preferences. *Psychological review*, 76(1):31.

Tversky, A. and Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive psychology*, 5(2):207–232.

Tversky, A. and Kahneman, D. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291.

Tversky, A. and Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business*, 59(4 pt 2).

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5(4):297–323.

Tversky, A., Sattath, S., and Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychological review*, 95(3):371.

Tversky, A. and Wakker, P. (1995). Risk attitudes and decision weights. *Econometrica: Journal of the Econometric Society*, pages 1255–1280.

Tyran, J.-R. and Sausgruber, R. (2006). A little fairness may induce a lot of redistribution in democracy. *European Economic Review*, 50(2):469–485.

Von Neumann, J. and Morgenstern, O. (1944). *Theory of games and economic behavior.* Princeton University Press.

WashingtonPost (2020). Subscription plans.

Weber, M. and Schram, A. (2017). The non-equivalence of labour market taxes: A real-effort experiment. *The economic journal*, 127(604):2187–2215.

Wilkinson, R. G. and Pickett, K. E. (2017). The enemy between us: The psychological and social costs of inequality. *European Journal of Social Psychology*, 47(1):11–24.

World Bank (2014). Income shares. https://data.worldbank.org/, accessed January 26, 2019.

Young, M. D. (1958). *The Rise of the Meritocracy*. Thames & Hudson.

Summary

Upon an introductory part (Chapter 1), this dissertation reports three individual experimental studies under the overarching topic of cognitive biases in expectation formation.

Chapter 2 brings existing inequality in South Africa (high) and Switzerland (low) to the lab to study how people's preferences for redistribution change with the level of income inequality, income mobility, uncertainty of initial income positions, and source of income (random or real-effort based). It is found that uncertainty and overconfidence about one's income position undermine demand for redistribution. The effect magnifies with larger income disparity. It further induces a *reverse* POUM effect: since wealth ambitions of rich aspirants are better preserved under low than under high mobility, demand for redistribution grows with the degree of mobility. These results combined propose an *inequality trap*: greater inequality today favors personal income overestimation. Demand for redistribution reduces, in particular with low mobility, which propels advanced inequality tomorrow.

Chapter 3 reports a series of Learning-to-Forecast experiments, which are found to replicate price volatility of demand-driven asset markets quite accurately. Yet, the scope of prior work rarely exceeded 50 decision periods or limited decision time substantially, and thereby neglected two central features of financial markets: long runtime and time pressure. This study investigates whether "bubble and crash" dynamics persist in the long run (150 periods) and how decision time (6 vs. 25 sec per decision) influences market volatility. For the treatment with low time pressure, it reports a tendency of prices converging to their fundamental value in the long run. Parallel to this change in dynamics, it identifies a switch from trend-extrapolating strategies to forecasting strategies that are more adaptive. In contrast, increasing time pressure limits trend-chasing behavior and coordination right from the beginning. Consequently, there is less price volatility and faster convergence to the fundamental value.

Chapter 4 explores a novel menu effect in the context of subscriptions that vi-

olates the transitivity principle of rational choice theory. Providers typically capitalize on arranging offers such that a longer, but costlier option is chosen over the cheaper, but shorter alternative. It is found that sizing the shorter subscription down to single-use raises its attraction. This suspects that the presence of a single-use option prompts rational evaluation based on a realistic estimate to use the subscription again. Instead, when both alternatives represent time spans, an irrational mind may discern them along the same category - referred to as *pigeonholing* - with the consequence that other comparative criteria come to the fore. Two-dimensional models, present in most behavioral theories, fail to explain this type of preference reversal. Inspired by the intuition of transaction utility and the availability heuristic the study proposes a generalization of salience theory to capture the effect of pigeonholing.

Zusammenfassung (German)

Nach dem einleitenden Kapitel 1 setzt sich die vorliegende Dissertation aus drei unabhängigen, experimentellen Studien zum übergreifenden Thema *kognitive Verzerrungen in der Erwartungsbildung* zusammen.

