University of Zurich[UZH]

**Zurich Open Repository and Archive**
University of Zurich
Main Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2017

# Three Essays on Economic Crises

Brune, Amelie

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: https://doi.org/10.5167/uzh-149103
Dissertation
Published Version

Originally published at:
Brune, Amelie. Three Essays on Economic Crises. 2017, University of Zurich, Faculty of Economics.

# Three essays on economic crises

**Dissertation**
**submitted to the**
**Faculty of Business, Economics and Informatics**
**of the University of Zurich**

to obtain the degree of
Doktorin der Wirtschaftswissenschaften, Dr. oec.
(corresponds to Doctor of Philosophy, PhD)

presented by

Amelie Brune
from Germany

approved in September 2017 at the request of

Prof. Dr. Thorsten Hens
Prof. Dr. Helga Fehr-Duda

The Faculty of Business, Economics and Informatics of the University of Zurich hereby authorizes the printing of this dissertation, without indicating an opinion of the views expressed in the work.

Zurich, 20 September 2017

The Chairman of the Doctoral Board: Prof. Dr. Steven Ongena

## Acknowledgements

This dissertation would not have been possible without the help and support of a number of people who accompanied me along this journey, and whom I would like to mention at this point. As such, I am thankful to ...

... my supervisor and co-author Thorsten Hens for allowing me to pursue my doctoral studies in a relaxed and flexible environment, to follow my research interests, for pushing me out of my comfort zone many times, as well as for finding funding for my doctoral studies. Financial support by the University Research Priority Program "Financial Market Regulation", sub-project "Investor Protection", of the University of Zurich is hereby gratefully acknowledged. I am also thankful to Helga Fehr-Duda, who arrived in just the right moment, for always being there to discuss and solve experimental issues, for making me think in a clear and simple way, and for making me reach the finishing line.

... my co-authors, Marc Oliver Rieger and Mei Wang, for encouraging me to publish the 'War Puzzle' at an early stage of my doctoral studies, and thereby to gain first-hand insights into the nitty-gritty of publishing academic work.

... current and former members of the Chair of Financial Economics at the Department of Banking and Finance for creating a pleasant working environment, a relaxing atmosphere, and many interesting discussions about and beyond academic research. A particular thank you goes to Ferdinand Langnickel for his interest in my research, extensive discussions, and for making me re-think little details that I might not have noticed otherwise. I would also like to thank Kremena Bachmann for her pointed questions about my current state of research, Remo Stössel for moral support, and Stefan Zeisberger for discussing early stages of my research.

... Ernst Fehr for taking the time to provide very valuable comments to some of my research, and for allowing me to present and discuss my work at his chair.

... Martine Baumgartner, the 'fairy godmother' of all administrative issues, for countless enjoyable discussions and for motivating me in the harder times of my doctoral studies. I am also thankful to Cornelia Schnyder for excellent administrative support in conducting my experiments at the Laboratory for Experimental and Behavioral Economics at the Department of Economics of the University of Zurich, and for finding session slots, participants, and money when needed most.

... Cornelia Rösler, my 'partner in crime' in pursuing the PhD studies, for a myriad of fun coffee breaks during the good times of our doctoral studies (and also for the more pensive ones when things just did not seem to work out once again ...), and Sabine Elmiger, the best long-term office colleague one could wish for.

Finally, I would like to thank my parents, Annelore Busche-Brune and Bernd Brune, for enabling and encouraging me to study in Zurich in the first place, and for always making me believe in myself and reach for the stars.

Amelie Brune, Zurich, September 2017

# Contents

[1]This paper is published as Brune, A., Hens, T., Rieger, M.O. et al. Int Rev Econ (2015) 62: 1. https://doi.org/10.1007/s12232-014-0215-7.

# Part I

# Dissertation overview

# Dissertation overview

Crises are an intensely debated topic both inside and outside scientific research. They are frequently observed in areas such as finance, society, and politics. This dissertation takes up the discussion by presenting three recent research projects on selected economic crises of topical interest. Each chapter deals with a particular and, in the broad sense of the term, economic crisis within financial networks, small organizations, or international politics. In this respect, it can be positioned at the intersection of societal and financial topics analyzed by means of experimental and empirical methods.

Chapter 1 is inspired by one of the worst economic and financial crises in modern history which we experienced only recently, and which is comparable only to the Great Depression taking its course with the historical stock market crash in the United States on 24 October 1929. Starting with the US subprime mortgage crisis in 2007, this financial crisis contagiously spread through the network of financial institutions, agents, and securities. One prominent example and element of the financial crisis is bank runs and their possible contagion through financial networks, which is studied in Chapter 1, titled "An experiment on the occurrence and propagation of silent bank runs through incomplete financial networks". There, we design a laboratory experiment where a minimum example of an incomplete network of banking institutions serves as the framework for analyzing the emergence and spread of silent bank runs over time. We find that the immediate experience of declines in net bank profits is an important driving factor in generating bank runs in a dynamic setting. Also, after experiencing a run on a bank with poor performance history, agents tend to withdraw larger amounts from other banks in the time after the run.

Chapter 2, titled "Can scapegoating explain persistent low-level group performance?", builds on the idea of a crisis emerging after a lost football match where the coach of the losing team is replaced. This phenomenon can be observed rather frequently in the daily press. Here, we propose a stylized experiment with a generalized setting to study whether the mechanism behind the dismissal of coaches after poor group performance is scapegoating, and if so, whether this affects future team performance. We find that up to 40% of groups with relative poor performance knowingly assign responsibility for their situation to their coach, when in fact they are fully aware that they caused this result themselves, and are thus scapegoating by definition. We also find that groups who scapegoat are less likely to increase their performance over time.

Finally, Chapter A, titled "The war puzzle: Contradictory effects of international conflicts on stock markets", is jointly written with Thorsten Hens, Marc

Oliver Rieger, and Mei Wang and published in the International Review of Economics in 2015. This chapter studies the response of international stock market indices toward changes in the climate of conflicts while distinguishing between foreseeable and sudden wars. We find that a larger probability of an imminent war decreases stock prices, whereas the actual outbreak of war increases them. In contrast, a sudden and unexpected outbreak of war decreases stock prices.

In conclusion, this dissertation unites research on three interesting topics in current crises on financial networks (Chapter 1), small organizations (Chapter 2), and international politics (Chapter A), with which we hope to be able to contribute to the current crisis literature.

# Part II

# Research papers

# Chapter 1

## An experiment on the occurrence and propagation of silent bank runs through incomplete financial networks

Amelie Brune

**Abstract.** The contribution of this paper is to study the circumstances under which silent bank runs occur and spread through incomplete financial networks. To this end, we design a dynamic laboratory experiment where subjects manage their deposit portfolios across three banks with different economic interdependencies. Treatments differ with respect to the information set available to the subject, consisting of knowledge about the bank projects' return distributions as well as underlying network structures between banks. Our findings suggest that rather than the information set available about banks, the immediate experience of declines in net profits is an important driving factor in generating bank runs in a dynamic setting. After experiencing a run on a bank with poor performance history, individuals tend to withdraw more funds from other banks in the subsequent period. As a tentative recommendation for policy makers, information issued about bank performance shall be organized in a careful manner.

**JEL classification:** C91, D81, G01, G02, G11, G21, G28, L14.

**Keywords:** Laboratory experiments, bank runs, contagion, networks, financial crises, financial market regulation, systemic risk, behavioral finance.

## 1.1 Introduction

Bank runs are an unnerving, but re-occurring phenomenon in financial markets. Like the burst of speculative bubbles, they create severe damage to the economy and have the potential to permanently destroy the agents' trust into the financial system. Not only the circumstance that they often seem to happen unexpectedly, but also the possibility of them spreading though the financial system both add a layer of complexity to this problem.

With the degree of interconnectedness in today's financial markets and society, the risk of contagion of bank runs seems unprecedented in economic history. From an economic point of view, the collapse of any given bank can have detrimental effects on up- and downstream financial institutions. Turning to the depositor's perspective, updating beliefs about bank performance can trigger a bank run, and eroding trust can lead to a spread of bank runs through the system. Therefore, it should be of utmost importance to identify the reasons for bank runs to both occur and spread, and to find ways of efficiently reducing the probability of both.

From a theoretical perspective, bank runs have been studied since the early 1980s. The literature distinguishes between two channels of bank run formation: Bank runs triggered through beliefs about other depositors' withdrawal behavior, and bank runs triggered through beliefs about the bank's solvency.

The first channel is described in the seminal contribution by Diamond and Dybvig [1983], who present their theoretical model of bank runs being caused by a change in individual expectations regarding the withdrawal behavior of other depositors. A bank run is one out of two possible symmetric Nash equilibria that this model generates, with both depositors deciding to withdraw their funds from the bank rather than keeping them. This type of bank run is generally known as self-fulfilling or panic-based, as the belief in an imminent bank run itself induces people to excessively withdraw deposits and therewith cause the bank to collapse. However, there is no objective reason to withdraw deposits from the respective bank such as, for example, the knowledge about an imminent, fundamentals-based bankruptcy. In order to prevent bank runs, the authors propose a temporary suspension of deposit convertibility as well as the introduction of a deposit insurance scheme provided by the government as potential mechanisms. The second channel is postulated by Chari and Jagannathan [1988], who suggest that bank runs are created through updated beliefs about a bank's degree of solvency, thus giving rise to information-based bank runs. They propose a model in which the observation of emerging queues in front of banks informs other depositors about a possible bank insolvency, and also come to the conclusion that policies such as the suspension of convertibility may be a possible means in order to prevent bank runs (cf. Diamond and Dybvig [1983]). We build on this perception of bank run emergence in our experimental design later on, however in the framework of silent bank runs, i.e., where adverse information about bank solvency is not conveyed through lines forming in front of a bank, but rather through individually accessible information about past bank performance.

For the past decade, researchers have been studying the determinants of bank runs by means of controlled laboratory experiments. Where field data of depositors' decisions in the time leading up to and during a bank run is either not available or rather difficult to analyze with respect to causal effects, laboratory experiments can allow for a precise identification of a bank run's trigger, and moreover provide an excellent possibility to generate other unique insights into the mechanism (see Madiès [2006], Schotter and Yorulmazer [2009], and Kiss, Rodríguez-Lara, and Rosa-García [2012]).

Madiès [2006] was the first to experimentally test the Diamond and Dybvig [1983] model of pure-panic bank runs. In his paper, he focuses on deposit insurance and temporary suspension of deposit convertibility, and investigates which degree of deposit insurance is capable of avoiding a self-fulfilling bank run. Abstracting the experimental situation in every period to two individuals whose possible decisions are to withdraw or not to withdraw their deposits at a given bank, this is a one-shot game with simultaneous decisions from a game theoretical point of view. Given the expected payoffs related to each decision and the possible decisions of the other player, each player chooses his best strategy. The result is two Nash equilibria in pure strategies, where both players make the same decision: If both players withdraw their deposits, a bank run occurs; if no player withdraws his deposits, the status quo is maintained. Based on the authors's finding that a decreasing level of partial deposit insurance would increase the probability of the occurrence of a bank run, he argues that only a complete deposit insurance scheme is effectively capable of preventing this type of coordination failure. Once a bank run has started, his experiments show that it can be stopped through a temporary suspension of deposit convertibility.

In later studies, most authors focus on the effects of various determinants on the probability of a bank run at an individual bank, again building on the coordination game by Diamond and Dybvig [1983]. Including the work by Madiès [2006], they provide evidence for a negative effect of liquidity (Madiès [2006]), temporary suspension of deposit convertibility (Madiès [2006]), information about the current crisis and the presence of insiders who know about the bank's asset quality (Schotter and Yorulmazer [2009]), the observability of a depositor's withdrawal actions by other depositors who are positioned behind him within a social network (Kiss, Rodríguez-Lara, and Rosa-García [2014a]), cognitive abilities (Kiss, Rodríguez-Lara, and Rosa-García [2016]), deposit insurance (Madiès [2006], Schotter and Yorulmazer [2009], Kiss, Rodríguez-Lara, and Rosa-García [2012]), and knowledge about group financial literacy in large banks (Campioni, Larocca, Mirra, and Panaccione [2017]). Schotter and Yorulmazer [2009] provide evidence even for the case of dynamic, information-based bank runs. Unlike their pure-panic counterpart, in this type of bank runs depositors' withdrawals are motivated by their knowledge about an actually impending insolvency of the bank and are thus objectively founded. A positive effect on the probability of bank run occurrence is found for macroeconomic uncertainty (Garratt and Keister [2009]), the number of withdrawal opportunities (Garratt and Keister [2009]), the degree of coordination required among depositors (Arifovic, Jiang, and Xu [2013]), the level of risk sharing, i.e., the expected payoff in the event of early withdrawals (Klos and Sträter [2013]), and the knowledge of group financial literacy in small banks (Campioni, Larocca, Mirra, and Panaccione [2017]). Public announcements in times of uncertainty may either increase or decrease the likelihood of bank runs depending on their wording and publication circumstances (Arifovic and Jiang [2014]). Gender differences do not seem to matter (Kiss, Rodríguez-Lara, and Rosa-García [2014b]). Also, the results about the effect of deposit insurance are still somewhat mixed, since Kiss, Rodríguez-Lara, and Rosa-García [2012] provide evidence for a positive impact on the reduction of the occurrence probability of bank runs. Moreover, the authors can show that the observability of decisions made by other depositors has an effect on the optimal degree of deposit insurance. This result has direct consequences for respective policies, as these can be assessed as the less expensive alternative to avoid bank runs in the case of sequential decisions of depositors.

If the depositors at any given bank lose their trust in their bank's sensible deployment of their deposits, a simultaneous decision to withdraw funds can lead to a bank run, which has the potential to spread through the financial system by means of beliefs updated with respect to bank fundamentals or other depositors' behavior. As such, even previously uninvolved and healthy banking institutions can suffer a run on themselves. A case where an initial local crisis is transmitted through the system is also known as contagion, which was studied by a smaller group of authors in the context of banks runs.

Chakravarty, Fonseca, and Kaplan [2014] and Brown, Trautmann, and Vlahu [2016] study the mechanisms and characteristics of bank run contagion within an experimental two-bank framework. Based on the bank run coordination game by Diamond and Dybvig [1983], both groups of authors construct a banking environment where two banks are either linked through identical asset quality or liquidation values, respectively, or have no such connection. Banks can have two different quality levels. Subjects at the first bank know their bank's fundamentals and make their withdrawal decisions. This outcome is observed by subjects at the downstream bank, upon which they proceed to make their own withdrawal decisions. They do not know any bank's fundamental values. In the paper by Brown, Trautmann, and Vlahu [2016], downstream depositors also know that upstream depositors are aware of their bank's quality when making their withdrawal decisions, as well as their bank's relation to the upstream bank. The authors find evidence for bank run contagion only in the event of existing economic linkages between the two banks. The channel of bank run transmission is reported to be the subjects' updated beliefs regarding the other depositors' likelihood to withdraw their funds, which relates to the concept of panic-based contagion. In contrast, Chakravarty, Fonseca, and Kaplan [2014] find support for contagion taking place even in absence of liquidity dependencies. If banks are connected, they report evidence for information-based contagion based on the mechanism that beliefs about the liquidity level of a bank are responsible for bank run contagion. In the case where banks are independent, they find support for panic-based contagion, i.e., bank runs spreading from the first to the second bank based on updated beliefs about other depositors' withdrawal behavior.

Although both studies are similar in their theoretical setup and research focus, their results are somehow inconclusive, which may also be due to differences in their designs, such as the explicit measurement of beliefs in Brown, Trautmann, and Vlahu [2016], but not in Chakravarty, Fonseca, and Kaplan [2014]. For example, there is evidence for panic-based contagion in both settings of connected and unconnected banks, which makes it rather difficult to draw precise conclusions about the occurrence of information-based and panic-based bank run contagion.

Duffy, Karadimitropoulou, and Parravano [2016] take the experimental analysis of bank run contagion one step further. With their goal of testing the model of financial contagion as proposed by Allen and Gale [2000], they conduct experiments with complete and incomplete symmetric network structures among four banks, where in the former each bank invests a fraction of deposits in every other bank within the network, while in the latter each bank invests only into its immediate neighboring bank. The network pattern is common knowledge among subjects. The authors find that, while a complete interbank network can partially hedge for bank run contagion as initiated through a liquidity shock to one particular bank, incomplete interbank networks always lead to complete contagion through the network. As such, full co-insurance between banking institutions with the aim of increasing their capacity to meet depositors' liquidity demand can be regarded as a means to efficiently share risk within the network. The authors do not address the issue of information-based versus panic-based contagion.

In contrast to the approach which was chosen by the authors of the three papers studying bank run contagion, our paper is not based on the model by Diamond and Dybvig [1983] or Allen and Gale [2000]. Rather than letting subjects play a series of one-shot coordination games, we decide to follow a novel approach by introducing an intrinsically dynamic setting where subjects are free to update their initial portfolio of bank deposits across all operating banks in each of 20 periods. Instead of performing the binary decision whether or not to withdraw all of their funds at a bank, subjects can both choose to withdraw and re-deposit a continuum of funds into any operating bank. As such, subjects are not restricted to deposit funds at one bank only. We also argue that, by assigning subjects to play the role of specific depositor types (namely, patient and impatient depositors), the use of the model by Diamond and Dybvig [1983] (or extended versions thereof) or by Allen and Gale [2000] in an experimental setting restricts subjects in their portfolio decisions. By lifting these assumptions, we hope to add more realism to the design, especially as these two types vary in each period in the paper by Chakravarty, Fonseca, and Kaplan [2014], thereby restricting realism even further. Moreover, nowadays – at least in the industrialized countries – bank runs do not necessarily have to manifest themselves through visible queues of depositors wishing to withdraw their money in front of cash machines. In times of fast and simple electronic cash transfers accessible for individual depositors, bank runs may be silent, such as the silent run on Wachovia in 2008, for instance. This topical situation is captured by our experimental design, as well: Subjects can only observe the banks' performance and decide on how to manage their private portfolio of cash and bank deposits. However, they are not able to observe other depositors' behavior and actions, mirroring the case of electronic money transfer. This is an important feature incorporated in our experiment.

Similar to Duffy, Karadimitropoulou, and Parravano [2016], we argue that bank run contagion within a network can only be studied effectively when more than two banks are involved. However, in this paper we aim to analyze contagion within an asymmetric network of banking institutions, where banks A and B are connected through economic linkages, but bank C is independent of either bank A or B. This follows up to the afore-mentioned authors' comment that future research could investigate asymmetric financial networks, which we aim to study in a simple setting to begin with. We also add to their contribution by addressing the issue of dynamics as elaborated on in the preceding paragraph.

The goal of this paper is to attempt a reconciliation of the results obtained by previous studies on bank run contagion by introducing an incomplete three-bank network, where bank A and bank B are economically linked while bank C is independent of both bank A and bank B. The idea of introducing networks into economics is nothing else than a generalization of competitive markets or games, which can both be modeled as special types of networks. Kirman [2011] (p. 29) puts it in a nutshell, saying that

> *"[a] perfectly competitive market might be thought of as a star with indi-*

*viduals linked to a central price setter, such as an auctioneer. A game-theoretical model in which the nodes are the individuals would have every node linked to every other one, since all the players consider the actions and reasoning of all the others. This would be a complete graph. The graph corresponding to real economic situations is somewhere in between, with densely linked local clusters linked by longer links to other clusters."*

A network can be characterized by the identity and size of its nodes, their proximity to each other, their connectivity through links as well as these links' length and strength. The propagation of information from one node to another through any network as well as its speed is essentially governed by these parameters. As implied by Kirman [2011], these links between nodes are intrinsically unstable, since they are created and destroyed by economic agents over time. The destruction probability of any link could hence be described as a function of the link's past quality, which can be interpreted as the degree of trust experienced between its corresponding nodes, for instance, in turn depending on the frequency of the link's use and the volume of information proceeded. These properties can be described by the link's thickness. An example for the endogenously triggered instability of any network is the sudden destruction of links connecting a given node to its neighbors. This sudden abandonment of this node can lead to a partial implosion of the network. In economic terms, the resulting destruction of value is even more detrimental if the considered node is located within a sufficiently large number of other nodes in close proximity to and with high connectivity to each other. In extreme cases, this setup can even trigger a collapse of the entire network structure. Since the underlying structure of today's financial markets can perfectly be described by a complicated network, the aforementioned parameters can play a decisive role in governing market outcomes. In fact, they can be regarded as one possible factor for endogenously emerging financial crises.

For our experimental framework, we decide to apply a compact version of the incomplete graph described by Kirman [2011]. The nodes in this network are three different banking institutions of equal size, where two of them are economically connected, which may be thought of as interbank lending, while the third bank is independent from the other two banks. One may argue that it does not seem to be too realistic to include a bank in the network without any connections to other financial institutions. However, experiments which are conducted in the laboratory have to abstract from reality to some extent in order to reduce information complexity for the subjects. The missing link to other banks may be interpreted as a minor connection whose economic or psychological value can be neglected without detriment of the basic framework. The lending volume and general characteristics of the relation between two banks represent the strength and implicitly also the length of this link, where a lower degree of financial activity between any two banking institutions can be represented by a link of increased length. The main difference between Kirman [2011]'s and our interpretation of networks is the different identity of nodes: In our experiment, the nodes in the core network are banking institutions rather than individuals. They establish the experimental framework. Subjects can be interpreted as an outer layer to the core network. As a consequence, the stability of the core financial network is determined exogenously rather than endogenously here. However, the outer layer to the core network does feature endogenous links, i.e., the individually determined connections to different banking institutions. These are in fact created and destroyed by subjects over time by means of their portfolio decisions. As mentioned above, the destruction probability of such a link depends on the degree of trust experienced between the subject and the bank in previous periods.

Our incomplete three-bank network may then be thought of as a minimal example of the graph described above. In this setup, there is scope for both information-based contagion (for instance, from A to B) and panic-based contagion (for instance, from A to C) within the same experiment. From a technical point of view, most previous experimental research on bank runs primarily focused on the outcome of one-shot coordination games set up against the background of bank runs. For a realistic analysis of bank runs, this approach falls short of two important aspects. One, bank runs emerge from a particular investment history combined with unsettled depositors. One-shot games do not adequately cater for the long-run perspective of bank runs, but are rather likely to produce mere happenstance results. And two, one-shot bank run games essentially are coordination games, i.e., the result is some sort of coordination on a certain outcome. The matter of coordination does not need to be the decision whether or not to withdraw deposits from a certain bank, but may evolve around any other type of coordination problem. Hence, any results obtained from such one-shots games do not uniquely have to be related to bank runs. In order to enrich previous research in this respect and make it more realistic, we use a dynamic game with a maximum duration of 20 periods where each bank invests into a profitable one-period project and subjects are free to update their bank portfolios each period.

To summarize, the ultimate purpose of this paper is to identify the circumstances under which bank runs arise and spread through a network of financial institutions. Problems such as the effect of bank runs and their spread to economically related banks on the state of economically independent banks, the probability of depositors running at their economically independent bank after having experienced a bank run at other banks in the system, or the question of how this depends on the knowledge of existing economic linkages are of utmost importance to all stake-holders. It may enable them to prepare for the event that a given part of the financial system experiences difficulties, and take appropriate precautions to shield healthy, independent banks from suffering from contagion. The knowledge of how exactly it is that depositors behave during a bank run, at a bank concerned or not, and how this is influenced by the existence and knowledge of network structures may help regulators and policy makers in designing more appropriate rules to mitigate the occurrence and spread of bank runs, and therefore reducing harm to the economy as a whole. The economic and psychological damage created through bank runs, or at least their near realization, can be mitigated and the protection of private depositors enhanced.

The remainder of this paper is organized as follows. Section 1.2 describes our experimental hypotheses. Section 1.3 describes the experimental design, Section 1.4 presents our behavioral predictions, and Section 1.5 analyses the data obtained from the laboratory experiment. Section 1.6 discusses the results obtained and offers concluding remarks.

## 1.2 Hypotheses

In this paper, we are interested in the occurrence and spread of bank runs within an incomplete three-bank experimental network. We suggest an explorative experiment, as described in the following section, which addresses a variety of issues in previous research as elaborated on in Section 1.1. We propose one variable each to characterize the occurrence and the spread of a bank run, respectively. Regarding the former, we suggest using information about the return distribution of a bank's background investment projects in order to capture bank performance and, thus, the likelihood that a bank may face a run for reasons of depositors' beliefs about deteriorating bank solvency. According to Hens [2008], negative signals about the quality of the bank's investment project can lead to a lack of trust in the bank, which can eventually lead to a bank run and to the bank's liquidation. Regarding the latter, i.e., the spread of a run, we suggest using information about the existence of network structures between banks. As such, we hypothesize that a depositor's knowledge about such existing or missing connections may drive his beliefs about neighboring banks suffering from contagion. Combining these two variables gives us a $2 \times 2$ design which is described in detail in Section 1.4. This is in order to ease the understanding of our behavioral predictions based on the features of the experimental design.