Kapitel 2 berichtet über eine Studie die bestehende Unterschiede in den Einkommensverteilungen in Südafrika (hoch) und der Schweiz (niedrig) mit Teilnehmern in einem Laborversuch simuliert. Darin wird untersucht wie sich der Grad der Einkommensungleichheit und -mobilität in einer Experimentalgesellschaft, die Unwissenheit der persönlichen Einkommensklasse und die Art der Einkommensgenerierung (zufällig oder auf Basis einer Geschicklichkeitsaufgabe) auf Umverteilungspräferenzen auswirken. Es zeigt sich, dass Einkommensungewissheit und eine sich daraus ergebende Überschätzung der eigenen Einkommensposition die Nachfrage nach mehr Umverteilung untergräbt. Der Effekt verstärkt sich mit wachsenden Einkommensunterschieden und induziert außerdem einen umgekehrten POUM-Effekt: Da die Wohlstandsambitionen oberer Einkommenspositionen bei geringerer Einkommensmobilität besser erhalten bleiben, steigt die Nachfrage nach Umverteilung mit dem Grad der Mobilität. Zusammengenommen legen diese Ergebnisse eine Ungleichheitsfalle nahe: Stärkere Einkommensungleichheit begünstigt Einkommensüberschätzung. Die Nachfrage nach Umverteilung sinkt, insbesondere in Gesellschaften mit geringer Einkommensmobilität, was die Ungleichheit vorantreibt.

Kapitel 3 berichtet über eine Reihe von Vorhersage-Experimenten, von denen bekannt ist, dass sie die Preisvolatilität von nachfrageorientierten Märkten recht genau nachbilden. Bisherige Studien haben sich allerdings auf 50 Entscheidungsperioden beschränkt oder die Entscheidungszeiten nicht begrenzt und damit zwei zentrale Merkmale von Finanzmärkten vernachlässigt: lange Laufzeiten und Zeitdruck. Diese Arbeit untersucht, ob "Blasen- und Crash"-Zyklen auf lange Sicht (150 Perioden) fortbestehen und wie die Entscheidungszeit (6 vs. 25 Sekunden pro Entscheidung) die Marktvolatilität beeinflusst. Unter geringem Zeitdruck konvergieren die Preise langfristig zu ihrem Soll-Wert. Parallel zu dieser Änderung der Marktdynamik wird ein Wechsel von trend-extrapolierenden Strategien zu eher adaptiven Prognosestrategien identifiziert. Zunehmender Zeitdruck begrenzt das Trendfolgeverhalten und die Koordination von vornherein. Folglich entsteht weniger Preisvolatilität und der Markt konvergiert schneller zum Soll-Wert.

Kapitel 4 erforscht eine bislang nicht dokumentierte Abo-Falle, die das Transitivitätsprinzip der Rational-Choice-Theorie verletzt. Anbieter von Abonnements können typischerweise Angebote so gestalten, dass eine längere, aber teurere Option einer billigeren, aber kürzeren Alternative vorgezogen wird. Es zeigt sich, dass eine Reduzierung der kürzeren Alternative auf eine einmalige Nutzung die Attraktivität im Vergleich zum längeren Abo erhöht. Dies lässt vermuten, dass die Option einer einmaligen Nutzung eine rationale Bewertung über die Wiederverwendungswahrscheinlichkeit des Abos begünstigt. Wenn hingegen beide Alternativen Zeitspannen (mehr als einmalige Nutzung) darstellen, könnte ein irrationaler Entscheider sie in die gleiche Kategorie einordnen (d.h. „Schubladendenken", in Englisch „Pigeonholing") mit der Folge, dass andere Vergleichskriterien in den Vordergrund treten. Zweidimensionale Modelle, denen die meisten Verhaltenstheorien zu Grunde liegen, können diese Art der Präferenzumkehr nicht erklären. Inspiriert von der Intuition des Transaktionsnutzens (Thaler, 1985) und der Verfügbarkeitsheuristik (Tversky und Kahneman, 1973) schlägt die Studie eine Verallgemeinerung der Salienztheorie vor, um den Effekt des Schubladendenkens zu erfassen.

Samenvatting (Dutch)

Na een inleidend deel, rapporteert deze dissertatie drie individuele experimentele studies onder het overkoepelende onderwerp van cognitieve biases in verwachtingsvorming.

Hoofdstuk 2 brengt bestaande ongelijkheid in Zuid-Afrika (hoog) en Zwitserland (laag) naar het lab om te bestuderen hoe de voorkeuren van mensen voor herverdeling veranderen met het niveau van inkomensongelijkheid, inkomensmobiliteit, onzekerheid van initiële inkomensposities, en bron van inkomen (willekeurig of gebaseerd op reële inspanningen). Gebleken is dat onzekerheid en overmoed over de eigen inkomenspositie de vraag naar herverdeling ondermijnen. Dit effect neemt toe naarmate de inkomensongelijkheid groter is. Het leidt verder tot een omgekeerd POUM effect: aangezien de welvaartsambities van rijke aspiranten beter overeind blijven bij lage dan bij hoge mobiliteit, neemt de vraag naar herverdeling toe met de mate van mobiliteit. De combinatie van deze resultaten leidt tot een ongelijkheidsval: grotere ongelijkheid leidt tot overschatting van het persoonlijk inkomen. De vraag naar herverdeling neemt af, met name bij geringe mobiliteit, waardoor de ongelijkheid morgen groter wordt.

In hoofdstuk 3 wordt verslag gedaan van een reeks "Learning to Forecast"-experimenten. Deze experimenten reproduceren de prijsvolatiliteit van vraaggestuurde activamarkten vrij accuraat. In eerder werk op dit gebied was het aantal perioden zelden meer dan vijftig, en werd ook de tijd waarin een beslissing genomen moest worden niet substantieel beperkt. Hierdoor worden twee belangrijke kenmerken van financiële markten verwaarloosd: het grote aantal periodes en tijdsdruk. Deze studie onderzoekt of de dynamiek van "bubbles" en "crashes" voortduurt op de lange termijn (rond de 150 periodes) en hoe de beslissingstijd (6 vs. 25 seconden per beslissing) de marktvolatiliteit beïnvloedt. Als de tijdsdruk laag is, convergeren prijzen op de lange termijn typisch naar hun fundamentele waarde. Tegelijkertijd verandert de wijze waarop deelnemers toekomstige prijzen voorspellen, van trend-extrapolerende voorspelstrategieën naar strategieën die een meer adaptief karakter hebben. Toenemende tijdsdruk beperkt daarentegen het trendvolgende gedrag en de coördinatie op vergelijkbare strategieën, al vanaf de eerste periodes. Dit leidt tot een lager niveau van prijsvolatiliteit en snellere convergentie naar de fundamentele waarde.

Hoofdstuk 4 onderzoekt een nieuw menu-effect in de context van abonnementen dat in strijd is met het transitiviteitsprincipe van de rationele keuzetheorie. Aanbieders zijn gewoonlijk in staat aanbiedingen zo te arrangeren dat een langere, maar duurdere optie wordt verkozen boven het goedkopere, maar kortere alternatief. Het blijkt dat de aantrekkelijkheid van het kortere abonnement wordt verhoogd als deze voor eenmalig gebruikm is. Dit doet vermoeden dat de aanwezigheid van een optie voor eenmalig gebruik rationele evaluatie uitlokt op basis van een realistische inschatting om het abonnement opnieuw te gebruiken. Wanneer beide alternatieven daarentegen tijdspannes vertegenwoordigen, kan een irrationele geest ze in dezelfde categorie indelen - dit wordt pigeonholing genoemd - met als gevolg dat andere vergelijkingscriteria op de voorgrond treden. Tweedimensionale modellen, aanwezig in de meeste gedragstheorieën, slagen er niet in dit type van voorkeuromkering te verklaren. Geïnspireerd door de intuïtie van transactie nut en de beschikbaarheid heuristiek stelt de studie een veralgemening van de salience theorie voor om het effect van pigeonholing te vatten.

Riepilogo (Italian)

Questa tesi è composta da tre studi di economia sperimentale dedicati al tema dei bias cognitivi nella formazione delle aspettative in vari ambiti delle decisioni economiche individuali. I tre studi sono preceduti da un capito introduttivo che li ricollega al tema generale.