We present our hypotheses in Table 2.1. To anticipate our predictions, we expect economic and psychological spillover effects to be observed between bank A and bank B depending on their timing within the dynamic 20-period framework.

| | |
|---|---|
| **H1:** | If the return distribution of bank projects is known and the network structure among banks is not known, no bank run occurs at any bank. |
| **H2 (a):** | If neither the return distribution of bank projects nor the network structure among banks is known, information-based bank runs at banks A and B occur. |
| **H2 (b):** | If neither the return distribution of bank projects nor the network structure among banks is known, an information-based bank run at bank A does not spread to bank B to cause a panic-based bank run there. |
| **H3 (a):** | If the return distribution of bank projects is not known, but the network structure of banks is, information-based bank runs at banks A and B occur. |
| **H3 (b):** | If the return distribution of bank projects is not known, but the network structure of banks is, an information-based bank run at bank A causes a panic-based bank run at bank B. |

Table 1.1: Experimental hypotheses.

Possible spillover effects with respect to bank C are expected to be of mere psychological nature. We now proceed to the description of the experiment.

## 1.3 Experimental design

Based on our hypotheses presented in Section 1.2, we design our experiment with three treatments. This experiment is a between-subjects design, i.e., each subject only participates in one treatment.

In treatment 1, we randomly assign subjects to groups of six. Within each group, subjects act as depositors within a simple network structure of three symmetric banks A, B, and C. All banks are framed to be independent, i.e., to have no economic linkages between each other. At each bank, each subject has an account at his disposal which is endowed with 1000 ECU[1]. This endowment is a non-accrual loan that has to be paid back at the end of the experiment. Endowing subjects with a given amount deposited on each bank account reduces the probability that they might invest only a small fraction of their available funds into banks at the beginning of the experiment, while waiting and earning the show-up-fee of 20 CHF without risk. We try to make subjects actively participating in the experiment from the beginning and thinking about their actions. Each subject also has a cash account with an initial balance of 0 ECU.

There is a maximum of 20 periods. In each period, each bank offers a fixed interest rate of 5% on deposits. These interest expenditures are financed by means of investment into a profitable one-period project. At the end of each period except the last, each bank invests total deposits into this project, which can be of either good or bad quality with probability $q$ and $1-q$, respectively. $q$ is equal to 0.5 and known to the subjects. Good-quality projects yield a return of 15% and bad-quality

[1] Experimental currency units.

projects yield a return of -5%, which is common knowledge. The expected return thus equals the offered interest rate on deposits. There are no interests on cash holdings.

Each bank has a different path of realized project returns (see Figure 1.1). The depicted bank trajectories are drawn from the population of trajectories based on independent and identically distributed bank project returns. This pre-selection of bank trajectories is motivated by the necessity to be able to compare banks across groups and treatments. Subjects are informed that a bank's project returns can increase or decrease with given probability $q$ and $1 - q$, respectively. The pre-selection of trajectories does not invalidate this information, as it is still true that trajectories were randomly formed in the first place. Figure 1.1 shows that bank A experiences a sequence of negative project returns for the first five consecutive periods, exemplifying a deteriorating economic situation. From period 6 onwards, bank A's project returns revert to a positive trend. Bank B starts with a rather good economic situation. However, from the recovery point of bank A, bank B follows the same downward development which bank A has experienced just before. From period 11 onwards, bank B's project returns also revert to a positive trend. Meanwhile, bank C follows a balanced path throughout the experiment. The figure also shows the theoretical development of deposits for any subject in complete absence of portfolio adjustments.

If subjects would keep their initial portfolio of 1000 ECU at each bank unchanged throughout the experiment, they would maximize their earnings and all banks would be fully able to pay out deposits plus interest income. However, this only serves as a theoretical guideline as subjects are unaware of this fact. The bank trajectories are constructed in such a way that the liquidity levels of all banks at the end of period 20 are identical and slightly larger than total deposits plus compound interest payments.

From the subject's perspective, each period is divided into two stages. In stage 1, subjects see an overview of their current deposits at each bank as well as their current cash holdings. For each bank, they also see a graph displaying the path of past net profits, which are defined as assets[2] minus deposits and compound interest rates. Current net profits are displayed as a number below each graph (see the respective screenshot in Appendix 1.7.2). Subjects then make their withdrawal decisions simultaneously, and any withdrawals are added to their cash account. If total withdrawal requests exceed current liquidity at any bank, a bank run occurs.

[2] Assets consist of deposits plus compound project returns. The terms assets and liquidity are used synonymously.

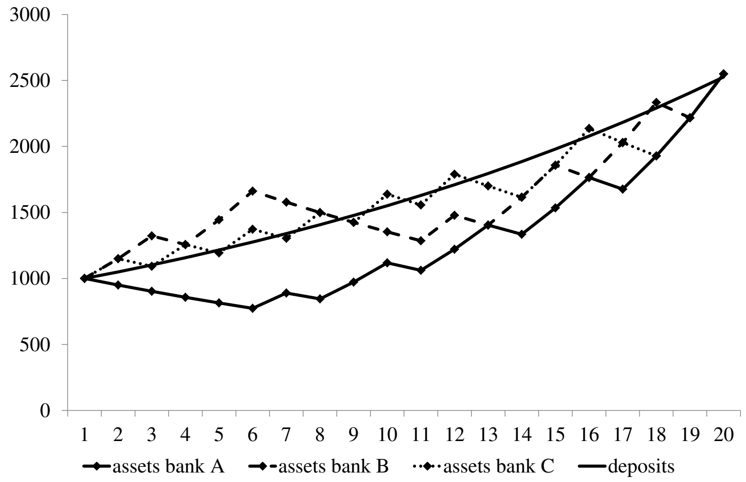



Figure 1.1: Bank trajectories and theoretical development of deposits over time. *Displayed is each bank's trajectory of realized gross project returns as well as the theoretical path of deposit development for any given subject in the case of an unchanged initial deposit portfolio. For each path, the starting point is 1000 ECU, i.e., the initial deposit of any given subject at any given bank in period 1.*

The bank is then liquidated[3] and each subject within the affected group receives the share of liquidity available that is equal to the share of his deposits relative to total deposits. If all three banks are liquidated, the experiment ends. In stage 2, updated information regarding bank balances, cash holdings, and bank performance is provided. The subjects may then use their cash holdings to re-invest into any bank until period 19. As the experiment ends after period 20 by the latest and banks do not invest into any more projects thereafter, there is also no further stage 2 for the subjects in the last round.

Treatment 2 only distinguishes itself from treatment 1 in that subjects do not know the return distribution of the banks' investment projects. Likewise, treatment 3 only distinguishes itself from treatment 2 in that bank A is said to have an economic link to bank B. This economic linkage between bank A and bank B is a framing effect, i.e., its nature is not explained in more detail in order to decrease information complexity for study participants. However, given the form of their return paths as shown in Figure 1.1, bank A and bank B can either be considered

[3]Note that this design feature does not allow us to conduct an analysis as performed in the paper by Chakravarty, Fonseca, and Kaplan [2014], where the persistence of bank runs within one bank over time can be analyzed due to their experimental structure of a repeated one-shot coordination game. However, our design as described before allows for a larger degree of realism.

as associated banks, where bank B's trajectory mirrors the one of bank A with a certain time lag (rather than immediately) and, thus, a decrease in returns of bank B can be interpreted as being fundamentals-based, or one can interpret them to show a mere happenstance correlation of project returns, which allows for relatively large flexibility in the interpretation of our design. The framing in treatment 3 then provides meaning to the return paths as they evolve for the subjects to see over time. In this sense, information given about the return distribution of bank projects as well as about economic connectivity can be both economically and psychologically meaningful.

At the end of each treatment, subjects are asked to fill in a questionnaire which is designed to collect data on socio-economic factors (i.e., gender, age, nationality, field of studies, and monthly disposable income), and furthermore to identify the degree of trust, risk aversion, ambiguity, social preferences, time preferences, loss aversion, and beliefs about the probabilities of a bank run.[4] For a detailed description of these variables, see Appendix 1.7.1.

The subject's earnings are composed of a show-up-fee of 20 CHF, the compound interest on deposits, and any earnings made from the sequence of lotteries as presented in the questionnaire. The exchange rate between ECU and CHF is set to 250:1, i.e., for each 250 ECU (or parts of it) earned, the subject receives (the respective proportion of) 1 CHF in cash. Applying this rate generates an upper bound of 18.32 CHF for the bank run experiment and expected lottery earnings of 0.40 CHF. Total earnings are subject to a lower bound of 10 CHF to ensure that subjects who might conclude the experiment with a particularly low amount are not disappointed and therefore will return to participate in future experiments.

## 1.4 Behavioral predictions

Our two treatment variables are the knowledge of the return probabilities of bank projects, $\mathbf{1}_q$, and the knowledge of economic linkages (or the network structure) between banks, $\mathbf{1}_n$, where $\mathbf{1}$ is the respective indicator function of $q$ and $n$. While $\mathbf{1}_q$ is indicative of the occurrence of bank runs, $\mathbf{1}_n$ is indicative of their spread through the network. In general, we expect bank runs to be more likely to arise if subjects are unaware of the underlying return distribution of bank projects (i.e., $\mathbf{1}_q = 0$), and more likely to spread if the network structure is known (i.e., $\mathbf{1}_n = 1$).

[4] As in Chakravarty, Fonseca, and Kaplan [2014] and Duffy, Karadimitropoulou, and Parravano [2016], we do not directly measure subjects' beliefs about bank solvency or other depositors' likelihood to withdraw funds explicitly. However, we can interpret individual withdrawals as a proxy for a subject's subjective probability that the bank from which he withdraws funds is about to become insolvent. This relates to the channel that beliefs about a bank's solvency are one out of two possible mechanisms for triggering bank runs, as proposed by Chari and Jagannathan [1988].

This leads us to a $2 \times 2$ design which is depicted in Table 1.2. In this paper, we look at three different treatments which we can categorize with respect to the amount of information given to subjects.

| | | $q$ **known** | |
|---|---|---|---|
| | | *No* | *Yes* |
| $n$ **known** | *No* | **treatment 2** (bank run, no spread) | **treatment 1** (no bank run, no spread) |
| | *Yes* | **treatment 3** (bank run, spread) | |

Table 1.2: Overview of the experimental $2 \times 2$ design.
*Displayed are our three treatments based on the $2 \times 2$ design, resulting from the intersection of the knowledge about a bank's project return probability, $q$, and the network structure, $n$. Expectations about the occurrence and spread of bank runs are indicated in parentheses. The intersection where both $q$ and $n$ are know, i.e., the theoretical combination of bank runs spreading after no bank run has occurred before, is not meaningful in practice and is thus not studied any further.*

In treatment 1, relating to our hypothesis H1, subjects have medium information (type 1). They know that each bank's project realizes good and bad outcomes with equal probabilities, and should thus anticipate that even after a sequence of negative returns, every path still has a 50% chance of trending upwards again in the subsequent period. Therefore, it is not rational to run at any point in time. All subjects should decide to trust the banking system and to keep their initial deposits unchanged for the entire duration of the experiment. This way, each subject would maximize his payoff from compound interests on deposits as shown in Figure 1.1. Furthermore, subjects are given the framing that banks do not feature economic linkages. Hence, even in the case that a bank run would occur at any bank, it should at least not spread through the network. This treatment serves as our control treatment.

In treatment 2, relating to our hypothesis H2, subjects have minimum information. The probability distribution of project returns is not disclosed. Risk is thus replaced by ambiguity, which is a key element of our design for the occurrence of bank runs. Note that by doing so, we are able to implement a fairly elegant solution in the sense that we do not have to alter the paths used for any bank, which allows for more experimental control and avoids undesirable treatment effects. We conjecture a negative relation between the project's probability of success, $q$, and the bank's probability of default, $p$, where $p$ depends on the subjects' withdrawal behavior over time and is thus endogenous. Therefore, the risk of default of any bank should now increase with the number of successive negative project returns. Based on the sequence of negative project returns for bank A between period 1 and

6 (see Figure 1.1), we expect to observe a bank run on bank A around period 6 with subjects redistributing their deposits to either the better performing banks B and C or to their cash account. With bank B engaging into a series of negative project returns between periods 6 and 11, we furthermore expect to observe a bank run on bank B. Basically, the timing of this second bank run is informative about the subjects' motivation for withdrawing their funds. If the run on bank B occurs around period 7, then we can characterize it as a panic-based bank run. At this state, there is no need to run from bank B since it only just entered into a period of bad project outcomes. However, if bank A collapses and subjects start panicking, then they might interpret a negative return at bank B as a sign for an imminent bank run. If the run on bank B occurs around period 11, however, then we can interpret it as an information-based bank run. In this case, funds are withdrawn based on the history of negative projects returns and redistributed to the remaining bank C or to the cash account. In this sense, a run on bank A around period 6 would also have to be characterized as information-based, whereas an earlier run would be labeled panic-based. However, since subjects do not know the network structure in this treatment, any observed downstream bank runs should theoretically be of information-based nature as spreads along the lines of economic linkages should not be observed.

Finally, in treatment 3, relating to our hypothesis H3, subjects have medium information (type 2). Knowledge about the underlying network structure is now added to their set of information. Hence, the considerations regarding the spread of bank runs through the financial network made in the previous paragraph should now be confirmed by the data: We expect to see bank runs at bank A, followed by contagion leading to bank runs at bank B.

To summarize, our general prediction is that the probability of a bank run is smallest in treatment 1 due to the considerations made above. While the probability of a bank run is assumed to be larger in both treatments 2 and 3, there is no ex ante conjecture about the relation of their magnitudes. Note that as a rational benchmark, there should be no Bayesian updating at work in this experiment. According to Bayes' theorem,

$$P(X|Y) = \frac{P(Y|X) \cdot P(X)}{P(Y)},$$

where $P(X)$ is the prior belief about $X$ and $P(X|Y)$ is the posterior belief about $X$ after having observed $Y$. Let $X$ be the event where the profitable project has a positive return and let $Y$ be the event where it has a negative return, i.e., $P(X) = P(\text{up})$ and $P(Y) = P(\text{down})$. In the control treatment, we have $P(\text{up}) = q = 0.5$ and $P(\text{down}) = 1 - q = 0.5$. In treatments 2 and 3, $q$ is unknown, but a reasonable heuristic is $q = 0.5$, which is in fact its true value. Inserting these values into Bayes' theorem, this yields

$$P(\text{up} \mid \text{down}) = \frac{P(\text{down} \mid \text{up}) \cdot P(\text{up})}{P(\text{down})},$$

where $P(\text{down} \mid \text{up}) = 0.5$. Obviously, we then have $P(\text{up} \mid \text{down}) = 0.5 = P(\text{up})$, as well. As each trajectory is created as a random sequence of positive and negative returns with equal probabilities, the probability of positive (or negative) returns in any period is independent from previous realizations. Hence, prior and posterior beliefs about the probability of good and bad projects should be equal in theory.

## 1.5 Results

The experiment was programmed using the experimental software z-Tree, the Zurich toolbox for ready-made economic experiments developed by Fischbacher [2007].[5] Subjects were recruited via the University Registration Center for Study Participants of the University of Zurich on condition that they had participated in a laboratory experiment at least once before and that they understand German. This is to ensure that subjects are familiar with the organizational procedures, understand the instructions, and therefore are unlikely to cause delays on the day of the experiment. 12 sessions with a total of 180 subjects were conducted in the Laboratory for Experimental and Behavioral Economics at the Department of Economics at the University of Zurich in May 2016. On average, each session lasted for one hour in which subjects earned 29.10 CHF, including a show-up fee of 20 CHF. The experiment was conducted in German. Detailed written instructions were distributed to the subjects before the start of the experiment (see Appendix 1.7.2). To ensure that each subject understood the upcoming procedures, a set of control questions was distributed which all subjects had to answer correctly (see Appendix 1.7.3). Any remaining open issues were explained by the experimenter.

Out of 180 subjects, 39% are male, 69% from Switzerland, and 37% report a monthly disposable income of up to 500 CHF, where the overall income distribution is right-skewed as it is typical for this type of subject pool. The average age is 23.8 years. Table 1.3 shows the distribution of reported fields of studies. Note that we decide not to exclude economics students from our experiment in order to generate a realistic cross-section of individual backgrounds in our data.

Each treatment was conducted with 60 subjects randomly distributed across 10 groups of 6 depositors each.

While Figure 1.2 shows the total number of bank runs observed at each bank

[5] The code is available upon request.

| Field of studies | |
|---|---|
| Arts, humanities, and political sciences | 13% |
| Computer sciences | 4% |
| Engineering and architecture | 11% |
| Law | 8% |
| Mechanical and electrical engineering | 7% |
| Mathematics | 2% |
| Medical sciences | 8% |
| Natural sciences | 32% |
| Economics | 6% |
| Other/not specified | 9% |

Table 1.3: Overview of the participants' reported fields of studies.
*Displayed is the percentage of subjects reporting a given field of studies.*

across treatments, Figure 1.3 shows the distribution of bank runs over time and across treatments.

Figure 1.2 is informative regarding the hypotheses whether or not bank runs should occur across the different treatments in the first place. Hypothesis H1 has to be rejected for both banks A and B, but is plausible for bank C: If only the bank's return distribution is known to subjects, bank runs do occur at those banks with a distinctive period of poor performance, i.e., at banks A and B, but not at banks which seem to perform relatively stable over time, as bank C does. Hence, subjects seem to react to sequences of bad performance that are visually presented more than to written information about the bank's fundamentals. Both hypothesis H2 (a) and H3 (a) seem plausible: If the bank's return distribution is not known to subjects, bank runs do occur at all banks. This is independent of the knowledge about economic linkages among banks. Moreover, we observe that in each treatment, bank runs at bank A are most frequent. Given our bank trajectories as depicted in Figure 1.1, subjects are probably most nervous about a possible run at bank A given its straightforward decline in assets from the beginning of the experiment. In contrast, at the time bank B experiences a sequence of negative returns, it has already accumulated a good balance. Hence, runs may be relatively less frequent as depositors may be less nervous about poor performance.

Figure 1.3 provides some primary information regarding the contagiousness of bank runs across banks over time. Hypothesis H2 (b) seems plausible: If neither the bank's return distribution nor the network structure among banks is known, there is no observational evidence for spreading bank runs in, as an average of 5 periods elapses between a run on bank A and a run on bank B. The time span leading up to a run on bank B can be interpreted as the time necessary for subjects to decide how to react to developments at bank B itself rather than to past events at bank A, as it is precisely five periods of consecutive poor performance that occur at bank B.

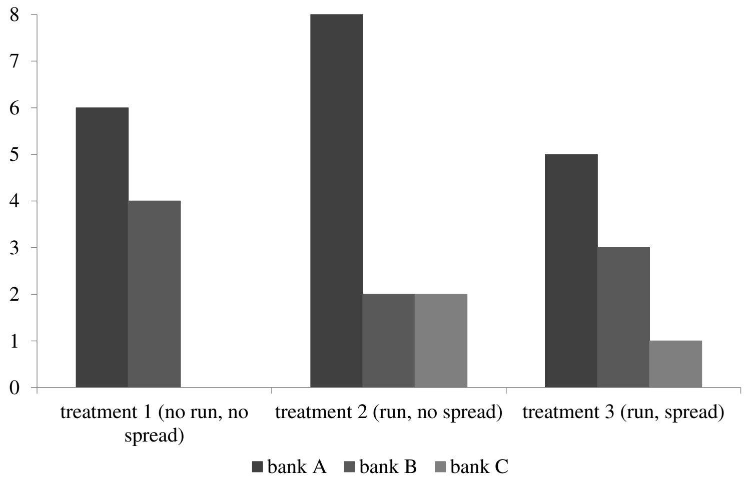
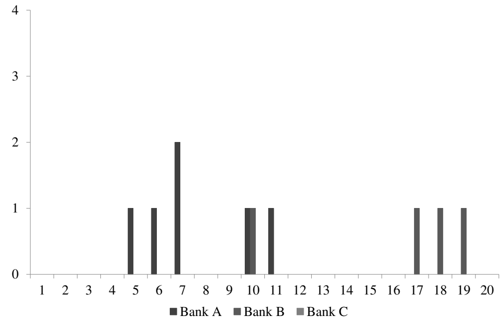


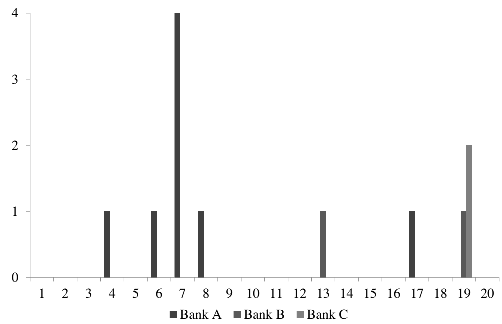

Figure 1.2: Total number of bank runs observed at each bank across treatments. *Displayed is the total number of bank runs observed for each bank in each treatment, with the respective treatment's expectations about the occurrence and spread of bank runs stated in parentheses.*

A contagious bank run is expected to happen sooner. Hypothesis H3 (b) does not seem plausible: The knowledge about network connections does not seem to foster contagion within the system, but rather induce the opposite: Here, an average of 9.5 periods elapses between any run on bank A and bank B, that is, more rather than fewer periods pass between any two runs. However, a caveat is in order at this point: The number of observations of successive bank runs in treatments 2 and 3 is very small, with only one respectively two observations available in the data. The impression of a plethora of bank runs as created by Figure 1.2 is due to the fact that many bank runs occurred in singularity within any group, or that a run on bank A was following rather than preceding a run on bank B.

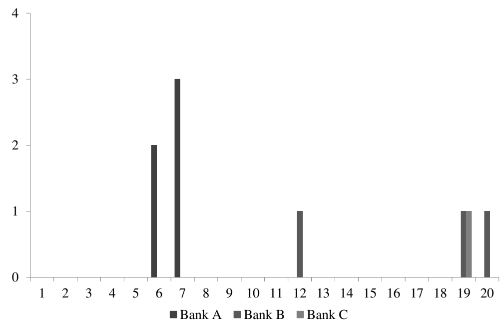
There may be a last-round effect present for bank C where most funds are accumulated by the end of the experiment when subjects might want to secure funds on their cash account, which we study in more detail in the panel analysis at the end of this section. An alternative explanation for late bank runs at bank C could be provided by Duffy, Karadimitropoulou, and Parravano [2016]'s finding that the timing of a downstream bank run is contingent on the concerned bank's proximity to the bank who suffered the initial liquidity shock in the financial system. As bank C is modeled to be independent from banks A and B, this reasoning could apply, as well.




(a) Treatment 1 (no run, no spread)




(b) Treatment 2 (run, no spread)




(c) Treatment 3 (run, spread)

Figure 1.3: Distribution of the number of bank runs over time and treatments. *The respective treatment's expectations about the occurrence and spread of bank runs stated in parentheses.*

The combination of results regarding our three hypotheses reveals, however, that there are no distinct treatment effects to be observed. Hence, in contrast to our initial hypotheses, each treatment features non-spreading bank runs. There are two potential reasons for missing treatment effects in the occurrence and spread of bank runs. Firstly, the number of observations relative to the number of groups is rather small. Re-running the experiment on a larger scale could lead to further insights in this matter. And secondly, the treatment effect is only designed as a frame within the written instructions. As such, subjects might neglect the detail of information about the bank's fundamentals given therein, and instead have focused on new information presented about net profits in each period. The salience of treatments thus could have been stronger. In particular, the distinction between treatments 1 and 2 was the knowledge about the bank's return distribution. If this was unknown as in treatment 2, a sensible benchmark would be to assume $q = 0.5$, which is indeed the true value of $q$. Hence, it is sensible to assume treatments 1 and 2 to be equal. Similarly, assuming $q = 0.5$ and neglecting the information given about network structures in the instructions for treatments 2 and 3, it is furthermore possible to assume these treatments to be equal, as well.

Based on these considerations, and after an in-depth descriptive analysis on potential variations across treatments, we find that the data is extremely homogenous. Therefore, we decide not to conduct any further analyses separately by treatment, but rather to pool the data from all treatments. We now divert our focus toward a unified analysis of withdrawal and depositing behavior and aim at identifying their determinants, which seem to be independent of the information given about fundamentals and/or network connections.

We conduct a unified panel analysis with fixed effects and use two sets of panel regressions with new deposits, $D$, and withdrawals, $W$, as the dependent variables, respectively, which also allows us to better discriminate between actual and quasi bank runs, where the latter describes situations in which total requested withdrawals are very close to assets available at the respective bank. Using withdrawals and new deposits instead of a binary variable indicating the occurrence of a bank run also increases the robustness of our results as we have a total of 180 observations. Our panels are balanced by construction.