Nel secondo capitolo viene condotto un esperimento per studiare come le preferenze delle persone per la tassazione redistributiva variano a seconda del livello di disuguaglianza (alta-Sudafrica o bassa- Svizzera), del grado di mobilità sociale, dell'incertezza sulla posizioni iniziali del reddito e sulle origini delle disuguaglianze (casuali o basate sulle capacità individuali). I risultati mostrano che incertezza ed eccessiva fiducia nelle proprie capacità riducono la domanda di redistribuzione. L'effetto si amplifica con maggiore disparità di reddito. Emerge inoltre un cosiddetto effetto di POUM (prospect of upwards mobilty) inverso: poiché le ambizioni di ricchezza di coloro che sovrastimano le proprie capacità si preservano meglio in condizioni di bassa che alta mobilità, la domanda di redistribuzione cresce con il grado di mobilità. Considerati nel complesso i risultati dell'esperimento suggeriscono l'esistenza di una trappola della disuguaglianza: un'elevata disuguaglianza oggi favorisce eccesso di fiducia e sovrastima del reddito personale; la domanda di redistribuzione perciò si riduce, in particolare con bassa mobilità, creando le condizioni per una ancora maggiore disuguaglianza domani.

Il capitolo 3 riporta una serie di esperimenti Learning-to-Forecast, che risultano replicare abbastanza accuratamente la volatilità dei prezzi dei mercati di beni guidati dalla domanda. Eppure, l'ambito del lavoro precedente raramente superava i 50 periodi di decisione o limitava il tempo di decisione sostanzialmente, trascurando così due caratteristiche centrali dei mercati finanziari: il lungo tempo di esecuzione e la pressione del tempo. Questo studio indaga se le dinamiche di "bolla e crash" persistono nel lungo periodo (150 periodi) e come il tempo di decisione (6 vs. 25 sec per decisione) influenza la volatilità del mercato. Per il trattamento con bassa pressione temporale, esso riporta una tendenza dei prezzi a convergere al loro valore fondamentale nel lungo periodo. Parallelamente a questo cambiamento nella dinamica, identifica un passaggio da strategie di estrapolazione dei trend a strategie di previsione che sono più adattive. Al contrario, l'aumento della la pressione temporale limita il comportamento e la coordinazione a caccia di tendenze fin dall'inizio. Di conseguenza, c'è meno volatilità dei prezzi e una convergenza più veloce verso il valore fondamentale.

Il capitolo 4 esplora un nuovo effetto menu nel contesto degli abbonamenti che viola il principio di transitività della teoria della scelta razionale. I fornitori tipicamente beneficiano per organizzare le offerte in modo tale che un'opzione più lunga, ma più costosa, venga scelta meno costosa, ma più corta. Si scopre che dimensionare l'abbonamento più breve a uso singolo aumenta la sua attrattiva. Questo sospetta che la presenza di un'opzione monouso spinga a una valutazione razionale basata su una stima realistica di utilizzare l'abbonamento di nuovo. Invece, quando entrambe le alternative rappresentano intervalli di tempo, una mente irrazionale può discernere lungo la stessa categoria - indicata come pigeonholing - con la conseguenza che altri criteri comparativi vengono in primo piano. I modelli bidimensionali, presenti nella maggior parte delle teorie comportamentali, non riescono a spiegare questo tipo di inversione delle preferenze. Ispirato dall'intuizione dell'utilità della transazione e dall'euristica della disponibilità, lo studio propone una generalizzazione della teoria della salienza per catturare l'effetto del pigeonholing.

Università
Ca'Foscari
Venezia

DEPOSITO ELETTRONICO DELLA TESI DI DOTTORATO

DICHIARAZIONE SOSTITUTIVA DELL'ATTO DI NOTORIETA'

(Art. 47 D.P.R. 445 del 28/12/2000 e relative modifiche)

Io sottoscritto …Frieder Neunhoeffer.....……………………………………………………

nato a …Tübingen (Germania)…… (prov. ..Baden-Württemberg) il …28.05.1989....……

residente a …Kirchentellinsfurt (Germania).... in …Schmiedstraße…............. n. …23...…

Matricola (se posseduta) …956400…………….Autore della tesi di dottorato dal titolo:

…Cognitive biases in expectation formation: ....………

…Lab evidence on redistribution preferences, financial forecasting, and subscription traps..