We analyze the following specifications of panel regressions:

$$D_{it}^{j} = \alpha_i + \beta_j \pi_t^j + \kappa_j \text{fear}_{it} + \theta_k \text{run}_t^k + \eta t_{20} + \varepsilon_{it} \tag{1.1}$$

and

$$W_{it}^{j} = \alpha_i + \beta_j \pi_t^j + \kappa_j \text{fear}_{it} + \theta_k \text{run}_{t-1}^k + \eta t_{20} + \varepsilon_{it}, \tag{1.2}$$

where $j \in \{A, B, C\}$ and

$$k = \begin{cases} B & \text{if } j = A, \\ A & \text{if } j = B, \\ A, B & \text{if } j = C. \end{cases}$$

Bank assets are given by the sum of deposits plus revenues from projects, and net profits shown to the subjects are determined by the difference between assets and the sum of deposits. Hence, assets are redundant and deposits cancel out. We therefore refrain from including either assets or the sum of deposits into the regression in order to avoid issues of multi-collinearity.

Table 1.4 defines and summarizes the variables used in panels 1.1 and 1.2.

| Variable | Definition |
|---|---|
| $D_{it}^j$ | new deposits made by a subject |
| $W_{it}^j$ | withdrawals made by a subject |
| $\pi_t^j$ | net profits of bank |
| $\text{fear}_{it}$ | dummy variable indicating whether a subject was worried that one or more banks would become bankrupt (see Appendix 1.7.1), with $\text{fear}_{it} = 1$ if the subject was worried about this |
| $\text{run}_t^k$ | dummy variable indicating a bank run, with $\text{run}_t^k = 1$ if a bank run occurred |
| $\text{run}_{t-1}^k$ | dummy variable indicating a bank run, with $\text{run}_{t-1}^k = 1$ if a bank run occurred in the previous period |
| $t_{20}$ | dummy variable indicating the last period, with $t_{20} = 1$ in period 20 |

Table 1.4: Definition of variables used in panels 1.1 and 1.2.
*Index $i$ refers to the subject (with $i \in \{1, 2, \ldots, 180\}$), $t$ to the period (with $t \in \{1, 2, \ldots, 20\}$), and $j$ and $k$ to the bank (with $j \in \{A, B, C\}$ and $k \in \{A, B\}$).*

Note that due to the time scale of events with respect to withdrawals, possible bank runs, and new deposits, panel 1.1 includes $\text{run}_t^k$ as a variable, while panel 1.2 includes its one-period time lag. For $j = k$, $\text{run}_t^k$ is omitted in each panel with $D_{it}^j$ or $W_{it}^j$ as dependent variable for trivial reasons: A run on bank $j$ causes its liquidation, hence no further deposits or withdrawals are possible – hence there is no effect to study. The effect of a run on bank C is hardly observed, and hence not studied.

Subject-specific and hence time-invariant characteristics such as socio-economic status are not included in the panel for reasons of multi-collinearity. Any variation in this respect is covered through fixed effects by $\alpha_i$.

The results for our panel regressions with new deposits and withdrawals at banks A, B, and C as dependent variables are reported in Tables 1.5 and 1.6, respectively, with columns 1, 2, and 3 showing coefficients, clustered standard errors, and levels of significance for all respective variables. We now proceed to interpret the significant coefficients' signs and magnitudes in order to grasp the overall structure underlying the data.

In Table 1.5, we see that there is a highly significant positive relation between

| | $D_{it}^A$ | $D_{it}^B$ | $D_{it}^C$ |
|---|---|---|---|
| $\pi_t^A$ | 0.0199** | −0.0338*** | −0.0473*** |
| | (0.0095) | (0.0117) | (0.0122) |
| $\pi_t^B$ | −0.0183* | 0.0454*** | −0.0575*** |
| | (0.0105) | (0.0056) | (0.0071) |
| $\pi_t^C$ | −0.0772*** | −0.0786*** | 0.0676*** |
| | (0.0200) | (0.0166) | (0.0109) |
| $\text{fear}_{it}$ | −125.7206 *** | −46.4367* | −15.6870 |
| | (33.7814) | (24.9356) | (50.1106) |
| $\text{run}_t^A$ | | 83.9980*** | 170.2416*** |
| | | (29.0827) | (21.3861) |
| $\text{run}_t^B$ | 329.9774* | | 177.6720*** |
| | (171.0208) | | (61.5187) |
| $t_{20}$ | −365.9074 *** | −208.1506 *** | −248.8644 *** |
| | (81.6162) | (41.2396) | (28.2088) |
| constant | 116.3138*** | 115.5748*** | 90.3945*** |
| | (14.8413) | (17.1810) | (17.7956) |
| within $R^2$ | 0.0714 | 0.0473 | 0.0565 |
| $n$ | 180 | 180 | 180 |
| $T$ | 20 | 20 | 20 |

Table 1.5: Fixed-effects panel regression results for new deposits at each bank as dependent variable.

** denotes significance at 10%, ** at 5%, and *** at 1% level. Clustered standard errors are given in parenthesis.*

net profits and new deposits at any bank $j$: The larger net profits at any bank, the larger new deposits at the respective institution. The magnitude of this effect ranges from 20 ECU at bank A to 68 ECU at bank C for each additional 1000 ECU in net profits. Looking at cross-bank effects, we find that an increase in net profits at any bank $k \neq j$ decreases new deposits at bank $j$. Again, the magnitude of this effect ranges from a decrease in new deposits at bank A by 18 ECU if net profits at bank B increase by 1000 ECU to a decrease in new deposits at bank B by 79 ECU if net profits at bank C increase by 1000 ECU. Thus, good performance at any bank causes people to slightly shift their funds from other banks to the well-performing bank, which supports rational behavior.

All coefficients of $\text{run}_t^j$ are positive and significant. A run on bank A causes new deposits to increase by 84 ECU at bank B and by 170 ECU at bank C. Similarly, a run on bank B causes new deposits to increase by 330 ECU at bank A and by 178 ECU at bank C. This is equivalent to saying that any run on a bank with a defined period of poor performance causes people to increase their deposits at other banks. However, in order to fully evaluate this cross-effect with respect to the spread of bank runs through the system, we need to consider the results of Table 1.6 in the following.

If subjects experience fear of an imminent bank run, new deposits significantly

| | $W_{it}^A$ | $W_{it}^B$ | $W_{it}^C$ |
|---|---|---|---|
| $\pi_t^A$ | 0.0544** | −0.0631** | −0.0137 |
| | (0.0218) | (0.0241) | (0.0241) |
| $\pi_t^B$ | −0.0032 | −0.0327*** | −0.0006 |
| | (0.0051) | (0.0075) | (0.0079) |
| $\pi_t^C$ | −0.0296** | 0.0484*** | −0.1665*** |
| | (0.0113) | (0.0134) | (0.0175) |
| $\text{fear}_{it}$ | 157.4749 | −40.7515 | 3.2999 |
| | (106.7036) | (36.3015) | (76.8981) |
| $\text{run}_{t-1}^A$ | | 247.9107*** | 4.8969 |
| | | (67.0356) | (50.7213) |
| $\text{run}_{t-1}^B$ | 296.3202** | | 134.5103 |
| | (139.9333) | | (101.5149) |
| $t_{20}$ | 222.1396 | 260.5101** | −304.9196 *** |
| | (153.8726) | (106.1226) | (62.6673) |
| constant | 229.9988*** | 41.8047 | 160.0316*** |
| | (40.6963) | (42.9036) | (42.7047) |
| within $R^2$ | 0.0446 | 0.0409 | 0.0897 |
| $n$ | 180 | 180 | 180 |
| $T$ | 20 | 20 | 20 |

Table 1.6: Fixed-effects panel regression results for withdrawals at each bank as dependent variable.

** denotes significance at 10%, ** at 5%, and *** at 1% level. Clustered standard errors are given in parenthesis.*

decrease at banks A and B (by order of magnitude of -126 ECU and -46 ECU, respectively). Hence, banks with a troubled history of net profits are more likely to face a decline in their assets than do relatively stable banks (bank C in this setting).

Finally, we find evidence for a last-round effect at all three banks between 208 ECU (bank B) and 366 ECU (bank A) on average.

Table 1.6 shows regression results for withdrawals at any bank, which may be interpreted as subjective probabilities of a bank run to occur, as discussed in Section 1.3. In unison with the results for new deposits, we find that an increase in net profits at banks B and C by 1000 ECU decreases withdrawals from these banks by 33 ECU and 167 ECU, respectively. However, we find the opposite effect for bank A: Here, an increase in net profits by 1000 ECU tends to increase withdrawals by 54 ECU on average. Similarly, the pattern observed for any cross-effects is both less clear and less strong than for new deposits. If bank C performs better, withdrawals from bank A slightly decrease and withdrawals from bank B slightly increase; if bank A performs better, withdrawals from bank B slightly decrease. These effects remain unexplained.

Withdrawals from both banks A and B strongly react to a bank run that occurred at the respective other bank in the previous period. A run on bank A in

$t-1$ causes withdrawals to increase at bank B by 248 ECU on average; a run on bank B in $t-1$ causes withdrawals to increase at bank A by 296 ECU on average. These results support our hypotheses H2 (b) and H3 (b), namely, that a run on any bank does provoke actions to withdraw funds from other banks. Combining these findings with those from Table 1.5, we see that the net withdrawal effect over any two successive periods, where net withdrawals are defined as $W_t - D_{t-1}$, is positive for bank A and negative for bank B, i.e., a run on bank B effectively decreases net withdrawals from bank A, whereas a run on bank A effectively increases net withdrawals from bank B. This means that a run on bank A has in fact a negative overall impact on deposits at bank B. This is evidence for a bank run at bank A causing bank B's assets to decrease in the subsequent period.

Unlike its effect on new deposits, the fear of an imminent bank run does not seem to have any effect on withdrawals themselves.

We observe last-round effects at banks B and C, with 261 ECU withdrawn from bank B and a decrease in withdrawals of 305 ECU at bank C in period 20. This indicates the action of shifting funds from bank B to the stable bank C, which may be regarded as a safe harbor for funds, at the end of the experiment. In this context, we also observe that subjects tend to keep a certain fraction of their wealth on their cash account throughout the experiment. The hypothesis of $H_0$: $\Delta C_t = 0$, where $\Delta C_t$ is defined as the difference between cash holdings at the end and at the beginning of each period, is rejected against the right-sided alternative at the 1%-level of significance (p-value equal to 0.0000). The magnitude of this variable increases with the occurrence of bank runs. Hence, there is mild evidence for a flight into cash in the face of troubled banks and possible runs.

The percentage of variation explained within subjects lies between approximately 4% and 9% for the two panels and is thus rather small, which can be attributed to the analysis of micro-data.

As a robustness check and for statistical control, we expand each panel regression by the inclusion of one individual time lag $\pi^j_{t-l}$ out of a set of five different time lags, where $l \in \{1, 2, \ldots, 5\}$. We find that the effect of a bank run at banks A or B on new deposits at bank C remains significant for all lags $l$. Similarly, their effect on withdrawals from bank C remains insignificant throughout. Including up to four time lags of $\pi^j_{t-l}$ (with the exception of $\pi^j_{t-1}$) keeps the effect of a bank run on bank B on changes in new deposits at bank A significant. A similar finding applies for withdrawals from bank A with up to three time lags keeping the effect significant. However, including any lags of net profit turns the effect of a bank run on bank A on new deposits at bank B insignificant. Only one time lag keeps significance for the respective effect on withdrawals. In conclusion, adding recent lags of net profits keeps our results relatively stable.

## 1.6 Discussion and concluding remarks

In this paper, we suggest a novel dynamic approach to study the occurrence and contagion of silent bank runs within an incomplete network of financial institutions in an experimental setting. We aim to contribute to the recent literature on bank run contagion by addressing a number of issues of previously used theoretical models, which we regard as noteworthy, through an explorative experimental design.

Based on our observations in the experiment, we can draw the following conclusions: When making depositing and withdrawal decisions, bank depositors do not seem to take into consideration objective, written information about a bank's situation, be it about the investment situation of the bank itself or its economic linkages to other banks. Hence, such information does not seem to be able to prevent bank runs. Instead, visually experienced information about current and previous net profits generated by the bank seems to matter more. We find a positive causal relation between net profits and new deposits at any given bank, and a negative relation with respect to new deposits at other banks. Overall, a recent run on a bank with a history of poor performance tends to decrease assets at other banks through increased withdrawals, thereby indicating a tendency for bank runs to create contagious effects through incomplete financial networks, even though the actual occurrence of bank runs does not spread itself. The fear of an imminent bank run also has a negative effect on withdrawal behavior. Healthy banks with relatively stable past performance are not affected by these factors overall. In this respect, regulators might want to think about what kind of bank information is disclosed to the public, and in which way this is communicated and presented.

Both Brown, Trautmann, and Vlahu [2016] and Chakravarty, Fonseca, and Kaplan [2014] build their experiments based on (a modified version of) the two-person coordination game as suggested by Diamond and Dybvig [1983]. While these studies use different variables to model economic linkages (namely, the bank's asset quality and its liquidation value, respectively), the underlying mechanism is the same: The respective variable is identical across banks in treatments with economic linkages, whereas missing economic linkages are realized through independent payoff schemes. This meaning of economic linkages (or the missing thereof) is communicated to subjects in the respective treatments. Building their design upon the model by Allen and Gale [2000], Duffy, Karadimitropoulou, and Parravano [2016] model economic linkages between two banks as either mutual or unilateral investments of one bank's deposits into another bank. In this sense, previous studies realized economic linkages either by means of identical bank fundamentals or by active depositing behavior of banks. In our paper, bank connections are implemented through trajectories of bank project returns, based on which banks may be flexibly interpreted as being linked or independent depending on the framing given in the respective instructions of a treatment. The framing thus provides meaning to the individual return paths. While offering an alternative way to create bank connections in the laboratory with our paper, we suggest that framing shall be made stronger and more salient in future studies.

Regarding the payout scheme in case of a bank run, both subjects in the two-person coordination game in the study by Brown, Trautmann, and Vlahu [2016] receive half of the liquidation value of their bank. In the equivalent situation in Chakravarty, Fonseca, and Kaplan [2014]'s ten-person setup, subjects are paid expected payoffs. In the paper by Duffy, Karadimitropoulou, and Parravano [2016], the payout upon the occurrence of a bank run is an equal share of bank assets for each withdrawing party. There is no direct sequential service constraint implemented. Hence, all three studies rely on a proportionate payout scheme in case of a bank run where all subjects involved decide to withdraw their funds in a given period. However, if the bank faces liquidation even with a smaller number of subjects requesting their funds, those who did not decide to withdraw at the same point in time obtain a payout of zero. This stands in contrast to our paper where proportionate payments in case of a bank run are made for all active depositors, irrespective of their participation in the bank run. Note that despite this circumstance, subjects still have an incentive to withdraw funds from their bank, as their withdrawal may be large enough to secure funds without triggering a bank run yet, given other subjects keep their deposits at that time. This design element in our paper may be regarded as an indirect deposit insurance scheme as previously studied in the literature on individual bank runs. We leave it to future research to further investigate the effect of indirect deposit insurance in banking networks.

Finally, one can discuss the rational behavior of subjects in treatment 1. We argue that agents should not run in this scenario, as they are aware of equal probabilities of good and bad bank projects. Therefore, in case of a succession of negative returns, they should be able to anticipate a trend reversion in the future and keep their deposits unchanged. One may discuss, however, that subjects update their beliefs in each period such that they expect a bank run to be more likely with each new negative return realization, especially since withdrawals do not entail transaction costs. While these different perspectives cannot be aligned easily at this point, we argue that given the construction of the trajectories of bank returns, a run in treatment 1 should not be rational. Subjects are assumed to be able to infer that keeping their deposits is the best long-run strategy. To further address this issue, we suggest running a complementary treatment where beliefs about changes in the probability of a bank run are elicited in each period and compared with project return probabilities.

Future research may focus on extending the current data set and check for original treatment effects. As an additional robustness check, one may include a treatment where withdrawals from banks are organized in a sequential instead of simultaneous manner within each group. In this respect, it would be possible to study the traditional "first come, first served" logic in a dynamic setting. Besides, a proportionate payout of bank assets in case of a bank run could be granted only to subjects involved in the current withdrawal process, for instance. It would furthermore be of interest to extend our modeling of financial linkages toward an incorporation of actual financial interaction between different banks. Such an additional treatment would allow for a robustness check regarding the impact of the presentation of information about economic linkages between banks on withdrawal behavior. Finally, one may consider a follow-up experiment on subjective awareness about network structures in which the determinants of network awareness would be analyzed and an economically desirable degree of awareness determined. In this respect, it would be possible to analyze the severity of the framing effect in this paper.

## 1.7 Appendix

### 1.7.1 Description of questionnaire variables

Trust is measured based on three questions of the German Socio-Economic Panel study (SOEP) youth questionnaire 2015 (Goebel, Kroh, Schröder, and Schupp [2015]), namely one, "One can trust people.", two, "In these times one can't rely on anybody else.", and three, "When dealing with strangers it is better to be cautious." For each question, subjects are asked to assess their perceived level of trust on a scale from 1 to 5, where 1 indicates complete disagreement and 5 indicates complete agreement. Note that the uneven number of items in this and the following questions ensures that truly neutral subjects do not have to choose an option that does not adequately reflect their individual perception. We assume subjects to evaluate items as equally spaced.

Risk aversion and ambiguity are measured based on a question of the SOEP youth questionnaire 2015 (Goebel, Kroh, Schröder, and Schupp [2015]) namely, "Generally speaking, how willing or unwilling are you to take risks?". Subjects are asked to indicate their level of risk aversion on a scale from 1 to 5, where 1 indicates complete risk aversion and 5 indicates complete risk seeking.

Social preferences are elicited through the question, "Generally speaking, how willing or unwilling are you to help others without expecting anything in return?". Again, subjects are asked to indicate their answer on a scale from 1 to 5, with 1 indicating complete unwillingness and 5 indicating complete willingness.

Time preferences are measured by the question "Are you generally an impatient person or someone who always shows great patience?". Again, subjects are asked to indicate their personal view on a scale from 1 to 5, where 1 indicates complete impatience and 5 indicates complete patience.

As to the determination of loss aversion, we use a modification of the approach suggested by Fehr and Goette [2007] which was first implemented by Brune [2011]. Each subject is presented with a sequence of five lotteries and has to decide for each of them whether to accept or to reject it. Each lottery is of type $(100, 0.5; \upsilon, 0.5)$, where $\upsilon \in \{-20, -40, \dots, -100\}$ ECU. Accepting a lottery triggers its realization and the respective amount is added to the earnings from the experiment, which is communicated to the subjects on the computer screen. Rejecting a lottery yields a payoff of 0 ECU. The threshold where each subject decides to move from the lottery to the safe option is then used to construct a measure of loss aversion, $l$, as the percentage of denied lotteries, i.e.,

$$l = 1 - \left( \frac{1}{5} \sum_{k=1}^{5} \mathbf{1}_{L,k} \right),$$

where $\mathbf{1}_{L,k}$ is the indicator function of lotteries accepted in stage $k$. A larger value of $l$ indicates a larger degree of loss aversion.

Beliefs about the probabilities of a bank run are determined by three questions, namely one, "During today's study, were you worried that one or more banks might become bankrupt?", two, "Can you remember in which round you were worried that one or more banks might become bankrupt?", and three, "Did you withdraw money from one of your bank accounts because you were worried that one or more banks might become bankrupt?". While questions 1 and 3 have to be answered with yes or no, the answer to question 2 requires the indication of at least one period between 1 and 20, or can be answered with no.

### 1.7.2 Instructions for all treatments[6]

*Note that the ultimate purpose of this experiment is explicitly not mentioned in any of the following instructions. This is in order to prevent subjects from generating an experimenter demand effect.*

#### General instructions for study participants

Welcome to and thank you for participating in today's study. The study will take approximately 90 minutes.

In this study you can earn money. The amount that you will earn depends on your decisions and on the decisions of the other study participants. All earnings that you will make in this study will be shown to you in Experimental Currency Units (ECU) first, where the following conversion between ECU and CHF applies:

**250 ECU = 1 CHF**
**resp.**
**100 ECU = 0.40 CHF**

At the end of today's study, all earnings that you will have made will be converted into CHF using the above conversion rate and paid to you **in cash**, including a participation fee of 20 CHF.

In order to earn money in this study, it is very important that you read the following instructions carefully.

**During the study, you are not allowed to talk to other participants or use electronic devices.** Non-observance of this rule will lead to exclusion from this study and all payments.

[6]These as well as all following instructions are written along the lines of the instructions used in Brune [2011]. The instructions distributed to the subjects were written in German. Differences between treatments 1, 2 and 3 are indicated by footnotes on page 37.

On the following pages, the course of this study will be described in detail. Please read this description carefully. On a supplementary sheet you will find a number of comprehension questions. Please answer these questions and raise your hand afterward. We will then come to your seat to check your answers and settle any remaining questions. As soon as the answers of all participants have been checked and there are no further questions, please wait until the study begins.

**Description of the study**

Today's study consists of maximally 20 rounds. All study participants will be randomly distributed into groups of six people each. You are a bank customer within one such group of six and you can deposit money at three banks A, B, and C.

You have a cash account as well as one account each at the three banks A, B, and C. At the beginning of round 1, the balance of your cash account is 0 ECU. The balance of each of your accounts at the three banks A, B, and C is 1000 ECU, hence 3000 ECU in total. These balances are to be understood as non-accrual loans. You will receive interests on your bank account balances and these will be your earnings from this study. At the end of this study, your final balance will be calculated which consists of your cash account as well as your bank account balances. The initial balance of 3000 ECU will be subtracted from your final balance. The remaining amount represents your earnings from this study. Should you have incurred losses during the study, that is, should your earnings be negative, then the difference will be subtracted from your participation fee. However, your final earnings from this study will be at least 10 CHF.

In each round, you can change your share of deposits over your three bank accounts as well as over your cash account in each round. At the end of each round, you will receive a fixed **interest of 5%** on your account balance at each bank. You will receive **no interests** on your cash account balance.

**Example on how to calculate your earnings**

- initial balance = 3000 ECU
- final balance = cash account balance + account balance at bank A + account balance at bank B + account balance at bank C
- your earnings = final balance - initial balance

**Example 1:** Assume that your cash account balance is 1000 ECU, that your account balance at bank A is 2500 ECU, that your account balance at bank B is 2000 ECU, and that your account balance at bank C is 1500 ECU at the end of the study. Then, your final balance is equal to

$$1000 \text{ ECU} + 2500 \text{ ECU} + 2000 \text{ ECU} + 1500 \text{ ECU} = 7000 \text{ ECU}.$$

Hence, your earnings amount to

$$7000 \text{ ECU} - 3000 \text{ ECU} = 4000 \text{ ECU}$$

or 16 CHF, respectively. Hence, and together with the participation fee of 20 CHF, you would have earned 36 CHF in total in this example.

**The three banks do not have connections among each other, that is, each bank is independent of the respective other two.**[7] In order to be able to finance the interest of 5% which you will receive for each of your bank account balances in each round, each bank invests all deposits of all study participants in your group into a project at the end of each round. **With a probability of 50% each, this project can either generate a return of 15% or a return of -5%.**[8] The liquid assets of a bank consist of the balances of the depositors in your group as well as the accumulated returns from projects. Depending on past realized returns, at a certain point in time the liquid assets of any bank can thus be larger, smaller, or equal to the sum of balances of the depositors. Net profits of any bank are then equal to the difference between the bank's liquid assets and the study participants' balances in your group. In each round, you will see a diagram showing the level of past net profits for every bank.

Please note that if a bank should have negative net profits in round 20, there are no sufficient liquid funds left to pay out all balances and interests at the end of the study, as every bank finances interest payments through their profits. In this case, every study participant in your group will be credited with a share of the bank's value that is equivalent to the share of his bank balance relative to the sum

[7] In Treatment 3, this sentence reads: **Bank A and bank B are connected. Bank C is independent from the other two banks.**

[8] In Treatments 2 and 3, this sentence reads: **This project can either generate a return of 15% or a return of -5%.**

**Example on how to calculate your earnings**

**Example 2:** Assume that your cash account balance is 500 ECU, that your account balance at bank A is 0 ECU, that your account balance at bank B is 1000 ECU, and that your account balance at bank C is 1000 ECU at the end of the study. Then, your final balance is equal to

$$500 \text{ ECU} + 0 \text{ ECU} + 1000 \text{ ECU} + 1000 \text{ ECU} = 2500 \text{ ECU}.$$

Hence, your earnings amount to

$$2500 \text{ ECU} - 3000 \text{ ECU} = -500 \text{ ECU}$$

or -2 CHF, respectively. Hence, and together with the participation fee of 20 CHF, you would have earned 18 CHF in total in this example.

of all bank balances.

A withdrawal means withdrawing a certain amount from a bank account. If in a given round the sum of all withdrawals requested from the participants in your group is larger than the liquid funds of the bank concerned, then this bank will be liquidated. Each study participant in your group will then be credited a share of the bank's liquidation value on his cash account that is equivalent to the share of his bank balance relative to the sum of all bank balances. In subsequent rounds, no more withdrawals from or deposits in this bank will be possible.

**Example on how a bank is liquidated**

**Example 1:** Assume that you have a balance of 4000 ECU at one of the banks. Also assume that the sum of deposits of all study participants in your group at this bank is equal to 40'000 ECU. Furthermore assume that, at the moment, the bank only has liquid funds (that is, the sum of all current deposits plus returns from previous rounds) in the amount of 30'000 ECU. Should the sum of all requested withdrawals now be equal to, for example, 35'000 ECU, then the bank is not able to pay out the entire requested sum. For each participant, his share of total deposits is now calculated (in your case, this would be equal to 4000 ECU / 40'000 ECU = 10%). Hence, you would receive and be credited on your cash account with 10% of the currently available 30'000 ECU, that is, 3000 ECU.

We will now once more explain to you the procedure of this study using the screens that you will see. In each round, there are two different screens respectively stages. In each stage, you will see the current round in the upper left and the remaining time in the upper right corner. For each stage you have a maximum of 30 seconds. If you haven't entered and confirmed any input yet after this time has expired, no changes will be made to your current account balances and you will be directed to the next stage respectively round. In each round, stage 1 (withdrawals) and stage 2 (deposits) together allow you to change the division of your deposits across your three bank accounts as well as your cash account.

**Stage 1:** You will see information regarding all past net profits of the three banks, regarding current net profits of the three banks as well as regarding your **current balance** in cash and at the three banks A, B, and C. In comparison to the respective previous round, your balance at each bank will have increased by an interest of 5%. You can **withdraw** any amount up to and including the balance of your bank account from one or more than one of your bank accounts and **add** it to your **cash balance**.[9]


Click on **"test"** to see your requested withdrawals. Then click on **"change"** to change your requested withdrawals. If you do not want to change your withdrawals, click on **"confirm"** to go to the next stage. Please note that your entries will only take effect once you have clicked on **"confirm"**.

Your bank balances must **not** be **overdrawn** here, that is, you cannot withdraw more from any given bank account than is available there. Should you overdraw one or more than one of your bank accounts, clicking on **"test"** would trigger the following notification:

In this case, click on **"adjust"** in order to go back to the withdrawal stage and

[9] The screenshots presented here are translations from the German experiment.

> **Notification**
>
> **Attention:** These withdrawals will overdraw your bank balance. Please adjust your withdrawals in such a way that you do not overdraw your bank balance.

adjust your withdrawals accordingly.

Should a bank be liquidated in a given round, then you will see the following notification after confirming your requested withdrawals:

> **Notification**
>
> Bank ... is bankrupt and will be liquidated. Appropriate to the share of your deposits relative to the sum of all deposits, you will be credited the amount of ... ECU on your cash account.

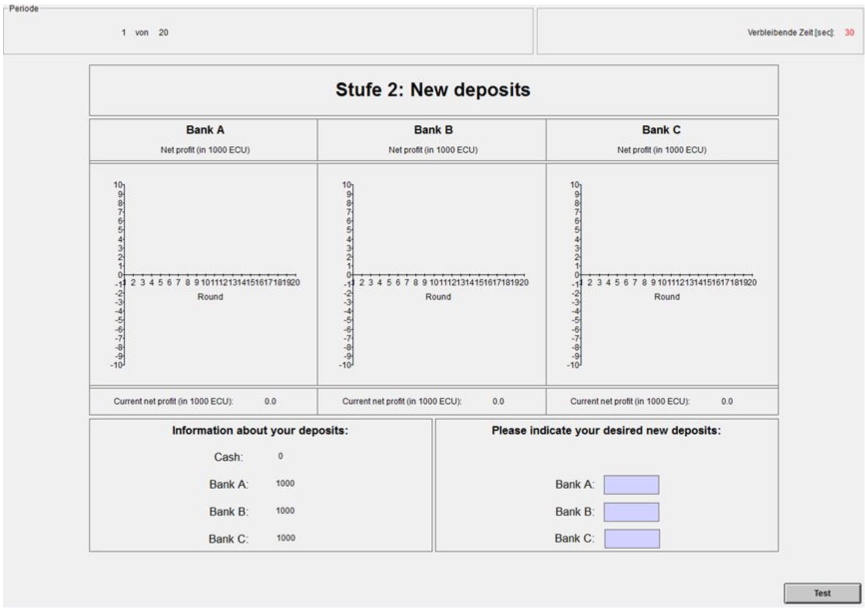
The liquidation of the respective bank is permanent, that is, withdrawals from and deposits into this bank will not be possible any more for the remaining duration of the study. Then click on **"OK"** in order to go to the next screen.

**Stage 2:** You will see information regarding all past net profits of the three banks, regarding current net profits of the three banks as well as regarding your **current balance** in cash and at the three banks A, B, and C. You can **newly deposit** any amount up to and including the balance of your cash account from your cash account at one or more than one of the three banks A, B, and C. The amount that you do not deposit at one or more than one of the three banks A, B, and C you will keep as a cash balance.

Your cash balance must **not** be **overdrawn** here, that is, the sum of your requested deposits at the banks must not be larger than the balance of your cash account. Should you overdraw your cash account, clicking on **"test"** would trigger the following notification:

> **Notification**
>
> **Attention:** These deposits will overdraw your cash balance. The sum of your deposits is larger than your available cash balance. Please adjust your deposits in such a way that you do not overdraw your cash balance.

In this case, click on **"adjust"** in order to go back to the deposit stage and adjust your deposits accordingly. Please note that there will be no more stage 2 in




round 20.

Should all three banks be liquidated before round 20, the following notification would appear:

> **Notification**
>
> All three banks are bankrupt and have been liquidated. The study is thus finished. Please click **"OK"** for as long as this notification appears. Then, please wait a moment until the closing questionnaire appears and until your will be called for your payment.

Finally, at the end of today's study you will be asked to fill in a short questionnaire on the computer screen.

## Comprehension questions

Now, please answer the enclosed comprehension questions regarding the procedure of today's study and raise your hand afterward. We will then come to your seat to check your answers and settle any further questions. As soon as all answers of all participants have been checked and there are no further questions, please wait until the study begins.

### 1.7.3 Control questions and solutions for all treatments

**Question 1:** Which interests do the four different depositing possibilities offer in each round?
**Answer:** bank A: 5%, bank B: 5%, bank C: 5%, cash: 0%.

**Question 2:** Assume that at the end of this study, and after subtracting the initial balance, you will have earned a total of 2500 ECU. How much will you be paid out in total in CHF then (that is, including the participation fee)?
**Answer:** $\frac{2500}{250} + 20 = 30$.

**Question 3:** When will a bank be liquidated?
**Answer:** A bank is liquidated if the sum of all requested withdrawals of the study participants in your group is larger than the liquid funds of the respective bank.

**Question 4:** Assume that a bank goes bankrupt. Die bank had deposits in the amount of 50'000 ECU, but liquid funds of only 40'000 ECU. Your balance at this bank was 10'000 ECU last. Which payout in ECU will you receive then?
**Answer:** $\frac{10'000}{50'000} \cdot 40'000 = 8000$.

**Question 5:** Assume that at the end of this study, and after subtracting the initial balance, you will have made a total loss of 1000 ECU. How much will you be paid out in total in CHF then (that is, including the participation fee)?
**Answer:** $\frac{-1000}{250} = 16$.

**Question 6:** Will this study last for 20 rounds in all cases?
**Answer:** No.

# References


Allen, F., and D. Gale (2000): "Financial contagion," *Journal of Political Economy*, 108(1), 1–33.

Arifovic, J., and J. Jiang (2014): "Do sunspots matter? Evidence from an experimental study of bank runs," Working Paper 2014-12, Bank of Canada.

Arifovic, J., J. Jiang, and Y. Xu (2013): "Experimental evidence of bank runs as pure coordination failures," *Journal of Economic Dynamics & Control*, 37(12), 2446–2465.

Brown, M., S. Trautmann, and R. Vlahu (2016): "Understanding bank-run contagion," *Management Science*, 63(7), 1–11.

Brune, A. (2011): "Endogenous authority: Design and analysis of a laboratory experiment," Master's thesis, Department of Economics, University of Zurich.

Campioni, E., V. Larocca, L. Mirra, and L. Panaccione (2017): "Financial literacy and bank runs: An experimental analysis," Working paper, CEIS Tor Vergata.

Chakravarty, S., M. Fonseca, and T. Kaplan (2014): "An experiment on the causes of bank run contagions," *European Economic Review*, 72, 39–51.

Chari, V., and R. Jagannathan (1988): "Banking panics, information, and rational expectations equilibrium," *Journal of Finance*, 43(3), 749–761.

Diamond, D., and P. Dybvig (1983): "Bank runs, deposit insurance, and liquidity," *Journal of Political Economy*, 91(3), 401–419.

Duffy, J., A. Karadimitropoulou, and M. Parravano (2016): "Financial contagion in the laboratory: Does network structure matter?," Working paper.

Fehr, E., and L. Goette (2007): "Do workers work more if wages are high? Evidence from a randomized field experiment," *American Economic Review*, 97(1), 298–317.

Fischbacher, U. (2007): "Z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, 10(2), 171–178.

Garratt, R., and T. Keister (2009): "Bank runs as coordination failures: An experimental study," *Journal of Economic Behavior & Organization*, 71(2), 300–317.

Goebel, J., M. Kroh, C. Schröder, and J. Schupp (2015): "SOEP 2015 – Erhebungsinstrumente 2015 (Welle 32) des Sozio-oekonomischen Panels: Jugendfragebogen, Altstichproben," SOEP Survey Papers 276, DIW Berlin.

Hens, T. (2008): "Wann sich die Rettung einer Bank lohnt," *Schweizer Bank*, (Dezember), 18.

Kirman, A. (2011): *Complex economics: Individual and collective rationality*, The Graz Schumpeter Lectures. Routledge.

Kiss, H., I. Rodríguez-Lara, and A. Rosa-García (2012): "On the effects of deposit insurance and observability on bank runs: An experimental study," *Journal of Money, Credit, and Banking*, 44(8), 1651–1665.

Kiss, H., I. Rodríguez-Lara, and A. Rosa-García (2014a): "Do social networks prevent or promote bank runs?," *Journal of Economic Behavior & Organization*, 101, 87–99.

Kiss, H., I. Rodríguez-Lara, and A. Rosa-García (2014b): "Do women panic more than men? An experimental study of financial decisions," *Journal of Behavioral and Experimental Economics*, 52, 40–51.

——— (2016): "Think twice before running! Bank runs and cognitive abilities," *Journal of Behavioral and Experimental Economics*, 64, 12–19.

Klos, A., and N. Sträter (2013): "How strongly do players react to changes in payoff parameters in an experimental bank run game?," Working paper, Kiel University and University of Münster.

Madiès, P. (2006): "An experimental exploration of self-fulfilling banking panics: Their occurrence, persistence, and prevention," *Journal of Business*, 79(4), 1831–1866.

Schotter, A., and T. Yorulmazer (2009): "On the dynamics and severity of bank runs: An experimental study," *Journal of Financial Intermediation*, 18(2), 217–241.

# Chapter 2

## Is scapegoating harmful for performance improvement?

Amelie Brune

**Abstract.** Whether in business, politics, or sports, groups replacing their managers, top politicians, or coaches after the group itself performed badly at a given event seems to be a widespread phenomenon. In this paper, we study whether the underlying mechanism for this behavior is scapegoating, and if so, what impact this has on future performance. Analyzing data obtained from a laboratory experiment which we designed in order to study these questions, we find that up to 40% of teams with relative poor performance knowingly assign responsibility for their situation to their coach, when in fact they are aware that they caused this result themselves, and are thus scapegoating by definition. Furthermore, we find that scapegoating teams are less likely to increase their performance over time relative to non-scapegoating teams. Our results provide first experimental evidence of the existence of scapegoating in teams, and aims at raising awareness with respect to the potentially adverse effects on future performance improvement.

**JEL classification:** C91, D71.
**Keywords:** Performance, responsibility, scapegoats.

## 2.1 Introduction

The following headline may be encountered rather frequently as news around the globe:

FOOTBALL COACH DISMISSED AFTER HORRIBLE DEFEAT

Most likely, the reader will have come across such news before. A football team has failed to win an important match and, subsequently, its coach is dismissed. Similar examples can be found in all areas of public life: Companies dismiss managers after detecting flaws in production, parties remove senior politicians after public scandals, and the government dismisses public service employees after bad operations.

Breaking down these examples into their organizational elements, one can observe a common underlying structure: An individual agent or entity guides a group of agents in producing a certain good, where the group's performance is measurable. If the group perceives its performance to be low at a certain trigger event, they decide to replace their guiding agent in expectation of better results in the future. Guiding agents may also choose or be forced to resign, but the resulting scenario is essentially the same. A related scenario is team production of a certain good by means of a given complementary production factor. Using an example from finance, one may consider two heterogenous traders working through a trading platform as facilitating entity. A classic example in economics is that of two workers producing a good in a factory, or even that of two oarsmen covering a certain distance in a rowboat, to mention a more illustrative example. Also from this angle, agents can – individually or as a group – claim the complementary non-human production factor to be the intrinsic cause for observed poor performance.

Among various possible psychological reasons for such behavior, our claim is that the underlying mechanism is scapegoating, i.e., the process of (at least partially) transferring responsibility for poor performance which is due to one's own fault to a third party, and that this is obstructive to improving performance in the future. Despite the far-reaching consequences that evidence for these claims would have for meritocracy, this exact phenomenon has not yet been studied by experimentalists to the best of our knowledge. Empirical support for our claim would offer insights into the motives for replacing managers, top politicians, and coaches, and help to understand why performance does not necessarily have to improve after such observed changes.

From a theoretical point of view in the psychology literature, Rothschild, Landau, Sullivan, and Keefer [2012] propose a dual-motive model of scapegoating, where the first channel they describe is the transfer of responsibility for a negative outcome to a third party in order to reduce feelings of guilt when the outcome was at least partially caused by the individual. The second channel is to increase perceived control over events where the outcome's result was beyond one's zone of influence. The authors also offer empirical support for their theory by conducting three survey-based experimental studies in which they show the existence of scapegoating by manipulating a given negative event, such as general environmental destruction, through either of the theoretical channels described above, test for possible solutions to scapegoating, and investigate possible obstacles in implementing such solutions. They find that, depending on whether or not the agent was responsible for the negative event, scapegoating can be motivated for both channels as described by their model. Within each channel of scapegoating, the opportunity to affirm either individual moral integrity or perceived personal control is shown to erase the previously established matter of scapegoating. However, there is no evidence for cross-effects, e.g., increasing perceived personal control does not cancel scapegoating motivated by personally influenced negative events. However, the authors also find that despite this possibility to erase scapegoating by means of fostering introspection in the respective domain, the mechanism fails to work if an attainable scapegoat is at hand. In our paper, we conduct an economically founded laboratory experiment that ventures along the first channel of scapegoating as suggested by Rothschild, Landau, Sullivan, and Keefer [2012], i.e., scapegoating after events that were influenced by the agent of interest in a negative way. Regarding the design of our study, instead of conducting surveys, we propose an interactive experimental design which directly places each subject into a tangible real-life situation of scapegoating. Moreover, we expand our research focus by asking which impact scapegoating has on future performance.

A commentary most closely related to the idea of our paper is written by Gamson and Scotch [1964], in which the authors argue that scapegoating is the intrinsic mechanism behind managerial changes in major league baseball clubs. They argue that the outcome of a game is essentially determined by the players' talent, while the quality of the team's coach can be considered as relatively constant. The authors then put forward that scapegoating may be used to reduce anxiety within the team, and that it shows the psychological connection between the dismissal of the manager and the hope to improve team performance in the future, while it is actually only the team who can effectively increase output. As such, the authors also conjecture that there should be no observable increase in performance which can merely be attributed to a change in managers. With our paper, we can contribute to this more than 50-year-old question by the design and implementation of a stylized laboratory experiment.

Among other papers which evolve around responsibility shifting, Chang, Solomon, and Westerfield [2016] use an individually performed trading game experiment to show that cognitive dissonance can explain the disposition effect, i.e., the phenomenon that agents tend to realize gains rather than losses, in stocks and the reverse-disposition effect, i.e., the phenomenon that agents tend to realize losses instead of keeping them, in delegated assets such as mutual funds. Cognitive dissonance, which describes a state of negative feelings caused by the realization that one's past decisions were sub-optimal, is shown to drive the blaming of managers for poor fund performance in order to restore individual perception about personal investment skills. The authors also show that the reverse disposition effect is stronger for higher salience of asset delegation. While our paper relates to the one of Chang, Solomon, and Westerfield [2016] in the sense that we also study unjustified blaming, or scapegoating, it differs in that we focus on group rather than individual performance, and generalize the latter through the implementation of a real-effort game. Moreover, we study the existence of scapegoating rather than its motivating factors. In an experiment using dictator games extended by delegation and punishment possibilities by Bartling and Fischbacher [2012], decision rights for potentially unfair allocations are shown to be delegated in order to avoid taking responsibility for the implementation of the allocation at hand. While our paper also aims at investigating the circumstances under which responsibility is transferred, we do not include the possibility to delegate decision rights. Using a survey-based approach, Brown and White [2010] find that a negative relation to one's organization increases the probability to attribute responsibility of a given crisis to this entity, and that scapegoating proves to be the most prevalent of four possible crisis response strategies analyzed. Again, our paper ventures along the same topic of scapegoating. However, we focus on group performance under a guiding agent rather than on the relationship to the latter as the driver of scapegoating. In an empirical panel study by Boeker [1992], the author finds that personal power reduces the likelihood for top managers to get dismissed during times of poor performance of their organization, as this enables them to shift responsibility for poor performance toward the next-lower managerial level, which then suffers from scapegoating. The author also touches upon the link between scapegoating and future performance, where the latter is believed to be improved after changes of managers from the perspective of the organization. While our paper aims at offering an experimental investigation of scapegoating and its effects on future performance, we do not consider the impact of power of the guiding agent. With our paper, we may however address Boeker [1992]'s note that his overall conclusion depends on the industry-specificity of his data. With our generalized, stylized experiment, we hope to be able to contribute to the analysis and understanding of scapegoating and its consequences for future performance. Finally, Brune [2011] analyzes the effect of different types of authorities on the quality of group coordination within the framework of a minimum coordination game. She distinguishes between the existence and type of authority, where the latter differentiates between either democratically elected or randomly imposed authorities, as well as the authority's capability to punish individual group members for defection. The author finds that the existence of an authority who recommends contributions to group members improves coordination, and that this is more important than the type of authority. While there are similar design elements such as the presence of a guiding agent and a group which is required to generate output, the main difference between Brune [2011] and this paper is that it is not the guiding agent who can dismiss group members for performing badly, but the group members who can dismiss their guiding agent if he is believed to be the main responsible party for a given outcome. In this respect, the tables of punishment power are turned in this paper.

In conclusion, while the current literature does address questions of responsibility attribution, there is no experiment on the existence of scapegoating in poorly performing groups of agents and its effect on future performance to the best of our knowledge. With our paper, we aim at filling this gap in the literature. In order to test our conjectures, we use an experimental setting in order to identify causality between poor performance, the replacement of guiding agents, and future performance development. With this approach, we are able to control the environment in which decisions are made, generate data which may be unobservable in the real world, and ensure validity through monetary incentives.

The remainder of this paper is organized as follows. Section 2.2 describes our experimental hypotheses, Section 2.3 presents the experimental design, and Section 2.4 describes our behavioral predictions. Section 2.5 shows and Section 2.6 discusses the results of our experiment, concludes, and offers directions for future research.

## 2.2 Hypotheses

We conjecture that scapegoating is the mechanism behind the replacement of guiding agents after poor team performance, and that, furthermore, if this is true, this is harmful to the improvement of performance in the future. While we propose an array of different real-life examples in Section 2.1 in order to motivate this paper, and while our leading example of a football team and its coach inspires our experimental design, we would like to emphasize that in the following, we abstract from them in constructing our experiment and do not attempt to replicate this situation in a direct manner. In this sense, the use of terms such as "coach" and "team", as well as connotations such as interaction, acquaintance, or the power of decision over coach replacements shall not be interpreted as a direct image of any of the afore-mentioned real-life examples. Our goal is to propose a stylized experiment in order to capture the non-trivial essence of the presented problem, and abstracting from real-life examples allows us to create a flexible experimental framework to study the unjustified attribution of responsibility for poor performance, which may be extended in various directions later on. Nevertheless, applications of our results to real-life situations can be made at the end of this paper for illustrative purposes.

Our conjectures stem from the observation of multiple real-life situations where changes on a managerial level often seem to follow after poor performance of respective subordinates. While we postulate our hypotheses about the underlying mechanism of such situations in Table 2.1, we provide a theoretical discussion including a theoretical benchmark against which to interpret our results in Section 2.4. This allows for an elaboration within the terminology of the experimental setting as described in the following section.

| | |
|---|---|
| **H1:** | Members of teams with poor relative performance which can be attributed to the team's own fault knowingly shift responsibility for this poor performance onto a third party. |
| **H2:** | If hypothesis H1 is true, then the difference between the improvement in team performance of non-scapegoating groups and the improvement in team performance of scapegoating groups is positive, i.e., scapegoating proves to be an obstacle in improving team performance. |

Table 2.1: Experimental hypotheses.

In the following, we describe an experimental design which allows us to test these hypotheses.

## 2.3 Experimental design

In order to test our hypotheses as presented in Section 2.2, we design a laboratory experiment which can be used to identify the existence of scapegoating as well as to study the development of future performance depending on how poor performance has been dealt with before. The experiment is divided into two parts: A pre-session and the main sessions. The former serves as a selection mechanism for some subjects to perform the role of a coach in the main sessions, where they are assigned to teams of two subjects each. Together, they play against another unity of team and coach by means of a real-effort game.

The pre-session is designed in order to select subjects for the roles of coach and replacement coach as required in the main sessions. Subjects individually perform a real-effort game, which consists of adding as many of ten sets of three double-digit numbers[1] as possible within one minute and where each correctly solved problem is given one point. This task is repeated over ten periods. Based on the individual average number of points obtained by the end of period 10, the top 20% of subjects

[1] Numbers are randomly drawn from a uniform distribution of integers between 10 and 99.

are selected, which are offered a bonus of 100 ECU[2] in order to create an additional incentive to exert effort in the game described. To operationalize the fraction of top performing subjects in a convenient way, the pre-session is conducted with 10 subjects, thus leaving the best two subjects with the roles of coach and replacement coach for the main experimental session. Whether the best or second best subject is given either role is random.[3] The result of one of the ten periods is randomly selected for payment, and this number also represents the coach's performance $x$ in the main sessions, where the coach is represented by his performance data only and, thus, his performance is exogenous.

The experiment in the main sessions consists of two periods and is designed as a between-subjects experiment.

In period 1, subjects are randomly assigned to teams of two, which is the smallest possible representation of a group or team and thus allows us to maximize the number of observations with respect to team choices. The coach who is determined in the pre-session is now assigned to each team, and each subject is endowed with 200 ECU. Half of all teams is labeled team A and the other half is labeled team B. One team A and team B each and their respective coaches are then randomly matched to play against each other. The task for each team and its coach is to maximize output through the same real-effort game as used in the pre-session, however with only one period thereof.[4] The task is performed individually by each subject at the same time. Each correctly solved problem is given one point. Team $j$'s total number of points, $g^j$, is then determined by the sum of points achieved by the two team members times the number of points obtained by the coach. The underlying production function can thus be described by

$$g^j = x^j \left( \sum_{i=1}^{2} y_i^j \right), \tag{2.1}$$

where for each team $j$, $x^j$ denotes the coach's performance and $y_i^j$ denotes the performance of subject $i$. Hence, the performance of the coach has a leveraging effect on team performance as his contribution enters the production function in a multiplicative way. Note that both $x^j$ and $y_i^j$ are determined by ability, which

[2] Experimental currency units.

[3] Recall that we abstract from the real-life examples stated in Section 2.1. While there are many possible interpretations and applications of our experiment to real-life situations, the experimental design is bound to abstract from both particular situations and their organizational details. As such, the selection mechanism for coaches is not intended to follow any particular real-life procedure, but rather to provide a basic legitimization for certain subjects to perform the role of the coach in the main sessions.

[4] A trial period consisting only of the real-effort game precedes the actual experiment. This is in order to allow subjects to familiarize themselves with the task, and to avoid potential learning effects in period 2. We understand learning as an increase in the number of correctly solved problems due to familiarity with the environment in which this task is performed. This is different from an increase due to higher effort.

is fixed in the short run, and effort, which is flexible in the short run. Hence, we assume it to be possible to increase individual output through increased effort between periods 1 and 2. Within each game, the team with the larger value of $g$ wins. It is communicated to each subject whether his team won or lost, as well as by how much they did so. Where a game ended in a draw, both teams involved are considered as winners. Each subject earns $\Delta g$ in ECU, which is the difference between $g^j$, generated by his own team $j$, and $g^k$, generated by the opposing team $k$. Hence, each subject in the winning team earns $\Delta g > 0$ which is added to his endowment. Likewise, each subject in the losing team loses the same amount (since $\Delta g < 0$ here) which is subtracted from his endowment. This is communicated to each subject, as well. The purpose of the endowment is to create a reference point from which any loss is perceived as more harmful than a gain of the same magnitude (see Kahneman and Tversky [1979]), such that subjects in the losing team experience strong disutility from losing. Subjects are therefore expected to exercise maximum effort in the real-effort game in order to avoid losing throughout the experiment, or at least to try to increase effort after losing in period 1. Due to the multiplicative form of equation 2.1, even small differences in either $y_1$ or $y_2$ can have a large impact on $g$. If $\Delta g < 0$, poor performance of one team member thus has the potential to significantly increase the distance to the other team, $|\Delta g|$, and thereby the magnitude of deduction from current endowments. In other words, this captures the degree of influence that the poorly performing team member has on the outcome of the game, which can be translated to the share of responsibility he effectively bears.