…………………………………………………………………………………………………

Dottorato di ricerca in …Economica…………………………………………………………………

(in cotutela con …Universiteit van Amsterdam..............…………………………………………………)

Ciclo …XXXIII…………………

Anno di conseguimento del titolo …2021…………

DICHIARO

di essere a conoscenza:
1) del fatto che in caso di dichiarazioni mendaci, oltre alle sanzioni previste dal codice penale e dalle Leggi speciali per l'ipotesi di falsità in atti ed uso di atti falsi, decado fin dall'inizio e senza necessità di nessuna formalità dai benefici conseguenti al provvedimento emanato sulla base di tali dichiarazioni;
2) dell'obbligo per l'Università di provvedere, per via telematica, al deposito di legge delle tesi di dottorato presso le Biblioteche Nazionali Centrali di Roma e di Firenze al fine di assicurarne la conservazione e la consultabilità da parte di terzi;
3) che l'Università si riserva i diritti di riproduzione per scopi didattici, con citazione della fonte;
4) del fatto che il testo integrale della tesi di dottorato di cui alla presente dichiarazione viene archiviato e reso consultabile via internet attraverso l'Archivio Istituzionale ad Accesso Aperto dell'Università Ca' Foscari, oltre che attraverso i cataloghi delle Biblioteche Nazionali Centrali di Roma e Firenze;
5) del fatto che, ai sensi e per gli effetti di cui al D.Lgs. n. 196/2003, i dati personali raccolti saranno trattati, anche con strumenti informatici, esclusivamente nell'ambito del procedimento per il quale la presentazione viene resa;
6) del fatto che la copia della tesi in formato elettronico depositato nell'Archivio Istituzionale ad Accesso Aperto è del tutto corrispondente alla tesi in formato cartaceo, controfirmata dal tutor, consegnata presso la segreteria didattica del dipartimento di riferimento del corso di dottorato ai fini del deposito presso l'Archivio di Ateneo, e che di conseguenza va esclusa qualsiasi responsabilità dell'Ateneo stesso per quanto riguarda eventuali errori, imprecisioni o omissioni nei contenuti della tesi;
7) del fatto che la copia consegnata in formato cartaceo, controfirmata dal tutor, depositata nell'Archivio di Ateneo, è l'unica alla quale farà riferimento l'Università per rilasciare, a richiesta, la dichiarazione di conformità di eventuali copie.

Data __09/02/2021________ Firma F. Neu[illegible]________________________________

AUTORIZZO

- l'Università a riprodurre ai fini dell'immissione in rete e a comunicare al pubblico tramite servizio on line entro l'Archivio Istituzionale ad Accesso Aperto il testo integrale della tesi depositata;
- l'Università a consentire:
    - la riproduzione a fini personali e di ricerca, escludendo ogni utilizzo di carattere commerciale;
    - la citazione purché completa di tutti i dati bibliografici (nome e cognome dell'autore, titolo della tesi, relatore e correlatore, l'università, l'anno accademico e il numero delle pagine citate).

DICHIARO

1) che il contenuto e l'organizzazione della tesi è opera originale da me realizzata e non infrange in alcun modo il diritto d'autore né gli obblighi connessi alla salvaguardia di diritti morali od economici di altri autori o di altri aventi diritto, sia per testi, immagini, foto, tabelle, o altre parti di cui la tesi è composta, né compromette in alcun modo i diritti di terzi relativi alla sicurezza dei dati personali;
2) che la tesi di dottorato non è il risultato di attività rientranti nella normativa sulla proprietà industriale, non è stata prodotta nell'ambito di progetti finanziati da soggetti pubblici o privati con vincoli alla divulgazione dei risultati, non è oggetto di eventuali registrazione di tipo brevettuale o di tutela;
3) che pertanto l'Università è in ogni caso esente da responsabilità di qualsivoglia natura civile, amministrativa o penale e sarà tenuta indenne a qualsiasi richiesta o rivendicazione da parte di terzi.

A tal fine:
- dichiaro di aver autoarchiviato la copia integrale della tesi in formato elettronico nell'Archivio Istituzionale ad Accesso Aperto dell'Università Ca' Foscari;
- consegno la copia integrale della tesi in formato cartaceo presso la segreteria didattica del dipartimento di riferimento del corso di dottorato ai fini del deposito presso l'Archivio di Ateneo.

Data __09/02/2021_______ Firma F. Nell

La presente dichiarazione è sottoscritta dall'interessato in presenza del dipendente addetto, ovvero sottoscritta e inviata, unitamente a copia fotostatica non autenticata di un documento di identità del dichiarante, all'ufficio competente via fax, ovvero tramite un incaricato, oppure a mezzo posta

Firma del dipendente addetto ....