Within each team, the poorly performing subject receives complete information about the composition of $g^j$, i.e., the coach's contribution $x^j$, the other team member's contribution $y^j_{\text{high}}$, as well as his personal contribution $y^j_{\text{low}}$. The well-performing subject only receives information about his personal contribution $y^j_{\text{high}}$. The instructions are written in such a way that the well-performing subject within each team cannot infer from the information presented to him that he must have performed better than the other team member. The respective passage only mentions that information about the number of points of some study participants involved is presented, but it does not clarify which information exactly is provided, nor does it say what the access to information will depend on (see Appendix 2.7.3). Hence, the non-existence of information about the other team member's or the coach's contribution cannot lead the well-performing subject to the conclusion that he holds the position of the better performing team member. The poorly performing subject knows that the better performing subject does not know about his full information set and is thus unaware of the actual cause underlying realized team output. Also, subjects cannot update their beliefs about the information structure over time, as information is only presented once in period 1.

Afterward, each subject is asked to provide his judgement about whom he deems responsible for the outcome of the game. He can choose between the coach, himself, the other team member, or the team (i.e., himself and the other team member as a group). Subjects make their responses simultaneously. Then, each subject's choice is presented to the other team member within his team. If the team lost and both team members make the coach responsible for this outcome, then the latter is replaced by the replacement coach for period 2. If the team won, the coach cannot be replaced. Other than this, winning and losing teams naturally also differ with respect to total team output, $Y$, which manoeuvres them in the respective winning or losing position. If it is true that poorly performing subjects are inclined to scapegoat when they have to reveal the probable underlying cause for bad team output, then the poorly performing subject should choose to tell the other subject that the coach was responsible for this outcome. If he does so, he does it in full knowledge that it is not true. Thus, this mechanism represents the very definition of scapegoating, i.e., to knowingly attribute false responsibility for individual failure to a third party in order to regain personal self-esteem within groups, and thereby allows for an isolation of scapegoating. To anticipate our empirical analysis, the action of indicating the coach is our indicator for the existence of scapegoating.

At the end of period 1, each subject is privately shown his updated endowment.

Period 2 is a repetition of period 1 with the exception that no further responsibility assignment through voting takes place, as the experiment ends after this period. Each team plays against the same team it was paired with in period 1. Also, if both members of the losing team in the previous period attributed responsibility to their coach, the replacement coach takes the position of his predecessor, about which the respective teams are informed at the beginning. At the end, total earnings are calculated and privately shown. The final payoff for each subject is composed of a show-up fee of 15 CHF, the endowment of 200 ECU, as well as the sum of earnings made from each period, which can be positive or negative. Total profits, i.e., the sum of endowment and per-period profits, is bounded by 0 from below, which is in order to protect subjects from extremely high losses which may occur if paired with a very well-performing group while their own performance is extremely low. At the same time, these parameters settle the distribution of $g$ over its most probable range for subjects with average problem solving abilities. The exchange rate between ECU and CHF is 20 : 1.

To conclude, note that the motivation behind conducting the pre-session is to provide the legitimization for subjects to be chosen to perform the role of coach in the main sessions. The coach has to bear a certain level of proficiency in performing the real-effort game, and thus the capability of genuinely improving team output as described earlier. Moreover, conducting the pre-session enables us to increase the number of observations in the main session, as we are able to economize on subjects who otherwise would have to act as potential replacement coaches in period 2. Also, it allows for comparability across games, as we use the same subjects to perform the roles of coach and replacement coach in each team, respectively.

At the end of the study, subjects are asked to fill in a short questionnaire to elicit socio-economic parameters, attitudes toward trust, risk, and losses, social preferences[5], their personally perceived effort made in solving the calculation problems presented, as well as their motivation underlying their choice of the entity they believed to bear main responsibility for the game's outcome in period 1[6].

## 2.4 Behavioral predictions

We can interpret our experimental design in the light of a repeated two-stage simultaneous-move game of complete information according to the classification of Gibbons [1992]. The experimental situation shall be classified as being of complete rather than incomplete information as the well-performing subject in each group is unaware that the other team member has access to more information, as this is not discussed in the instructions. Moreover, both subjects know the respective other player's payoff, as both obtain $\Delta g$. Within this framework, we can analyze in which cases it is rational to replace the coach, and in which cases it is irrational. The latter is the situation where scapegoating, i.e., the wilful and unjustified attribution of responsibility for poor group performance to the coach can arise.

The payoff structure for each team member is determined by the value of the production function $g$ of his own team, $j$, and by that of the opposing team, $k$, (see equation 2.1), or, in short, by his team's distance to the opposing team, $\Delta g$. Each subject's best strategy is to maximize individual effort in the real-effort game given their personal ability and given the circumstances at the beginning of the experiment. Ex ante, the probability to win the game in period 1 is $\text{Prob(win)} = \frac{1}{2}$, as subjects are randomly distributed over teams such that individual ability and the propensity to exert effort are evenly distributed, and coaches are equally good across groups, which is both common knowledge. Applying Bayesian updating, the probability of a team to win in period 2 given that it has already won in period 1 is $\text{Prob(win} \mid \text{win)} > \frac{1}{2}$, given the persisting effect of individual ability. Likewise, the probability of a team to lose in period 2 given that is has already lost in period 1 is $\text{Prob(lose} \mid \text{lose)} > \frac{1}{2}$. Ex ante, poorly performing teams in period 1 are more

[5]These are the same standardized questions as used and explained in Chapter 1 (see Appendix 1.7.1). The lotteries presented in the lottery choice task which are used to elicit the degree of loss aversion were scaled down by a factor 10 in order to match the scale of resulting lottery earnings with previous earnings from the scapegoating game using the same exchange rate between ECU and CHF.

[6]With the exception of this last question, the questionnaires for the pre-session and the main sessions are the same.

likely to stay in a losing position in period 2 given fixed ability in the short run, unless more effort is exerted in order to compensate for lower ability. While any number $\tilde{y}_i$ might be the individually perceived maximum possible effort in period 1, after observing his team losing the first game, this figure may shift to a higher level $\hat{y}_i$ (with $\hat{y}_i > \tilde{y}_i$), such that effort is still able to increase. Returning to the ex ante situation, the probability to switch status over time is smaller than $\frac{1}{2}$, i.e.,

$$\text{Prob(win} \mid \text{lose)} < \frac{1}{2} \qquad \text{and} \qquad \text{Prob(lose} \mid \text{win)} < \frac{1}{2}.$$

While an increase in perceived maximum possible effort from $\tilde{y}_i$ to $\hat{y}_i$ has the potential to affect these probabilities, changing the coach does not, as his performance is fixed and on a top level. Hence, expected payoffs cannot be increased (or decreased) by changing the coach. Therefore, the rational benchmark is that

$$\text{Prob(change coach)} = 0.$$

This can be considered as the Nash equilibrium. Any contrary observations in the experiment, i.e., any replacements of coaches should thus be considered as irrational. Note that teams who scapegoat would believe that Prob(win | lose) $> \frac{1}{2}$ in case of a new coach, whereas it is effectively still smaller than $\frac{1}{2}$ or could only pass the threshold of $\frac{1}{2}$ through higher effort, but not through a new coach.

A subject could rationally dismiss his coach if he finds him to be less skilled than expected. The information set that the poorly performing subject has about the coach is that his performance can have a theoretical range between 0 and 10, and that he was drawn from among the top 20% of subjects in the pre-session. Assuming a normal distribution of correctly solved calculation problems with $\mu = 5$ and $\sigma = 1.5$, for instance, the top 20% would encompass any number between 6.4 and 10. Therefore, it would be rational to replace the coach for reasons of unrealized quality expectations if $x \leq 6$ is revealed at the end of period 1. Regarding the well-performing subject within the losing team, while he has no access to information about the actual performance of the coach, he can still perform the same reasoning about the expected quality using $\mathcal{N}(\mu, \sigma)$ and his personal (good) performance as a benchmark, intrinsically knowing whether he solved the task well or badly.

Based on these considerations, we make the following conjectures for our experiment: One, the percentage of losing teams who (unanimously) assign responsibility for the outcome of the game to the coach is significantly larger than zero, i.e., scapegoating can be observed. In particular, the fraction of poorly performing subjects within losing teams which assigns responsibility to the coach is significantly larger than that of well-performing subjects in the same teams. And two, using a difference-in-difference approach, the improvement in performance over time in scapegoating losing teams is significantly smaller than in non-scapegoating losing teams. Both conjectures are motivated by real-life observations as elaborated on in Section 2.1 and can be checked against the theoretical benchmark as discussed in this section.

## 2.5 Results

The experiment was programmed in z-Tree (Fischbacher [2007])[7] and conducted in March and April 2017. We recruited subjects using the University Registration Center for Study Participants of the University of Zurich, where we required participants to have taken part in at least one previous study and to understand German well. This is because subjects who participated in an experiment before know the laboratory environment and the experimental procedures, which simplifies the realization of the study, and because the experiment was conducted in German. There were 10 subjects in the pre-session and a total of 116 subjects[8] distributed over five main sessions, all of which were conducted in the Laboratory for Experimental and Behavioral Economics at the Department of Economics at the University of Zurich. Each session lasted 45 minutes on average.

Detailed written instructions were distributed to each subject before the start of the experiment for them to read carefully (see Appendix 2.7.1 and 2.7.3). A set of control questions had to be answered correctly in order to check their understanding of the procedures of the experiment. Answers were checked by the experimenter, and false ones indicated and explained individually (see Appendix 2.7.2 and 2.7.4).

Subjects earned 30.80 CHF in the pre-session and 25 CHF on average[9] in the main sessions. In the pre-session, the average age was 25.3. 30% were male and 80% were of Swiss origin. With 40% of subjects, the majority reported monthly disposable income within the bracket of 1000 CHF and 1500 CHF. In the main sessions, the average age of study participants was 23.9. 40% were male and 77% were of Swiss origin. 34% reported monthly disposable income of up to 500 CHF with income skewed to the right overall. The distribution of the reported fields of studies of all participants is shown in Table 2.2. We find that both reported fields of studies and income brackets show no abnormalities for the typical range of participants recruited from our data bank.

[7] The code is available upon request.

[8] This number includes 16 subjects from a pilot session, where the show-up fee was 10 CHF and $x = 8$ (as compared to 15 CHF and $x = 9$) as initial parameter values. Changes in the fixed show-up fee should not alter subjects' behavior during the experiment, and both values of $x$ are within a probable range for a performance to be among the top 20% of the population (for instance, assuming a normal distribution with $\mu = 5$ and $\sigma = 1.5$, any number between 6.4 and 10 lies within the upper 20% of the distribution. Hence, as these differences are negligible, we decided to integrate the data obtained from the pilot session into the data analysis in order to increase the number of observations.

[9] This figure takes into account an additional 5 CHF for the pilot session.

| Field of studies | Pre-session | Main sessions |
|---|---|---|
| Arts, humanities, and political sciences | | 14% |
| Computer sciences | | 2% |
| Engineering and architecture | | 12% |
| Law | | 8% |
| Mechanical and electrical engineering | | 5% |
| Mathematics | | 3% |
| Medical sciences | | 9% |
| Natural sciences | 60% | 24% |
| Economics | 10% | 9% |
| Other/not specified | 30% | 14% |

Table 2.2: Overview of participants' reported fields of studies.
*Displayed is the percentage of subjects reporting a given field of studies in the pre-session and in the main sessions.*

In the pre-session, the random selection of one period outcome of each of the two players in the top 20% yielded $x = 9$ each, which are the data that were used for the contributions of coach and replacement coach in the main experimental sessions. As such, the team is solely responsible for the outcome of the game as determined by the production function (see page 51).

We study losing teams against the theoretical benchmark as discussed in Section 2.4. Winning teams are studied as an additional control and in order to gain further information about these subjects' behavior.

We collected observations for 29 games involving 58 teams, out of which 25 teams lost and 33 teams won the game or ended undecided. In the following, winning teams denote teams that won their game in period 1. This label stays fixed for period 2 in order to allow for a comparison of these groups' change in performance over time. The analogous holds true for losing teams. Similarly to the definition of ties, subjects within any team who obtained the same number of points in any period were both labeled a good player or well-performing subject.

| | Winning teams | Losing teams | $\Delta\bar{Y}$ |
|---|---|---|---|
| Period 1 | 6.41 | 4.96 | 1.45 |
| Period 2 | 6.79 | 5.70 | 1.09 |
| $\Delta_t$ | 0.38 | 0.74 | |

Table 2.3: Average contributions of winning and losing teams.
*Displayed is the average number of correctly solved calculation problems in the real-effort game within each team. $\Delta\bar{Y}$ denotes the difference in average contributions between winning and losing teams in each period. $\Delta_t$ denotes the differences in average team performance over time for winning and losing teams, respectively.*

Table 2.3 reports average contributions as well as differences in contributions for winning and losing teams over time. Winning teams make larger average contributions than losing teams in both periods ($\Delta\bar{Y} > 0$). In each group, performance slightly improves over time ($\Delta_t > 0$), although not in a statistically significant way: For both groups, the null ($H_0$: $\Delta_t = 0$) cannot be rejected in favor of the right-sided alternative (p-value equal to 0.0336 for winning teams and 0.0029 for losing teams). The null of equal performance improvements across groups ($H_0$: $\Delta_{t(\text{winning teams})} - \Delta_{t(\text{losing teams})} = 0$) cannot be rejected.

| | Winning teams | | Losing teams | |
|---|---|---|---|---|
| | Bad player | Good player | Bad player | Good player |
| Oneself | 29% | 17% | 77% | 32% |
| Other player | 13% | 3% | | 14% |
| Team | 29% | 66% | 14% | 18% |
| Coach | 29% | 14% | 9% | 36% |

Table 2.4: Distribution of responsibility attribution by type of player.
*Displayed is the percentage of subjects within each of four different categories of teams who indicated themselves, the respective other player in their team, their team as a whole, or their coach to be the main responsible party for the game's outcome.*

Table 2.4 summarizes the fractions of poorly and well-performing subjects in winning and losing teams which attributed responsibility for their game's outcome to different parties. Interestingly, with 77% the vast majority of all poorly performing subjects among those teams who lost stated that they see the responsibility for their situation in themselves and thus their personal performance. 9% were truly scapegoating in our original definition of the term: These subjects hold the coach responsible for losing the game, even though they know that he contributes 9 units, that the other team member contributed more units than the subject itself, and thus that the game's unfavorable outcome has to be attributed to their own performance, as the majority correctly acknowledges. However, 36% of the better performing subjects among losing teams deem the coach responsible. They do not know the contributions made by the coach or the other team member, yet they decide to blame the coach. This may be interpreted as a milder form of scapegoating: Under incomplete information, blame is falsely attributed to the coach, however the respective subjects could have assumed that the coach's performance was rather good, given that he was drawn from the top 20% of all subjects in the pre-session. While he is the better performing subject in his team, they still lost the game. At this point, a careful interpretation of this finding is the second channel of scapegoating as postulated by Rothschild, Landau, Sullivan, and Keefer [2012], i.e., the aim to regain perceived control through scapegoating after negative events that were not directly possible to influence by the agent of interest. Among winning teams, 66% of well-performing subjects indicated that they deem the team as a whole responsible for the favorable outcome of the game, whereas poorly performing subjects show a more even distribution of responsibility assignments. In fact, in this group 29% each assign responsibility to either the coach, oneself, or the team as a whole.[10] An analysis of the motivation which subjects stated to lie behind their respective choices supports these findings. Good players in winning teams credit their win to the team, as they place more attention to the value and work ethic of the team. Bad players in losing teams credit their loss to themselves, as they not only have a very realistic perception of their own calculation abilities, but also acknowledge their realized weak performance within the team.

A closer look at the data of those subjects who were among bad players in losing teams and who attributed responsibility to the coach reveals that, in fact, they belonged to the bottom 36% of the distribution of points obtained by all poorly performing subjects in losing teams. This is hence an indication for poorly performing individuals to indeed put blame onto a third party. In particular, it has to be noted that their distance in performance relative to their teammate was between 4 and 7 points and relative to their coach between 6 and 8 points.

To obtain a tangible measure of scapegoating, we define $s$ as the number of coach dismissals, $d$, relative to the number of groups, $n$, for all losing teams:

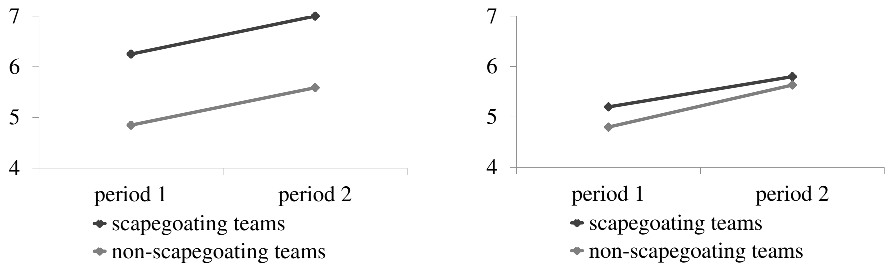
$$s = \frac{d}{n}.$$

Since $s \in [0, 1]$ by construction, we are thus able to generate a probability measure for scapegoating. We can furthermore distinguish between a strong (first-order) and a weak (second-order) version of this measurement, where we define $s_1$ for all losing teams where both members and $s_2$ where at least one member voted in favor of dismissing their coach. We find that $s_1 = 8\%$ and $s_2 = 40\%$ in support of our hypothesis H1, i.e., that members of teams with poor relative performance which can be attributed to the team's own fault knowingly shift responsibility for this poor performance onto a third party.[11] However, we also find that the fraction of better performing subjects in losing teams who scapegoat is actually larger than the fraction of worse performing subjects. Teams who scapegoated in the strong definition of the term stated their wish to be allocated a new coach, or suspected their coach of poor performance. Scapegoating teams in the weak definition also stated that they did not wish to blame their teammate for poor results, and would rather choose to blame their coach instead.

Figure 2.1 shows the development of team performance among scapegoating and

[10] Note that an individual's choice of the coach as the main responsible party for the outcome of the game bears intrinsically different connotations in winning and losing teams. If a subject in a losing team indicates the coach as the main responsible party, it is an manifestation of scapegoating as elaborated on in Section 2.3. However, if a subject in a winning team makes the same choice, it has a rather positive connotation. As such, the implications of any option chosen in winning and losing teams has to be assessed accordingly.

[11] The magnitude of first-order scapegoating, $s_1 = 8\%$, is not to be confused with the fraction of poorly performing subjects in losing teams attributing responsibility to the coach as shown in Table 2.4. $s_1$ captures the fraction of losing teams where both subjects voted for the coach.




(a) $s_1$ (b) $s_2$

Figure 2.1: Difference in improvement of team performance between scapegoating and non-scapegoating losing teams.

*Displayed is the development of average team performance over time in scapegoating and non-scapegoating teams, with (a) and (b) displaying the respective values for the strong and weak measure of scapegoating,* $s_1$ *and* $s_2$.

non-scapegoating losing teams for our two different measurements of scapegoating, $s_1$ and $s_2$. While all losing teams are able to improve their performance over time on average (as shown in Table 2.3 for a general overview), it can also be seen that the difference in improvement between scapegoating and non-scapegoating teams, once differentiating losing teams with respect to these two sub-groups, is rather small for both definitions. First-order scapegoating yields a difference-in-difference measure of $\Delta_1 = -0.01$, where $\Delta_1$ is defined as

$$\Delta_1 = \left(\bar{Y}^{\text{nsg}}_{t=2} - \bar{Y}^{\text{nsg}}_{t=1}\right) - \left(\bar{Y}^{\text{sg}}_{t=2} - \bar{Y}^{\text{sg}}_{t=1}\right).$$

Again, $\bar{Y}$ denotes average team performance, and nsg/sg abbreviates non-scapegoating and scapegoating teams, respectively. Our hypothesis H2 conjectured that $\Delta_1 > 0$, i.e., if hypothesis H1 is true, then the difference between the improvement in team performance of non-scapegoating groups and the improvement in team performance of scapegoating groups is positive, i.e., scapegoating proves to be an obstacle in improving team performance. While our estimate for $\Delta_1$ shows the opposite sign, we do find a positive difference for second-order scapegoating teams, where $\Delta_2 = 0.23$. However, we find that this value is not significantly larger than zero. For this to be the case, we would need to find a value of at least 1.3 at a significance level of 5% in our data. On another note, it is interesting to observe that in both periods, scapegoating teams perform significantly better on average in the strong definition of the scapegoating measure: The null of $H_0$: $\bar{Y}^{\text{sg}}_t - \bar{Y}^{\text{nsg}}_t = 0$ can be rejected against the right-sided alternative at the 10%-level of significance (with p-values equal to 0.0715 for $t = 1$ and 0.0870 for $t = 2$).

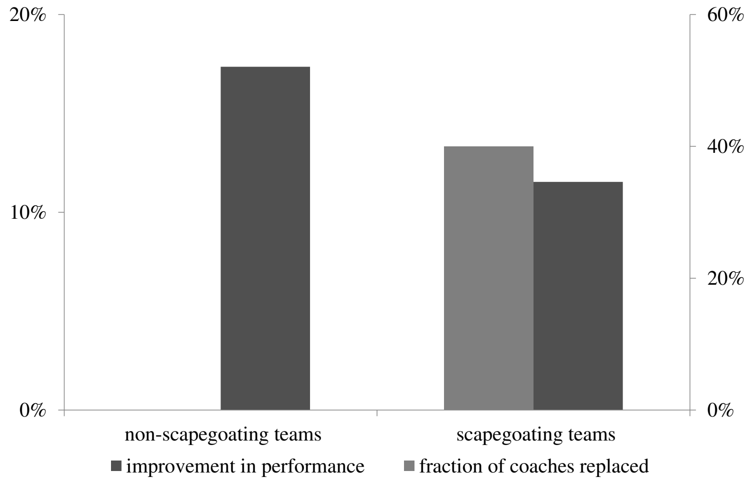



Figure 2.2: Summary of findings for second-order scapegoating.
*Displayed is the improvement in performance over time (measured on the left-hand axis) and the fraction of coaches blamed for losing the game (measured on the right-hand axis) for non-scapegoating and scapegoating teams, respectively).*

Figure 2.2 summarizes our findings for hypotheses H1 and H2 for second-order scapegoating. We find that, one, the fraction of losing teams who blame the coach for losing the game is significantly larger than zero, and that, two, their increase in performance over time tends to be smaller than that of non-scapegoating teams.

Finally, we perform some complementary analyses to obtain a better insight into the relation between the propensity to scapegoat on the one hand and trust, risk and loss aversion and social preferences on the other, as well as some robustness checks.

To this end, we define four sub-groups according to their vote in period 1 (oneself/other player/team/coach). For each sub-group, we then define an index to further distinguish between the four different categories of performers (good/bad players in winning/losing teams). We then compute correlations between each such index and the degree of trust[12], risk and loss aversion, and social preferences, respectively, as reported in the questionnaire.

We find that poorly performing subjects who attribute responsibility for losing

[12] We compute correlations for all three measures of trust employed in the questionnaire (as explained in Appendix 1.7.1). The scale of the second and third measures of trust are mirrored in order to make them comparable with the first measure, i.e., so that a larger value of the respective measure indicates a larger level of trust toward other individuals. Using more than one measure of trust serves as a robustness check.

the game to the coach have a considerably lower level of trust than other subjects (correlations of up to $r = -0.3239$). The same is true for poorly performing subjects in winning teams, although to a lesser extent (correlations of up to $r = -0.0869$). Well-performing subjects in winning teams who believe the team as a whole to be responsible for winning prove to have more trust in others (correlations of up to $r = 0.1360$). Good players in both losing and winning teams who vote for their teammate have more trust toward others (correlations of $r = 0.1667$ and $r = 0.1295$, respectively). Scapegoating subjects also tend to be more risk averse. The correlation between voting for the coach and showing poor performance in losing teams is $r = -0.2653$, which proves to be the largest correlation with respect to risk observed across all sub-groups. They also tend to exhibit a lower degree of social preferences, as documented by the negative correlation of $r = -0.2209$, i.e., they have lower preference for helping others without expecting anything in return. Subjects who scapegoat in the strict definition of the term tend to be more loss averse than others ($r = 0.1095$).

With one exception, all subjects indicated that they believe to have made a strong effort in solving the calculation exercises (a value of at least 4.3 on a scale from 1 to 5), which is an indicator for well working financial incentives in our experiment.

To conclude, comparing the performance data of periods 1 and 2 with those of the trial period, we find that the difference in the number of points obtained from the real-effort game, $y$, between the trial and the first period is always larger than the difference between the first and the second period. This indicates the presence of a learning effect from the trial period to period 1, which supports our decision to include a trial period in order to reduce the possibility to confound learning effects with increased effort from period 1 to period 2. We therefore assume that the remaining observed difference between performance in period 1 and period 2 can be traced back not to learning, but to increased effort, as discussed in Section 2.4 on page 55.

Finally, in the spirit of a small cross-sectional study, we can make use of the circumstance that the coach's contribution was $x = 8$ in the pilot study and $x = 9$ in the main sessions in order to check for rationality of coach dismissals. We find that the fraction of poorly performing subjects wanting to dismiss the coach under $x = 8$ is more than six times larger than under $x = 9$. While ultimate conclusions from this comparison have to be made with care due to the small number of observations from the pilot study, it nevertheless gives a direction toward the behavior of poorly performing subjects in losing teams: Coaches tend to be made responsible more frequently if their performance is lower. This supports the rational benchmark as described in Section 2.4, i.e., that coaches can be rationally dismissed, but only if $x \leq 6$. As $x = 8$ in our data, the fraction of coaches being made responsible can be attributed to irrational scapegoating.

## 2.6 Discussion and concluding remarks

In this paper, we study whether scapegoating is the mechanism behind the replacement of coaches after poor team performance, and what effect this has on future group performance. To address these questions, we design a laboratory experiment where two teams consisting of two members each and their respective coach play against each other in a real-effort game. Among losing teams, we find that up to 40% of teams blame their coach for the game's unfavorable outcome, while they are in fact aware that he cannot be held responsible for this. Moreover, we find that the increase in future performance of these scapegoating teams is smaller as compared to non-scapegoating teams. However, this effect is not statistically significant at conventional levels. Our results thus provide evidence for the existence of scapegoating in poorly performing teams and its potential negative consequences for future performance development.

For illustrative purposes, we can apply our experimental design to our leading example of the situation of a football team and its coach by considering the position and strategic behavior of the coach as well as the decisions made by each individual within the team and the team as a whole. In our design, the coach has a leveraging effect on team output and thus on final team payoff. The incentive of each subject should be to maximize individual effort, as this increases the probability of winning the game and thus final payoffs. This could mirror the real-life situation where the coach of a football team can provide strategic advice tailored to achieve optimal results, but has no direct influence over the players' performance in it. Also, the real-effort game used as performance task could reflect the situation during a football match: Not only do team members have to maximize individual effort, but also, the effort exerted by the weakest player is most likely to represent the binding constraint of the overall result that the team can achieve. Other applications can be made with appropriate care.

Having shown that there is mild evidence for scapegoating (see hypothesis H1), one can now venture forth into studying possible reasons which underlie this phenomenon. One possible conjecture is that scapegoating may be triggered through cognitive dissonance, a mental state describing the discrepancy between perceived individual ability to perform well on a given task and the realization that the corresponding outcome does not measure up to one's expectations. Restoring self-esteem is then accomplished by blaming a third party in order to feel better about oneself (see Chang, Solomon, and Westerfield [2016]). The attempt to reduce cognitive dissonance after bad performance could be one reason for losing teams to replace their coaches. The first channel of scapegoating described in the model by Rothschild, Landau, Sullivan, and Keefer [2012] may provide a theoretical foundation for such a project. Hence, one may expand the experimental design by adding an appropriate treatment to investigate this aspect.

While our data indicates that there might be evidence for hypothesis H2 as we find effects that go in the postulated direction, we do not find that these are significantly different from zero just yet. However, this might change if more data are available. Hence, we suggest to re-run the experiment on a larger scale in order to obtain more observations and thereby improve statistical power for the negative effect of scapegoating on future team performance. Along these lines, one may also expand the current analysis toward an econometric assessment of individual responsibility attribution choices of which an additional advantage is to control for the team members' ability in solving the calculation problems presented. This procedure would allow for a robustness check to establish that increased team contributions actually stem from increased effort, as postulated in Section 2.4, as well as for studying whether individuals with higher ability are able to improve faster over time than individuals with lower ability.

Some scholars may argue that the problem of convex payoffs could arise from the theoretical existence of ambiguity regarding the coaches' effective quality. Their line of argument would evolve around the statement that coaches are drawn from the top 20% of subjects in the pre-session, which could give rise to gambling, emphasizing the relevance of individual beliefs about the distribution of coach performance. Being unable to alter personal ability in performing the real-effort game in the short run, the only option for the poorly performing subject in each team to increase the probability of winning the subsequent game is to vote for a replacement of the coach, thereby speculating to obtain a better coach in period 2. Hence, the dismissal of any coach is said to be rational if there exists a positive probability of the replacement coach to perform better than the previous coach. We can address this conflicting point of view in three ways. One, our results show that 77% of poorly performing subjects in losing teams correctly believe themselves to be the main responsible party for their game's outcome. This may be regarded as an indication toward non-gambling behavior in the majority of the subject population. Two, the dismissal of coaches would only be rational for $x \leq 6$, which was not realized among the top 20% in the pre-session. However, an additional control treatment where coaches make contributions of $x \leq 6$ could provide more insights about the salience of this argument. And three, a control treatment with fixed-ability coaches and replacement coaches could be conducted in order to check whether the observed scapegoating pattern changes if subjects know that the replacement coach cannot have better (or worse) contribution power than their initial coach. A simple implementation of this idea is to re-run the experiment while making $x = 9$

common knowledge.

An alternative way to test for rational dismissals is the introduction of an active coach into the main sessions of the experiment, where one could investigate whether a coach's performance increases upon his dismissal.[13] If this is the case, rationality of the previous dismissal could be confirmed, which would provide evidence against scapegoating. This builds on our rational benchmark that poorly performing coaches can be rationally dismissed. While a concluding statement is not possible at this point, note that the performance of both the coach and the replacement coach proved to be identical in the pre-session ($x = 9$). Hence, from an objective point of view, a subject's decision to opt for the dismissal of his coach cannot be interpreted as being rationally motivated, since there is technically no scope for improvement in coach performance through replacement. As such, scapegoating as a psychologically founded reason for voting in favor of the replacement of the coach is sensible.

If a coach knows that he has to resign if his team fails to win an important match, he may be more likely to approach the match with the respective negative mindset. He might either decide not to put maximum effort into preparing the team, or make bad decisions under stress. Hence, in an experimental setting with an active coach, one could also study the behavior of the latter in more detail.

Likewise, the implementation of a control treatment where the replacement of coaches is not possible under any circumstance can provide additional insights into behavioral patterns when scapegoating is not feasible. A comparison with the results presented in this paper could be informative about whether guiding agents should stay fixed for a certain period of time in order to improve group performance.

Another group of scholars might regard prospect theory as a possible explanation for scapegoating behavior. According to the four-fold pattern of risk preferences, agents are relatively more risk seeking after experiencing losses with medium-sized probabilities. With a probability of $\frac{1}{2}$ of losing the game in period 1, where losing means a reduction of the respective subject's initial endowment, our setting captures the pattern's prerequisites. Dismissing the coach could thus be motivated by gambling behavior as predicted by the four-fold pattern. However, our results suggest that the underlying mechanism seems to be more complex than conjectured at first glance: We find that subjects who scapegoat tend to be relatively more (rather than less) risk averse, or, in other words, they tend to be relatively less (rather than more) risk seeking, which is contradictory to what prospect theory predicts.

Finally, one could also study whether scapegoating is independent of the level

[13] Recall that the experimental setting in this paper is such that coaches are only represented by their performance data from the pre-session in which they were selected for their respective positions based on their placing among the top 20% of study participants. Hence, their performance is intrinsically constant over time in the main sessions. An analysis of coach performance development is thus not feasible in this setting.

of performance, and if so, why this might be the case. Quite bizarre an example is that of the Icelandic national football team's coach who resigned after the team lost the match in the round of the last sixteen at the European Championships 2016, despite the fact that this has been by far the best result of the Icelandic national team in its entire history. When actually this should have been considered a huge success, given that this team had never even qualified for the Europeans before, it was in fact treated as a failure because the team lost in the end. Hence, it would be interesting to study whether scapegoating can arise at any relative level of losing or failure.

Hence, future research may be directed toward a wide array of different aspects. Scapegoating seems to be an important and potentially harmful, yet not very frequently studied aspect of team dynamics. We believe this to be a very interesting area of research in behavioral economics which should be addressed in more detail, and targeted at increasing team performance in all areas of public life.

## 2.7 Appendix

### 2.7.1 Instructions for the pre-session[14]

#### General instructions for study participants

Welcome to and thank you for participating in today's study. The study will take approximately 60 minutes.

In this study you can earn money. The amount that you will earn depends on your decisions and on the decisions of the other study participants. All earnings that you will make in this study will be shown to you in Experimental Currency Units (ECU) first, where the following conversion between ECU and CHF applies:

**20 ECU = 1 CHF**
**resp.**
**1 ECU = 0.05 CHF**

At the end of today's study, all earnings that you will have made will be converted into CHF using the above conversion rate. In any case, you will receive a participation fee of 15 CHF. You earnings will be paid to you **in cash**.

In order to earn money in this study, it is very important that you read the following instructions carefully.

**During the study, you are not allowed to talk to other participants or use electronic devices.** Non-observance of this rule will lead to exclusion from this study and all payments.

On the following pages, the course of this study will be described in detail. Please read this description carefully. On a supplementary sheet you will find a number of comprehension questions. Please answer these questions and raise your hand afterward. We will then come to your seat to check your answers and settle any remaining questions. Please wait until the answers of all participants are checked and there are no further questions. The study will begin afterward.

#### Description of the study

Today's study consists of **10 rounds**. In each round, your task is to **solve calculation problems in your head** within a certain time period. Within **one minute**, you should try to solve as many calculation problems as possible. Each calculation problem consists of adding three two-digit numbers between 10 and 99. Here is an

[14]The first section is written along the lines of the instructions used in Brune [2011]. The instructions distributed to the subjects were written in German.

example:

$$21 + 45 + 83 = \ldots$$

**Ten** such problems will be shown in each round. For each correctly solved problem you will receive **one point**. At the end of each round, the number of points that you obtained in this round will be shown to you on the screen.

At the end of round 10, the average number of problems that you solved correctly (and thus your **average number of points**) over all rounds will be calculated. If this average number of points puts you in the **top 20% of all participants of today's study**, you will receive a **bonus of 100 ECU**. Furthermore, your payment will include a remuneration of **40 ECU per point** for the number of points that you obtained in **a randomly selected period**.

In order to prepare for the task (solving calculation problems in your head), you will perform a **trial round** on the screen before the beginning of the actual study. This is for you to familiarize yourself with the task's format and type before the actual study begins.

Your earnings at the end of today's study thus consists of your participation fee of 15 CHF as well as the number of points from a round being randomly selected from all ten rounds, which will be remunerated with 40 ECU per point. If you are among the top 20% of all participants of today's study, you will also receive the bonus of 100 ECU.

Finally, at the end of today's study you will be asked to fill in a questionnaire on the computer screen.

**Comprehension questions**

Now, please answer the enclosed comprehension questions regarding the procedure of today's study and raise your hand afterward. We will then come to your seat to check your answers and settle any further questions. As soon as all answers of all participants have been checked and there are no further questions, please wait until the study begins.

### 2.7.2 Control questions and solutions for the pre-session

**Question 1:** How many rounds are there in today's study?
**Answer:** 10.

**Question 2:** Assume that in any round, you solved 9 problems and 8 of those were solved correctly. Also assume that this round is selected for your payment. How many ECU will you receive in this case?

**Answer:** $8 \cdot 40 = 320$.

**Question 3:** Assume that your average number of points at the end of the study is six, and assume that this puts you in the top 30% of all participants of today's study. In this case, will you receive the bonus of 100 ECU for sure?
**Answer:** No: for the bonus it is required to be among the top 20%.

**Question 4:** Assume that you solved 7, 8, 6, 9, 4, 7, 5, 9, 10, and 6 problems correctly in the respective rounds, and that with the resulting average number of points of 7.1 you will be among the top 20% of all participants of today's study. Also assume that the randomly selected number of points is 5. In this case, which payment will you receive at the end of today's study (including the participation fee)?
**Answer:** $\frac{5 \cdot 40}{20} + \frac{100}{20} + 15 = 30$.

### 2.7.3 Instructions for the main session[15]

**General instructions for study participants**

Welcome to and thank you for participating in today's study. The study will take approximately 60 minutes.

In this study you can earn money. The amount that you will earn depends on your decisions and on the decisions of the other study participants. All earnings that you will make in this study will be shown to you in Experimental Currency Units (ECU) first, where the following conversion between ECU and CHF applies:

**20 ECU = 1 CHF**
**resp.**
**1 ECU = 0.05 CHF**

At the end of today's study, all earnings that you will have made will be converted into CHF using the above conversion rate. In any case, you will receive a participation fee of 15 CHF.[16] You earnings will be paid to you **in cash**.

In order to earn money in this study, it is very important that you read the following instructions carefully.

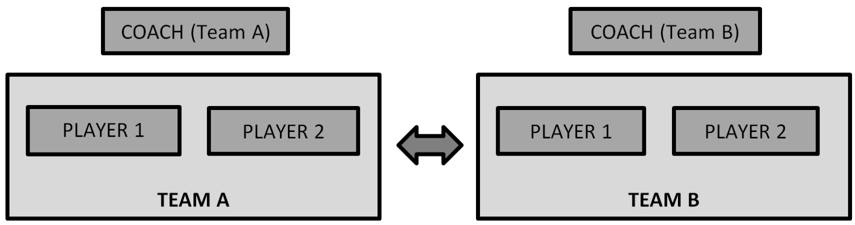
**During the study, you are not allowed to talk to other participants or use electronic devices.** Non-observance of this rule will lead to exclusion from this study and all payments.

[15] The first section is written along the lines of the instructions used in Brune [2011]. The instructions distributed to the subjects were written in German.
[16] In the pilot session, the show-up fee was 10 CHF.

On the following pages, the course of this study will be described in detail. Please read this description carefully. On a supplementary sheet you will find a number of comprehension questions. Please answer these questions and raise your hand afterward. We will then come to your seat to check your answers and settle any remaining questions. Please wait until the answers of all participants are checked and there are no further questions. The study will begin afterward.

**Description of the study**

Today's study consists of two rounds. You are the member of a team of two team members in total. You will be randomly assigned to either team A or team B. Moreover, each team has a coach. The task for each team and its coach is to collect as many points as possible during a productive game. Your team and your coach will play against another team with their coach:




The team with the larger number of points wins in this round.

The productive game consists of **solving calculation problems in your head** within a certain time period. Within **one minute**, you should try to solve as many calculation problem as possible. Each calculation problem consists of adding three two-digit numbers between 10 and 99. Here is an example:

$$21 + 45 + 83 = \ldots$$

**Ten** such problems will be shown in each round. For each correctly solved problem you will receive **one point**. The same applies for the other team member and the coach. The total number of points for your team is then calculated as follows:

**team points = coach's points * (your points + your teammate's points)**

This number of points of the team determines the game's result. The team with the larger number of points will win.

Each study participant starts out with an **endowment of 200 ECU**. If your team wins, then you and the other team member will each receive the difference between the number of points of your team and the number of points of the other team in ECU. If your team loses, this amount will be subtracted from your and the other team member's endowment:

| **distance to other team = your team points - other team's points** |
|---|

**Here is an example:** Assume that in round 1, your team generates 80 points more than the other team and thus wins the game. In this case, you and the other team member will receive 80 points each to be added to your endowment of 200 ECU. Your endowment thus **increases** to 280 ECU. Each member of the other team, who lost the game by 80 points, will lose 80 points each of their endowment. Their endowment thus **decreases** to 120 ECU.

The respective coach is not sitting with you in today's study, but already participated in a previous study. There he solved the **same** problems in the **same** time as you will do today, however over 10 rounds. One of his number of points was randomly selected from among these 10 rounds, saved, and used for this study. Your coach was randomly selected from among the **top 20% of participants** of this previous study based on the average number of points achieved at the end of these 10 rounds.

After the game, the number of points of each team is calculated and thus the game's result determined. You will be informed about whether you **won or lost** against the team that your played against, as well as **by how many points** you won or lost. You will also receive **information** about the amount by which this **changes** your **current endowment** and **about the number of points of some study participants involved**.

Afterward, you will be asked **who, in your opinion, bears the main responsibility for your game's outcome**. You will have the following response options:

- ○ the coach
- ○ yourself
- ○ the other team member
- ○ the team (that is, yourself and the other team member together)

You and the other team member will choose their answers **at the same time**. **Afterward**, your response will be send to the other team member and **at the same time**, you will receive the response of the other team member.

If your team **lost** and both you and your teammate think that the coach is responsible for this, then he will be **replaced** in the next round. Your team will then get a **new coach** who also belonged to the **top 20% of participants** of the previous study. If your team **won**, the coach **cannot** be replaced.

A second round of the same game follows, in which you will play against the **same** team as before. However, there will be no more voting as the game ends after two rounds.

In order to prepare for the task (solving calculation problems in your head), you will perform a **trial round** on the screen before the beginning of the actual study. **This will only cover the calculation part.** This is for you to familiarize yourself with the task's format and type before the actual study begins.

Your earnings at the end of today's study consists of your participation fee, your initial endowment of 200 ECU as well as your gains or losses from the two games in rounds 1 and 2.

Finally, at the end of today's study you will be asked to fill in a questionnaire on the computer screen.

**Comprehension questions**

Now, please answer the enclosed comprehension questions regarding the procedure of today's study and raise your hand afterward. We will then come to your seat to check your answers and settle any further questions. As soon as all answers of all participants have been checked and there are no further questions, please wait until the study begins.

### 2.7.4 Control questions and solutions for the main session

**Question 1:** How many rounds are there in today's study?
**Answer:** 2.

**Question 2:** Out of how many study participants does your team consist?
**Answer:** 2.

**Question 3:** You start out with an endowment of 200 ECU. If your team wins by 50 points in the first round, what is your updated endowment after round 1?
**Answer:** $200 + 50 = 250$.

**Question 4:** You start out with an endowment of 200 ECU. If your team loses by 90 points in the first round and wins by 10 points in the second round, what is your updated endowment at the end of round 2?
**Answer:** $200 - 90 + 10 = 120$.

**Question 5:** Assume that your updated endowment at the end of the study is 240 ECU. In this case, which payment will you receive in CHF (including the participation fee)?
**Answer:** $\frac{240}{20} + 15 = 27$.

**Question 6:** If your coach contributes 10 points and you and the other team member contribute 8 points together, what is the number of points for the team?
**Answer:** $10 \cdot 8 = 80$.

**Question 7:** Assume that your coach solved 7, 8, 6, 9, 4, 7, 5, 9, 10, and 6 problems correctly in the respective rounds, and that with the resulting average number of points of 7.1 he was among the top 20% of all participants of the previous study. How many points will he be able to contribute as a minimum and as a maximum then?
**Answer:** The coach will be able to contribute at least 4 points and at most 10 points, where his contribution is a random selection from among the ten rounds.

**Question 8:** In which case will the coach be replaced for round 2?
**Answer:** The coach will be replaced if both members of a losing team attribute responsibility for this outcome to the coach in the first round.

# References

Bartling, B., and U. Fischbacher (2012): "Shifting the blame: On delegation and responsibility," *Review of Economic Studies*, 79(1), 67–87.

Boeker, W. (1992): "Power and managerial dismissal: Scapegoating at the top," *Administrative Science Quarterly*, 37(3), 400–421.

Brown, K., and C. White (2010): "Organization–public relationships and crisis response strategies: Impact on attribution of responsibility," *Journal of Public Relations Research*, 23(1), 75–92.

Brune, A. (2011): "Endogenous authority: Design and analysis of a laboratory experiment," Master's thesis, Department of Economics, University of Zurich.

Chang, T., D. Solomon, and M. Westerfield (2016): "Looking for someone to blame: Delegation, cognitive dissonance, and the disposition effect," *Journal of Finance*, 71(1), 267–302.

Fischbacher, U. (2007): "Z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, 10(2), 171–178.

Gamson, W., and N. Scotch (1964): "Scapegoating in baseball," *American Journal of Sociology*, 70(1), 69–72.

Gibbons, R. (1992): *A primer in game theory*. Prentice Hall, England.

Kahneman, D., and A. Tversky (1979): "Prospect theory: An analysis of decision under risk," *Econometrica*, 47(2), 263–291.

Rothschild, Z., M. Landau, D. Sullivan, and L. Keefer (2012): "A dual-motive model of scapegoating: Displacing blame to reduce guilt or increase control," *Journal of Personality and Social Psychology*, 102(6), 1148–1163.

# Part III

# Appendix

# Appendix A

## The war puzzle: Contradictory effects of international conflicts on stock markets[1]

Amelie Brune, Thorsten Hens, Marc Oliver Rieger, Mei Wang

**Abstract.** We study stock market reactions to large international military conflicts since World War II. Using a news analysis proxy for the estimated likelihood that a conflict will result in a war, we find that an increase in the war likelihood tends to decrease stock prices, but the ultimate outbreak of a war increases them. In cases when a war starts surprisingly, however, the outbreak of a war decreases stock prices. We show that this puzzle cannot be explained by risk or ambiguity aversion or by expectations about a quick end of the war.

**JEL classification:** G11, G14, G19.

**Keywords:** International conflicts, war, stock market reaction, news analysis, behavioral finance.

[1] This paper is published as Brune, A., Hens, T., Rieger, M.O. et al. Int Rev Econ (2015) 62: 1. https://doi.org/10.1007/s12232-014-0215-7.

## A.1 Introduction

"Buy on the sound of the cannon, sell on the sound of the trumpet." is an old proverb from the Napoleonic wars, attributed to London financier Nathan Rothschild. It suggests that the start of a war is a good time to buy stocks and that they should be sold once the war is over. The rationale behind this advice is that investors tend to overreact to the bad news of a coming war, leading to underpricing, and that they overreact to the good news of the end of a war in a similar way, leading to overpricing.

Economists have been concerned with the causes and consequences of international military crises for more than 100 years, one of the first contributions being the work of Keynes [1919]. Further work focused on the interplay between economy and war (Holsti and North [1966], Russett and Hanson [1975], and Collier and Hoeffler [1998]), as well as on the impact of war on financial markets throughout history (Frey and Kucher [2000]).

Recently, the financial market reaction induced by international conflicts has received substantial interest in finance. News analysis has been used to reflect the perceived risks and consequences of wars. In an early pioneering study using content analysis of news, Holsti and North [1966] demonstrate that security prices are related to rising international tensions during the outbreak of World War I. In a more general framework, Niederhoffer [1971] codes the headlines on world events on a seven-point good-bad scale in order to examine the impacts of news, and Cutler, Poterba, and Summers [1998] analyze the stock market reaction to world news (including wars) from 1926 to 1985. To quantify the effect of international conflicts on stock prices, Goldstein [1992] constructed a scale to code conflictive versus cooperative events that can be used to assess the impact of war risk, as it was pointed out by Schneider and Troeger [2006]. There are further studies on the relation between world news and financial markets, in particular during times of crises: Elmendorf, Hirschfeld, and Weil [1996] study British bond market reactions to news by comparing weeks with important news to those without such news. Similarly, Rigobon and Sack [2005] distinguish the days with war-related news to those without these news to measure the impact of war risk on financial markets.

Whereas all of these studies either consider particular conflicts or the impact of war risk on financial markets during (at first glance) peaceful times, in this paper we want to focus on a slightly different aspect, namely wars with a "prologue", i.e., wars which follow a period of tension. In this type of war, we distinguish two phases: In the first phase, where cannons are still silent, there is an increasing danger of a war, possibly interrupted by times of hope for a peaceful resolution. Afterward, in the second phase, the "sound of the cannon" starts and war breaks out.

The surprising finding of our analysis is that often the reaction of the stock market to the likelihood of a war is different between these two phases: Whereas in the pre-war phase an increased likelihood of war *decreases* market prices, the outbreak of the war itself, so to speak the increase in the likelihood from 99 to 100%, *increases* market prices. On the day of the outbreak of some specific war, news coverage is clearly focused on this particular event. The development of the stock market index then reflects the market's reaction to and evaluation of the event, which is thus far from random.

We observe this striking puzzle for a number of larger wars, including the Iraq War and World War II. In all cases, we study the impact on the US stock market as measured by the S&P 500 or Dow Jones Industrial Average. To complement our results, we also analyze all other international crises with large impact since World War II, in particular wars that started unexpectedly, such as the Korean War.

This paper is organized as follows. Section A.2 discusses different proxies to estimate the likelihood of war. Section A.3 analyzes the Iraq War in 2003 using different likelihood measurements and shows that a puzzle in the development of stock prices along the evolvement of this international crisis can be observed. Using further examples of international wars in Section A.4, we show that this phenomenon does not seem to be restricted to the war in Iraq. Section A.5 discusses possible explanations for the puzzle and rejects some – at first glance natural – explanations. Section A.6 concludes.

## A.2 Estimating the likelihood of war

To analyze market reactions to changes in war likelihood, we first need to find appropriate proxies for this likelihood. To this end, we analyze the Iraq War as several independent estimates for the probability of war are available here. This war then serves as a benchmark to establish a simple yet robust proxy using news that we can then readily apply to earlier wars where other data are not available.

The Iraq crisis started to become serious on January 28, 2003, when President George W. Bush announced a possible attack on Iraq even in the absence of a UN resolution legitimating this course of action. After the ultimatum proposed to Saddam Hussein on March 17 to leave Iraq within 48 h had expired, the first military intervention by the USA started immediately in the early morning of March 20. Later that morning, President George Bush formally announced the Operation Iraqi Freedom.

Following Wolfers and Zitzewitz [2009], we use two independent estimates for the probability that a war would take place in Iraq: The first is the so-called Saddameter, an expert estimate for the likelihood of an invasion of Iraq which was published on a daily basis by William Saletan on www.slate.com. This estimate provides us with data from November 2002 to March 18, 2003. The second is the so-called Saddam Security, a security that existed on the online exchange www.tradesports.com and which was designed to pay a certain amount if and only if Saddam Hussein, president of Iraq, were still in power at a certain date. As a war was expected to end his rule over Iraq, Saddam Security prices gave a good probability estimate for the likelihood of war before the designated date (see Amihud and Wohl [2004]). This security was available with different maturity dates, where only March 2003 and June 2003 securities had a long enough price series. We use the June security, as the pattern of the March security is somehow obfuscated by the fact that even though the probability for a coming war increased, it was at times not at all clear whether the eventual war was over before the end of March.

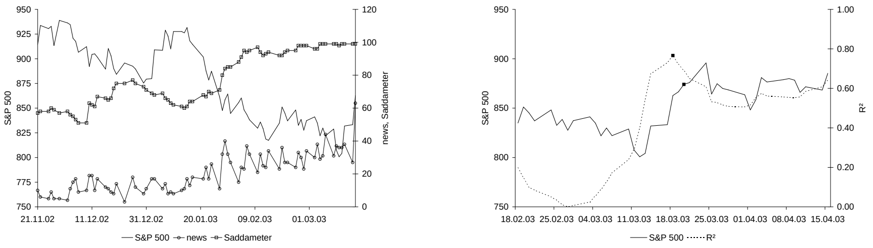
| | Saddam Security June | Iraq War news |
|---|---|---|
| Saddameter | $\rho = 0.86$ | $\rho = 0.66$ |
| | $N = 56$ | $N = 81$ |
| Saddam Security June | | $\rho = 0.64$ |
| | | $N = 78$ |

Table A.1: Correlations between different proxies for the probability of war. All correlations are significant at the 1% level and use the largest available number of days from the discussed time series.

As another proxy we studied news data from the New York Times. We counted how many articles with the key words “war” *and* “Iraq” appeared in each day’s issue. To show that this variable is closely related to the above-mentioned probability measurements and that it provides the same qualitative indications about the war likelihood, we regress each of the two measurements (i.e., the Saddameter and the Saddam Security) on our news proxy. The results are reported in Table A.1. We see that the news variable is positively correlated with the Saddameter, that is, the expert measurement on the war likelihood and the number of (mostly negative) news items run in the same direction. The comovement between the Saddameter and the news variable between November 2002 and March 2003 is depicted in Figure A.1(a) and is negatively related to the development of the S&P 500. Similarly, the news proxy is positively related to the Saddam Security.

## A.3 The Iraq War puzzle

As mentioned earlier, a puzzle can be observed when analyzing data of stock market prices prior to and at the onset of a war. The relation between stock market prices and the probability of an international conflict developing into a war is negative




(a) Probability of a war

(b) Structural break analysis

Figure A.1: Development of the S&P 500 at the beginning of the Iraq War. The largest local $R^2$ was reached for $\tilde{t}$ = March 18, 2003, while the invasion of Iraq took place two days later on March 20, 2003.

as long as the conflict is still evolving. However, it becomes positive at the actual onset of the war, thus a positive shock in the stock market can be observed when the war likelihood increases to 100%. In other words, an increasing war likelihood seems to decrease stock prices, while the outbreak of the war itself seems to increase them. This is what we call the "war puzzle".

To provide empirical evidence for this puzzle, we begin with an analysis of the Iraq War in 2003. The main reason for this approach is the fact that for this war, we are able to use two alternative measurements of the probability of war, namely the Saddameter and the news variable described in Section A.2. As previously explained, this has the advantage that we are able to conduct econometric analyses for both types of probability measures in the same war. Hence, showing that the news variable is a valid proxy for war likelihood, we can then use this variable to analyze periods of war where no alternative probability measures are available.

| | | Intercept | Intercept and trend |
|---|---|---|---|
| Levels | SP500 | −1.5167 | −2.5662 |
| | news | 0.1771 | −7.0122*** |
| | Saddameter | −0.5129 | −1.8901 |
| First differences | SP500 | −9.6884*** | −9.6221*** |
| | news | −7.7683*** | −7.8207*** |
| | Saddameter | −7.8386*** | −7.7840*** |

Table A.2: Iraq War: ADF test statistics for unit roots. * denotes significance at 10%, ** at 5%, and *** at 1% level.

Analyzing the characteristics of the S&P 500, news, and the Saddameter, we find that these time series are non-stationary in levels, but stationary in first differences (see Table A.2 for augmented Dickey-Fuller test statistics for unit roots). Moreover,

| | | Intercept | Intercept and trend |
|---|---|---|---|
| Levels | $\varepsilon_t = \text{SP500}_t - \beta_0 - \beta_1 \text{news}_t$ | −3.0972** | −5.4364*** |
| | $\varepsilon_t = \text{SP500}_t - \beta_0 - \beta_1 \text{Saddameter}_t$ | −3.8389*** | −3.7985** |
| First differences | $\varepsilon_t = \text{SP500}_t - \beta_0 - \beta_1 \text{news}_t$ | −13.2052*** | −13.1416*** |
| | $\varepsilon_t = \text{SP500}_t - \beta_0 - \beta_1 \text{Saddameter}_t$ | −9.2592*** | −9.2003*** |

Table A.3: Iraq War cointegration tests: ADF test statistics for unit roots in residuals $\varepsilon_t$.

we find evidence for a cointegrating relationship between the S&P 500 and each probability measure of order one (see Table A.3). This leads us to the choice of an error correction model to appropriately model the relation between the US stock market and war likelihood. The econometric models employed to study the stock market in the pre-war phase are

$$\begin{aligned}\Delta\text{SP500}_{t-1} &= \beta_1\Delta\text{SP500}_{t-2} + \beta_2\Delta\text{news}_t + \gamma[\text{SP500}_{t-2} - \alpha_0 - \alpha_1\text{news}_{t-1}] + \varepsilon_t \\ \Delta\text{SP500}_{t-1} &= \beta_1\Delta\text{SP500}_{t-2} + \beta_2\Delta\text{Sadd}_{t-1} + \gamma[\text{SP500}_{t-2} - \alpha_0 - \alpha_1\text{Sadd}_{t-1}] + \varepsilon_t\end{aligned}$$

where ΔSP500 measures absolute daily changes in the S&P 500, and the error correction parameter $\gamma$ captures the degree to which deviations from an equilibrium in the previous period (captured by the error term) affect current values. Hence, $\gamma$ is expected to be negative if a long-run equilibrium relationship prevails. Note that the functional form of most models used in this and the following sections is such that the explanatory variable is lagged forward relative to the dependent variable. While this may seem slightly counterintuitive at first sight, the reason is that there is a certain delay between the stock market pricing an event and its actual publication in the newspaper.[2]

The estimated coefficients are presented in columns 1 and 2 of Table A.4. Both news and the Saddameter are highly significant at the 1 and 5% level, respectively, and indeed indicate a negative relationship between stock prices and war likelihood. An increase in the probability of war is thus associated with decreasing stock market prices.[3] In particular, an increase in the difference of the number of news items, $\Delta\text{news}_t$, by one unit leads to a decrease in the difference in the S&P 500, $\Delta\text{SP500}_{t-1}$, by 0.507 points. Similarly, an increase in the difference of the Saddameter value, $\Delta\text{Sadd}_{t-1}$, by one percentage point leads to a decrease in the

[2]Standard predictive models are often constructed such that changes in variables at time $t+1$ are explained by changes in variables at time $t$. In our case this approach would require to use a variable for the war likelihood which realizes prior to actual changes in the stock market. However, since our argument is to use the news proxy as a flexible probability measurement which is available for different wars over time, the functional forms used here slightly differ from standard models due to the delay of news publication. This, however, does not alter the models' predictive power.

[3]Also, the error correction parameter $\gamma$ is significantly negative as expected for both types of probability measurements.

| War | Iraq | Iraq | World War II | Vietnam War | Gulf War |
|---|---|---|---|---|---|
| Method | ECM | ECM | first differences | first differences | first differences |
| Likelihood proxy | news | Saddameter | news | news | news |
| $\Delta \text{news}_{t-1}$ | | | −0.1946** | −0.1160*** | −0.2207 |
| | | | (0.0741) | (0.0384) | (0.1324) |
| $\Delta \text{news}_t$ | −0.5072*** | | −0.2844*** | −0.0456 | −0.2885** |
| | (0.1835) | | (0.0863) | (0.0384) | (0.1281) |
| $\Delta \text{news}_{t+1}$ | | | −0.2651*** | | |
| | | | (0.0697) | | |
| $\text{news}_{t-1}$ | −3.6611*** | | | | |
| | (0.5644) | | | | |
| $\Delta \text{Sadd}_{t-1}$ | | −1.1147** | | | |
| | | (0.4927) | | | |
| $\text{Sadd}_{t-1}$ | | −2.3051*** | | | |
| | | (0.1892) | | | |
| $\Delta \text{SP500}_{t-2}$ | −0.0530 | 0.0512 | | | 0.1588 |
| | (0.1104) | (0.1110) | | | (0.1170) |
| $\Delta \text{DJI}_{t-2}$ | | | −0.0004 | 0.1519*** | |
| | | | (0.1181) | (0.0553) | |
| $\gamma$ | −0.2274*** | −0.3940*** | | | |
| | (0.0681) | (0.1020) | | | |
| Constant | | | 0.1827 | 0.2032 | 0.1788 |
| | | | (0.1859) | (0.1609) | (0.3735) |
| $R^2$ | 0.1680 | 0.2767 | 0.2276 | 0.0510 | 0.1098 |
| $N$ | 78 | 78 | 69 | 311 | 72 |
| From | Nov 21, 2002 | Nov 21, 2002 | May 25, 1939 | May 6, 1964 | Oct 10, 1990 |
| To | Mar 18, 2003 | Mar 18, 2003 | Aug 31, 1939 | Mar 12, 1965 | Jan 1, 1991 |

Table A.4: Regression results for the effect of the war likelihood on stock prices. Standard errors are given in parenthesis. * denotes significance at 10%, ** at 5%, and *** at 1% level.

difference in the S&P 500, $\Delta \text{SP500}_{t-1}$, by 1.112 points. The one-day lag in the news variable can be attributed to the time required for physical publication of the newspaper. Since both probability measures yield the same implications, they may both be regarded as valid proxies, which ensures that we can employ the news variable as a valid proxy also in periods where alternative probability measures are not available.

To study the behavior of the stock market at the onset of the war, we conduct an analysis for structural breaks in the development of the stock market index. To this end, cumulative returns of the S&P 500 are regressed on a binary dummy $D_t$ defined as

$$D_t = \begin{cases} 0 & \text{if } t < \tilde{t} \\ 1 & \text{if } t \geq \tilde{t} \end{cases}$$

where $\tilde{t}$ indicates a potential structural break, and a time window of size 51[4] around $\tilde{t}$ is used to obtain a set of models of type

$$\text{SP500}_t = \beta_0 + \beta_1 D_t + \varepsilon_t.$$

[4]That is, the data set ranges from $\tilde{t} - 25$ to $\tilde{t} + 25$ days around $\tilde{t}$.

That is, $\tilde{t}$ is varied over time until the largest associated $R^2$ is found, where we suppose that the actual onset of a war, and hence the observed structural break denoted by $t_0$, lies in close proximity.[5] Note that while we are aware of potential problems of non-stationarity in the cumulative returns, our analysis of structural breaks would not be feasible using first differences here. Our interest lies in potential increases in stock market values at the beginning of a war, which has to be separated from the question of average increases or changing trends. The problem thus is to weight an interesting question with the obvious problems of not using differences in the model in order to resolve the issue.

The results are illustrated in Figure A.1(b). The largest $R^2$ obtained amounts to 0.7665 and is associated with March 18, 2003, while on March 20, 2003, the invasion of Iraq took place ($R^2 = 0.6898$). Hence, the true structural break can be relatively well predicted.[6] The S&P 500 increases by 46.4 points at the time of the predicted structural break and by 47.8 points at the time of the invasion two days later. Although it may prove difficult to interpret this absolute change in the stock market index at the onset of the war and compare it across conflicts, we refrain from calculating relative changes, for instance. As the volume of news items clearly differs over time, a comparison of relative values over time would be meaningless and thus reveal no further insights for the analysis.

## A.4 The war puzzle: A general phenomenon

The results of the previous section could either be a general phenomenon or just an idiosyncrasy of the Iraq War. In this section, we therefore extend the analysis to other wars with large international impact since World War II. We face two problems here: First, we need to decide which wars to include, and second, the number of large scale wars during this time was – fortunately – not big enough to use a single econometric approach for all wars. Instead of performing a unified analysis, we are forced to study the conflicts separately. As any event study involving data on stock market indices during international military conflicts, our subsequent analysis is necessarily based on a small number of observations.

We solve the above-mentioned selection problem by using the list of the most costly wars to the USA (and thus arguably to the US economy) as composed by

[5] We use this type of structural break analysis since we are interested in whether there exists an instantaneous jump in the stock market index at the onset of a war. An alternative question would be whether there are possible changes in the trend of the stock market. For that purpose, a model testing for changes in expected growth rates of stock returns may be appropriate (see Amihud and Wohl [2004]).

[6] Note that the estimated structural break on March 18 coincides with the beginning of the US ultimatum to Saddam Hussein. Investors might have anticipated that this ultimatum would not be agreed upon which caused an early rise in stock market values.

Stephen Daggett[7] (see Table A.5). In the following, we will study each of the wars from this list, where we distinguish between wars with a surprising start and wars with a longer prelude. Since the wars took place at very different times, with news traveling at different speeds, the econometric models we apply necessarily differ in details.

| | Costs in billion 2008 US$ | % of GDP |
|---|---|---|
| World War II | 4114 | 35.8 |
| Korean War | 320 | 4.2 |
| Vietnam War | 686 | 2.3 |
| Gulf War | 96 | 0.3 |
| Iraq War | 648 | 1.0 |
| Afghanistan War | 171 | 0.3 |

Table A.5: List of the most costly wars to the USA (source: Stephen Daggett, CRS Report for Congress, Costs of Major U.S. wars, July 2008).

### A.4.1 World War II (1939-1945)

To analyze World War II, two starting dates can be studied: On the one hand, the German invasion of Poland on September 1, 1939, and on the other, the Japanese attack of Pearl Harbor on December 7, 1941, dragging the USA into the war.

We first study the former event as it had a large prelude while we will study the latter rather surprising one in Section A.4.6. To analyze it in an econometric setting, a first differences approach of the form

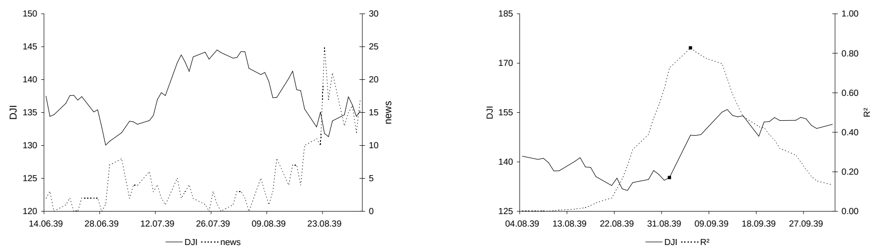
$$\Delta\text{DJI}_{t-1} = \beta_0 + \beta_1 \Delta\text{DJI}_{t-2} + \beta_2 \Delta\text{news}_{t+1} + \beta_3 \Delta\text{news}_t + \beta_4 \Delta\text{news}_{t-1} + \varepsilon_t$$

is most appropriate. The stock market is described by absolute daily changes in the Dow Jones Index, ΔDJI, and the war likelihood is from now on captured by the news proxy as established in Section A.2, i.e., in this case the number of articles in the New York Times featuring the key words "war" *and* "Poland". Figure A.2(a) depicts the relation between the Dow Jones Index and the probability of war between June and September 1939. Column 3 of Table A.4 shows the regression results obtained from estimating this equation. The marginal effect of the news variable is significantly negative for all different lags $\Delta\text{news}_{t-1}$, $\Delta\text{news}_t$, and $\Delta\text{news}_{t+1}$, and ranges from -0.195 ($\Delta\text{news}_{t-1}$) to -0.284 ($\Delta\text{news}_t$) points. Figure A.2(b) illustrates the results obtained from the structural break analysis of type

$$\text{DJI}_t = \beta_0 + \beta_1 D_t + \varepsilon_t.$$

[7]CRS Report for Congress, Costs of Major U.S. wars, July 2008.

September 5, 1939, is identified as the most likely structural break in the data set ($R^2 = 0.8274$). In fact, this lies close to the German attack of Poland being the critical event on September 1, 1939 ($R^2 = 0.7244$). The time lag may be explained by the slower speed of news publication at that time.



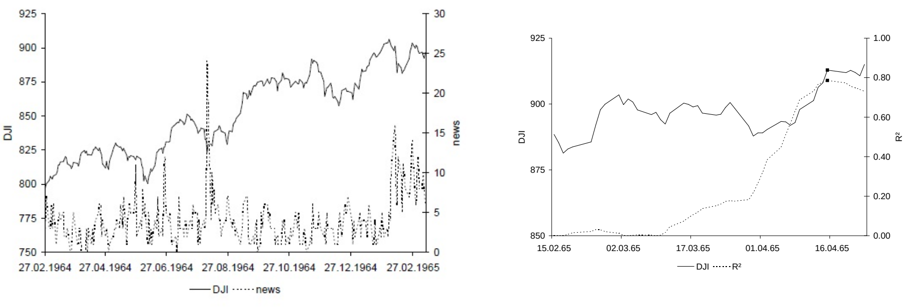

(a) Probability of a war

(b) Structural break analysis

Figure A.2: Development of the Dow Jones Index at the beginning of World War II. The largest local $R^2$ was reached for $\tilde{t}$ = September 5, 1939, while the German attack of Poland took place on September 1, 1939.

### A.4.2 Vietnam War (1955-1975)

The Vietnam War distinguishes itself from other wars by the fact that its time line is less clear, but the onset of the large scale US involvement is set in 1965. The development of the Dow Jones Index and the news variable between February 1964 and March 1965 is depicted in Figure A.3(a).




(a) Probability of a war

(b) Structural break analysis

Figure A.3: Development of the Dow Jones Index during the Vietnam War. The largest local $R^2$ was reached for $\tilde{t}$ = April 15, 1965, which coincides with the first US and South Vietnamese bombing of Viet Cong positions on that day.

The most suitable model for this war is a time series approach of type

$$\Delta\text{DJI}_{t-1} = \beta_0 + \beta_1 \Delta\text{DJI}_{t-2} + \beta_2 \Delta\text{news}_t + \beta_3 \Delta\text{news}_{t-1} + \varepsilon_t.$$

The regression results for this model are presented in column 4 of Table A.4. Again, we see a significant negative relation between the war likelihood and the development of stock prices. Figure A.3(b) illustrates the results for the analysis of structural breaks. We see that the ad hoc regression identifies April 15, 1965, as the event associated with the largest coefficient of determination, $R^2 = 0.7860$. In fact, one critical event coincides with this date: It marks the first US and South Vietnamese bombing of Viet Cong positions. The jump in cumulative returns is significantly positive at the 1% level and amounts to an increase in the Dow Jones Index by 28.01 points.

Note that, for the war in Vietnam, the analysis does not exactly describe the pre-war period and the distinct outbreak of the war itself, but nevertheless proves to be a useful method to analyze the reaction of cumulative stock returns to war news within a prolonged period of tensions.

### A.4.3 Gulf War (1990-1991)

The Gulf War[8] started with the invasion of Kuwait by Iraq on August 2, 1990, and ended with the defeat of Iraq by the USA and their allies on February 28, 1991. The relevant date for our analysis is January 17, 1991, when the Operation Desert Storm (with the goal of liberating Kuwait) was started by massive aerial bombing. The Gulf War moved into its second and much larger phase then. As such, our news series covers the time span up to this date, which can be interpreted as a pre-war phase from the perspective of the USA and their allies.

Analyzing this relation with a time series approach, column 5 of Table A.4 shows the results. We see once more that there is empirical evidence for a significant negative relation between news and stock market prices. The analysis for structural breaks in the development of cumulative returns (see Figure A.4) identifies February 4, 1991, as the break point ($R^2 = 0.8457$) which, however, lies some days after the initial aerial bombing. In fact, the closest real event to this indication is the Battle of Khafji being the first major ground battle in the Gulf War on January 29, 1991.

### A.4.4 Afghanistan War (2001)

In contrast to the conflicts analyzed in the previous sections, the Afghanistan War, starting in October 2001, had too short a prelude (beginning only after Afghanistan's involvement in the attack on September 11, 2001, became clear)

[8]The Gulf War is often also referred to as Second Gulf War or Persian Gulf War.

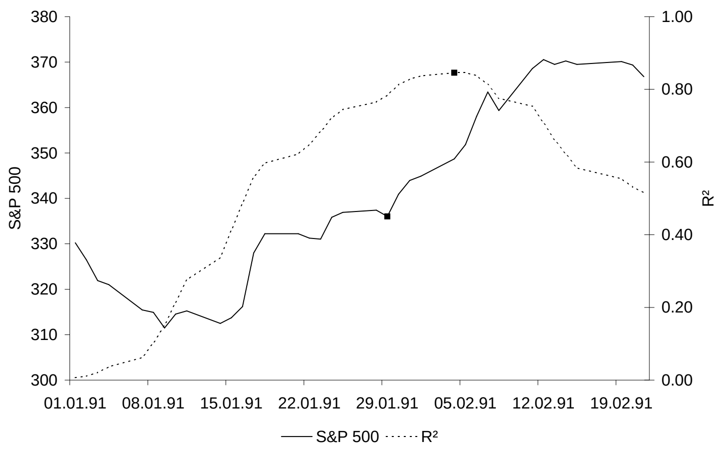


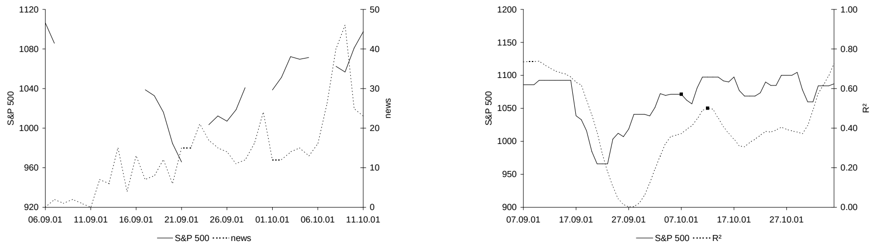

Figure A.4: Development of the S&P 500 and explanatory power of structural break analysis at the beginning of the Gulf War. The largest local $R^2$ was reached for $\tilde{t}$ = February 4, 1991, while the Battle of Khafji took place on January 29, 1991.

to be analyzed in an econometric setting. However, we can still perform an analysis for structural breaks and show that there exists an increase in stock market prices at the onset of the war. Figure A.5(a) depicts the development of the S&P 500 and news between September and October 2001. Figure A.5(b) shows the empirical results from the analysis for structural breaks. Again, an increase in the S&P 500 around the time of the start of the war can be identified.

### A.4.5 Behavior of other large stock market indices

Do stock markets in other countries behave similarly to the US market? For the two recent wars with a sufficiently long prelude, i.e., the Gulf War and the Iraq War, Figure A.6 shows that Japanese, German, and European stock indices mirror the development of the S&P 500 quite well. The analysis for structural breaks (see Table A.6) shows particularly for the Gulf War simultaneous up-moves at the onsets of war.[9] The fact that the USA, Europe, and Japan are strongly interconnected in economic terms may further account for the clear comovement of Nikkei, DAX, and Eurostoxx which can be observed during these two wars in the Middle East. A

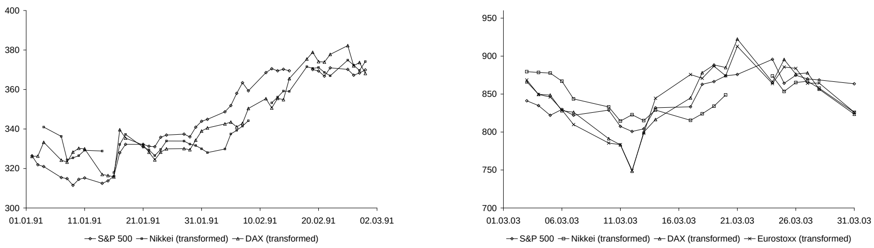
[9]Although it would be interesting to see how stock markets evolved in the countries were war actually took place, this analysis is often not feasible. Either stock exchanges were not yet institutionalized, or controlled by the government so that data would not be reliable, or data is not available.




(a) Probability of a war (b) Structural break analysis

Figure A.5: Development of the S&P 500 during the Afghanistan War. The largest local $R^2$ was reached for $\tilde{t}$ = October 12, 2001, while the air attack took place on October 7, 2001.

similar, though less pronounced, comovement also exists for the Afghanistan War.[10]




(a) Gulf War (b) Iraq War

Figure A.6: Comovement of other stock market indices. For the sake of comparability, the data series of Nikkei, DAX, and Eurostoxx are linearly transformed such that all series have a mean equal to the mean of the S&P 500.

### A.4.6 Wars "out of the blue"

So far we have analyzed wars that had a more or less pronounced prelude. However, there are also some wars with a very fast and unexpected onset.[11] Most notably, these are:

[10] Due to issues of data availability, we could not perform a similar analysis for global wars having occurred earlier in history.

[11] Li and Sacko [2002] find that an unexpected onset of a military dispute reduces bilateral trades more severely.

| | Critical event | Estimated breaks | | | |
|---|---|---|---|---|---|
| | | S&P 500, DJI | Nikkei | DAX | Eurostoxx |
| Iraq War | Mar 20, 2003 | Mar 18, 2003 | Mar 7, 2003 | Apr 7, 2003 | Feb 25, 2003 |
| Gulf War | Jan 29, 1991 | Feb 4, 1991 | Feb 5, 1991 | Jan 30, 1991 | N/A |

Table A.6: Analyses for structural breaks in other large stock market indices.

- The start of the involvement of the USA in World War II (Pearl Harbor) on December 7, 1941,
- The Korean War, which started with the rather unexpected invasion of South Korea by North Korea on June 25, 1950,
- And finally the Gulf War, or more precisely, the invasion of Kuwait by Iraq on August 2, 1990.

Pearl Harbor can indeed be interpreted as a surprising conflict. Although US relations with Japan were deteriorating in 1941, the actual start of the war came as a surprise: A representative survey by Gallup conducted a few days before the attack on Pearl Harbor found that only 52% of Americans actually expected "that the United States will go to war against Japan some time in the near future" (Canadian Institute of Public Opinion [1941]). What happened was that, not "some time in the near future", but just a few days later, the war started – and not by the USA going to war, but by the USA being attacked. This is in stark contrast to the situation, e.g., in the Iraq War where the likelihood of a war was already estimated to be above 90% several weeks before the start of the war (see Figure A.1(a)).

In all of these cases, we cannot expect to find the same pattern which we observed before the outbreak of the wars analyzed in the previous section. There is simply no pre-war phase that could be analyzed and where an increase in the probability of war could decrease stock market prices.[12]

What we *can* observe, however, is whether the onset of a war led to a sudden increase or decrease in stock market prices. A structural break analysis shows that this is indeed the case for all three wars. However, the pattern is inverted: In all three cases, stock market prices went down significantly at the outbreak of the war. In the case of Pearl Harbor, the Dow Jones Index went down by nearly 3% in one day, in the case of the Korean War by nearly 5%, and in the case of the Gulf War by around 5.7% within a week.

### A.4.7 What's puzzling about the war puzzle

Our analysis so far has revealed quite a peculiar pattern: On the one hand, stock prices tend to fall when the probability of a war increases and tend to rise when there are signs for a peaceful resolution. However, the actual onset of the war will

[12] This holds true unless we studied intra-day data which in most of these cases is not available.

increase stock market prices.

While this by itself is already puzzling, it becomes even more puzzling considering the evidence that we found about surprising wars (wars without a lengthy prelude): Here, we observe that stock market prices tend to decrease once the war breaks out. In other words, whether stock prices decrease or increase when a war breaks out does not seem to depend much on the particular war but more on the previous history or, more precisely, whether the war was surprising or not. But even this observation cannot explain the discrepancy that one and the same war first suggests a pattern that an increase in the likelihood of war decreases stock market prices, but then ultimately, the onset of the war itself increases them. This cannot easily be accommodated with the idea that stock market prices reflect expectations of the future economic development of a country.

Table A.7 summarizes the stock markets' different reactions towards news for all conflicts studied.

| | Stock market reaction to: | | |
|---|---|---|---|
| | Increasing probability | Expected start | Surprising start |
| probability range (schematic) | 0-99% | 99-100% | 0-100% |
| World War II (start in Europe) | ↘ | ↗ | |
| World War II (Pearl Harbor) | | | ↘ |
| Korean War | | | ↘ |
| Vietnam War | ↘ | ↗ | |
| Gulf War (Kuwait invasion) | | | ↘ |
| Gulf War ("desert storm") | ↘ | ↗ | |
| Afghanistan War | ⋆ | ↗ | |
| Iraq War (2003) | ↘ | ↗ | |

Table A.7: Summary of stock markets' reactions to news. ⋆ indicates a lack of data in order to investigate this point.

## A.5 Attempts to explain the puzzle

### A.5.1 War as a stimulus package for the US economy

The first idea about this puzzle might be that it is none: A war can be considered as a stimulus package for the US economy and thus lead to an increase in stock market prices.

However, this line of argument does not work for two reasons: First, Section A.4.5 showed that the puzzle also seems to be present for countries which were not involved in the respective war, such as, for instance, Germany in the Iraq War. Second, the explanation fails to explain why *before* the war, prices fall whenever the war becomes more likely.

Falling prices when war is looming and rising prices when war is starting: Only the combination of these two observations makes the war puzzle a puzzle.

### A.5.2 Expectations about a quick end of the war

A natural idea is to consider the time dimension (Schneider and Troeger [2006]): Once the war breaks out, it is clear that the trouble will be over soon, and thus, investors buy stocks again. For example, the Saddam Security can be interpreted as the likelihood of a coming war, but also as expected length of the war (Amihud and Wohl [2004] and Wolfers and Zitzewitz [2009]). Chappell and Eldridge [2000] also suggested psychological explanations such as "despair" and "renewed hope" regarding the UK stock index reactions to the two sub-periods during World War II. There are, however, at least three arguments that lead us to discard this natural looking idea to explain our findings as well:

- Since the Saddam Security works well as a proxy for the estimated probability of war, it is clear that investors did not expect the war to take place in the distant future: Recall that the particular security that we studied would only pay out if Saddam Hussein was out of power by June – not long after the war indeed took place!
- While the positive expectation of a quick war that ends the costly and lengthy tensions before might be true, e.g., for the war in Iraq (although with hindsight it turned out to be overly optimistic), this cannot explain the pattern observed at the onset of World War II, given that it was generally not expected that this war would end quickly, given the experience of the four-year-long World War I.
- The expectation of a quick end to the war cannot explain why investors do not seem to appreciate the increase of the likelihood of a coming war earlier, as this brings the war (and hence its end) closer.

### A.5.3 Expectations about a devastating outcome of the war

For the Iraq War, another explanation for the fact that stock prices increased while the war took its course is that, in this case, everybody was afraid of a truly devastating outcome by means of weapons of mass destruction. As it turned out that this did not happen, stock prices eventually increased.

However, this reasoning is neither able to explain why stock prices increased already at the very start of the war – and not a bit later while Saddam Hussein still had time to use weapons of mass destruction if he had them – nor can it explain why the same pattern is observed in World War II, for instance, where this argument fails to be applicable.

### A.5.4 The role of the USA within each conflict

There exists a large body of literature on the structure of conflicts from a political perspective. The basic modeling idea is that an aggressor makes an initial surprising move towards his opponent, who then reacts as a defender. Both aggressor and defender can have allies to their aid.

One of the most influential publications within this literature is Schelling [1966]. Against the background of the potential use of nuclear weapons during the Cold War, he argues that the purpose of military strategy became intrinsically based on deterrence rather than ultimate victory over the opponent. Deterrence can thus be interpreted as bargaining power to coerce the opponent to behave in a particular way. While the author argues that this power is most successful when it is not actually performed, but rather used as a threat, he also notes that, in reality, this bargaining power is often conveyed by some performance of it. War is thus interpreted as a particular type of bargaining process. Other authors argue that war itself is a sign of the previous breakdown of the peaceful bargaining process (see Powell [2002]).

In this paper, we use a data set on US news items and stock prices. An obvious question to ask is thus which role the USA played in each of the conflicts analyzed and to what extent this may have an impact on our results.

While there is no obvious relation between a possible aim of deterrence on the part of the USA and the development of stock prices, a closer look at the overall position of the USA reveals an interesting pattern: The USA played an active role in the Vietnam War and the Gulf War (as ally to the defender) and in the Afghanistan War and the Iraq War (as main attacker). In all of these conflicts, stock prices increased at the onset of the war. At first sight, one may argue that it is the active role of the USA which influenced the positive development of stock market prices. However, all of these wars did also have a preceding period of increased tension which makes it difficult to disentangle US participation and pre-war phases as possible causes for the direction of stock market developments. This argument can be supported by looking at all unexpected conflicts, namely Pearl Harbor, the invasion of South Korea by North Korea, and the invasion of Kuwait by Iraq. The USA was the attacked country in the case of Pearl Harbor, and an ally to the defender in the Korean War as well as the Gulf War. In all three cases, the USA had not yet assumed an active role within the conflict when stock prices started to decrease. The same problem arises again: We cannot say whether it is the type of role that the USA assumed or the character of the crisis that influences the direction of stock price developments. However, there is one exception: When stock prices began to increase at the onset of World War II in September 1939, the USA had not yet assumed an active role, either. This can be interpreted as an indication toward the presence of the war puzzle, i.e., it is the structure of the conflict that matters for stock markets, and not the military position or the goals of the USA. But even if the driving force behind increasing stock prices at the onset of a war is the participation of the USA, this does not resolve the "war puzzle", but rather raises the new question of how the active role of the USA influences the stock market.

Also note that one might argue that the fact that catastrophic outcomes were unlikely for the USA during World War II contradicts the fact that Germany proceeded with comprehensive propaganda related to weapons of mass destruction. Their use did eventually not materialize. This seeming contradiction can be resolved by the interpretation of German propaganda as a non-plausible threat to the USA. It may also be regarded as a deterrence towards the USA (see Schelling [1966]). In particular, propaganda may have been used as a bargaining power to prevent the USA from intervening.

### A.5.5 Ambiguity-averse investors

An alternative explanation that seems to be natural is to assume that investors show ambiguity aversion. At first, the ambiguity about the probability of a war breaking out makes people shy away from the stock market and hence leads to lower stock prices. This effect stops as soon as it becomes clear that a war is indeed starting for sure and uncertainty is reduced (Schneider and Troeger [2006]). Guidolin and La Ferrara [2010] suggest that the initiation of conflict can been seen as a sign of resolve and investors tend to show positive reactions by buying stocks.

However, this seemingly natural behavioral explanation cannot accommodate our empirical data, as ambiguity cannot make an ambiguous situation seem worse than its worst possible outcome. In this situation, that would mean that, even though the probability of the start of a war is uncertain and hence ambiguous, an ambiguity-averse investor would still prefer this situation over a situation where the war has started for sure.

### A.5.6 Uncertainty about investment decisions

Another idea to resolve the puzzle is to consider the uncertainty about a war as an uncertainty about the decision between different portfolio allocations: In the case of a peaceful resolution, an investor should have invested in different assets than in the case of a war. While it is uncertain which of the two outcomes will occur, it might be optimal for an investor to stay out of the stock market entirely, or at least to reduce their holdings substantially.

To study this idea more in detail, we describe a small model.

When facing the possibility of the outbreak of a war, an investor has to decide between investing in two different portfolios, a "war portfolio" that will be successful if a war breaks out, and a "peace portfolio" that will be successful if the conflict ends in a peaceful resolution. We summarize the four potential outcomes in the following table[13]:

| | War breaks out (probability $p$) | Peaceful resolution (probability $1-p$) |
|---|---|---|
| War portfolio | $a$ | $-b$ |
| Peace portfolio | $-c$ | 1 |

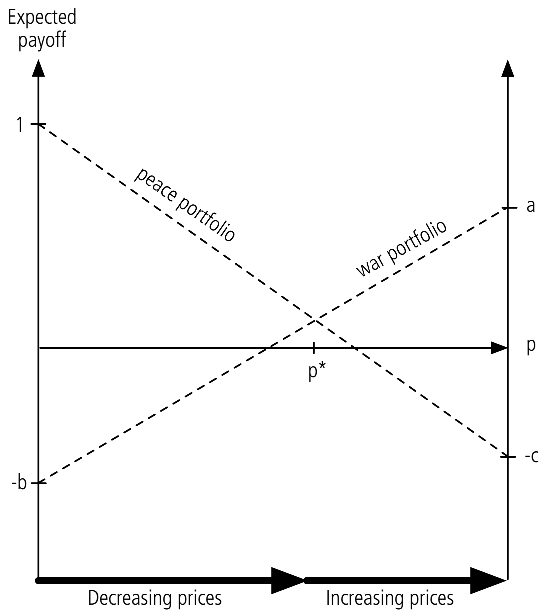



Figure A.7: Investors switching from a "peace portfolio" to a "war portfolio" when the probability of war increases beyond a threshold $p^*$ could explain the war puzzle.

Assuming that it is in any case bad for the investor to speculate on the wrong outcome, we have $a, b, c \geq 0$.

The expected payoff of the investor now depends on his portfolio decision. If he decides on the war portfolio, it is $ap - b(1-p)$. If he decides instead on the peace portfolio, it is $(1-p) - cp$ (see Figure A.7). Obviously, the war portfolio is better for large values of $p$ and worse for small values. As a short computation shows, the cut-off point is at

$$p^* = \frac{1+b}{1+a+b+c}.$$

This model indeed predicts the observed pattern: An increase in $p$ makes the

[13] For simplicity, we normalize the investor's return for a "peace portfolio" to one in the case of a peaceful resolution.

peace portfolio less and less attractive. Consequently, investors will sell the corresponding stocks. At the same time, however, the war portfolio is still even less attractive, thus the falling prices of the peace portfolio stocks cannot be matched by rising prices of stocks in the war portfolio. Once the probability of war reaches $p^*$, the pattern changes: Now, war stocks become more and more attractive, and thus, investors buy them and the overall market starts to rise.

In order to find a probability $p^*$ that is close to one and thus in line with our empirical evidence, $b$ needs to be sufficiently large, and thus, the potential loss when choosing a war portfolio, but encountering a peaceful resolution, should be fairly large.

Theoretically, we can explain the puzzle in this way, but how is it in reality? How can we test empirically whether this model indeed explains the war puzzle? The general idea is that different stocks should show a different pattern during the different phases before an outbreak of the war: Stocks that are in the "peace portfolio", i.e., stocks that would benefit from a peaceful resolution of the conflict, should initially be sold whenever the likelihood of a war increases. Thus, we would expect them to follow the generally observed pattern of the stock market, but only before the outbreak of the war: At that point, these stocks should not increase significantly, as there is no reason for investors to purchase them once it is clear that the "war portfolio" is the right investment.

Stocks in the "war portfolio" instead should follow the increase in the stock market at the start of the war, but not its previous decrease whenever the war became more likely.

A simple empirical test is now possible by looking at differences between sectors that should clearly be in the war portfolio (weapon-related industry) and sectors that should clearly be in the peace portfolio (e.g., travel-related industry). It turns out, however, that as convincing as the theoretical idea is, it thoroughly fails this simple but clear-cut test: In fact, whereas sectors that one would undoubtedly assign to the peace portfolio (e.g., travel stocks, like airlines) increased *most* at the outbreak of the war, stocks in the war portfolio (weapon manufacturers) did not.

To sum up, this explanation looks good in theory, but unfortunately does not help to resolve the puzzle in reality.

### A.5.7 Mean-variance preferences

Finally, there is a rather unexpected possible explanation for the phenomenon: Classical mean-variance preferences. Indeed, this concept could explain the observed data surprisingly well. One could argue that people do not buy when it is unsure what is going to happen due to their variance aversion, even though the expected return might still be a little bit better than when the war has finally started. In this sense, the observed phenomenon could be related to the mean-variance paradox.

To formalize this idea, we assign average expected returns $\mu_W$ and $\mu_P$ to the two potential outcomes (war and peace). We denote the probability of a war by $p$ and compute the variance of the two-outcome lottery ($\mu_W$ with probability $p$ and $\mu_P$ with probability $1-p$) as

$$\mathrm{var}(p) = (1-p)p(\mu_P - \mu_W)^2.$$

The variance is zero for $p = 0$ (i.e., peace is sure) or $p = 1$ (i.e., war is sure) and maximal for $p = 1/2$ (see Figure A.8).

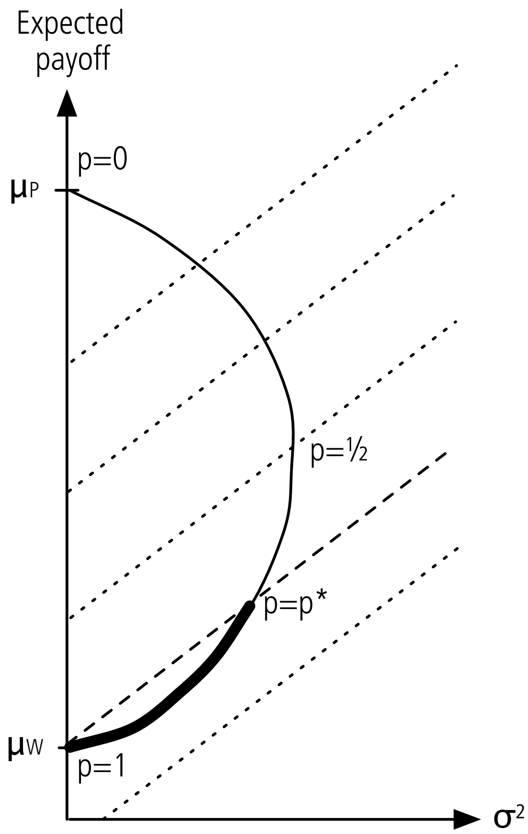



Figure A.8: Mean-variance preferences can explain the war puzzle. Here, $\mu_P$ and $\mu_W$ are the returns in case of peace and war, respectively. $\sigma$ is the variance of the situation, depending on the probability of war. It is maximal when the probability is 1/2 (maximal degree of uncertainty). The *dashed lines* are indifference curves, i.e., they mark combinations of mean and variance on which an investor would be indifferent. A certain war is preferred over an uncertain situation with a high probability of war (*points on the thick line*).

Typical indifference curves for a mean-variance investor with utility $\mu - \alpha\sigma^2$ are also shown in Figure A.8. Looking at the indifference curve through the point $(0, \mu_W)$ (i.e., a certain war), we see that this investor would prefer a certain war $(p = 1)$ over a situation where $p$ is large, but still smaller than one.

When the probability $p$ increases from zero to one, at first the situation worsens for the investor (who consequently would value stocks less), but after a certain point (in our model at $p^* = (\alpha - 1)/\alpha$, as a short computation shows), there is indeed a perceived improvement for the investor, explaining the increase in stock prices at the outbreak of a war.

## A.6 Conclusions

We have observed that stock market prices react very sensitively to the probability of the outbreak of a major war. The pattern that was shown for several wars during the last century demonstrates that an increase in the likelihood of war decreases stock prices and vice versa. What is puzzling, however, is that once a war breaks out, stock market prices do not decrease further, but do the very opposite and increase significantly. This was true for all wars with a more or less lengthy prologue. Wars that occur "out of the blue" show a different pattern in that their sudden outbreak tends to decrease stock market prices. These results are certainly a challenge for classical asset pricing models.

Mean-variance preferences of investors might be able to explain the observed pattern. There could, however, be different explanations for this puzzle based on other behavioral factors. Future research, e.g., with the help of laboratory experiments, might be able to shed more light on this puzzle.

### Acknowledgements

We thank our students Marco Lang, Laura Leoni, Raphael Müller, and Thomas Rast for their help with the data collection and analysis as well as Florian Herzog from swissQuant, Yakov Amihud, and Christian Ehm as well as anonymous referees for useful comments. Support by the National Centre of Competence in Research "Financial Valuation and Risk Management" (NCCR FINRISK), Project A1, "Behavioural and Evolutionary Finance," and by the University Priority Program "Finance and Financial Markets" of the University of Zurich is gratefully acknowledged.

# References

AMIHUD, Y., AND A. WOHL (2004): "Political news and stock prices: The case of Saddam Hussein contracts," *Journal of Banking & Finance*, 28(5), 1185–1200.

CANADIAN INSTITUTE OF PUBLIC OPINION (1941): "Gallup poll found 52 p.c. of Americans expected war," *The Evening Citizen*, 99(147), 1.

CHAPPELL, D., AND R. ELDRIDGE (2000): "Evidence of market inefficiency in a war environment," *Applied Financial Economics*, 10(5), 489–492.

COLLIER, P., AND A. HOEFFLER (1998): "On economic causes of civil war," *Oxford Economic Papers*, 50(4), 563–573.

CUTLER, D., J. POTERBA, AND L. SUMMERS (1998): "What moves stock prices?," *Journal of Portfolio Management*, 15(3), 4–11.

ELMENDORF, D., M. HIRSCHFELD, AND D. WEIL (1996): "The effect of news on bond prices: Evidence from the United Kingdom, 1900-1920," *Review of Economics and Statistics*, 78(2), 341–344.

FREY, B., AND M. KUCHER (2000): "History as reflected in capital markets: The case of World War II," *Journal of Economic History*, 60(2), 468–496.

GOLDSTEIN, J. (1992): "A conflict-cooperation scale for WEIS events data," *Journal of Conflict Resolution*, 36(2), 369–385.

GUIDOLIN, M., AND E. LA FERRARA (2010): "The economic effects of violent conflict: Evidence from asset market reactions," *Journal of Peace Research*, 47(6), 671–684.

HOLSTI, O., AND R. NORTH (1966): "Comparative data from content analysis: Perceptions of hostility and economic variables in the 1914 crisis in comparing nations," in *The use of quantitative data in cross-national research*, ed. by R. Merritt, and S. Rokkan, pp. 169–199. Yale University Press, New Haven.

KEYNES, J. (1919): *The economic consequences of the peace*. Macmillan, London, reprint edn. 1971.

LI, Q., AND D. SACKO (2002): "The (ir)relevance of militarized interstate disputes for international trade," *International Studies Quarterly*, 46(1), 11–43.

Niederhoffer, V. (1971): "The analysis of world events and stock prices," *Journal of Business*, 44(2), 193–219.

Powell, R. (2002): "Bargaining theory and international conflict," *Annual Review of Political Science*, 5(1), 1–30.

Rigobon, R., and B. Sack (2005): "The effects of war risk on US financial markets," *Journal of Banking & Finance*, 29(7), 1769–1789.

Russett, B., and E. Hanson (1975): *Interest and ideology: The foreign policy beliefs of American businessmen*. Freeman, San Francisco.

Schelling, T. (1966): *Arms and influence*. Yale University Press, New Haven.

Schneider, G., and V. Troeger (2006): "War and the world economy: Stock market reactions to international conflicts," *Journal of Conflict Resolution*, 50(5), 623–645.

Wolfers, J., and E. Zitzewitz (2009): "Using markets to inform policy: The case of the Iraq War," *Economica*, 76(302), 225–250.

# Part IV

# Curriculum vitae

# Curriculum vitae

Amelie Brune, born 5 September 1986 in Bad Pyrmont (Germany)

## Education

| | |
|---|---|
| 09/2011 – 09/2017 | Doctoral programme in Finance, University of Zurich, Switzerland |
| 09/2009 – 10/2011 | Master of Arts in Economics, University of Zurich, Switzerland *(magna cum laude)* |
| 07/2008 – 08/2008 | Certificate in Advanced Econometrics, London School of Economics and Political Science, United Kingdom *(grade A)* |
| 10/2006 – 10/2009 | Bachelor of Arts in Economics, University of Zurich, Switzerland *(magna cum laude)* |
| 06/2006 | German High School Diploma ("Abitur"), Viktoria-Luise-Gymnasium Hameln, Germany *(grade 1.0)* |

## Professional experience

| | |
|---|---|
| 01/2015 | Visiting Lecturer, ISM University of Management and Economics, Vilnius, Lithuania |
| 07/2010 – 08/2010 | Internship in Economic Research, Ifo Institute for Economic Research, University of Munich, Germany |
| 05/2008 – 10/2017 | Research and Teaching Assistant to Prof. Dr. Thorsten Hens, Department of Banking and Finance, University of Zurich, Switzerland |