THE BONN JOURNAL OF ECONOMICS

Volume II(2) December 2013

Department of Economics

University of Bonn

Legal Information

Publisher	The Bonn Journal of Economics Adenauerallee 24-42 53113 Bonn http://bje.uni-bonn.de bje@uni-bonn.de
V. i. S. d. P.	Prof. Dr. K. Sandmann Head of Department of Economics University of Bonn
Place of Publication	Bonn, Germany
ISSN	2195-3449

The Bonn Journal of Economics

Founded in 2012

VOLUME 2

ISSUE 2

Founded by

Johannes Hermle Justus Inhoffen Tobias Ruof

Department of Economics
University of Bonn

The Bonn Journal of Economics

Tax Compliance and Whistleblowing – The Role of Incentives

Ludger Breuer*

Introduction

With sovereign debt levels reaching record heights, governments face a growing pressure to effectively combat tax evasion. In recent years, tax authorities have turned their hands to a novel method for detecting tax evaders: the purchase of CDs with incriminating bank information about large-scale tax evasion. Most notably, German tax authorities have been avid for such CDs that were offered for sale by employees of Swiss, Liechtenstein, and Luxembourgian banks[1]. At present, legality, morality, and effectiveness of the actions taken by the German

* Ludger Breuer received his degree in Economics (Diplom, M. Sc. equivalent) from the University of Bonn in 2012. The present article refers to his diploma thesis submitted in August 2012

[1] In January 2006, a former employee of Liechtenstein's LGT Bank was the first to offer data on several hundred affluent tax evaders. The German Government paid roughly € 4.2 million for the CD and used the data to start a tax evasion probe in February 2008. The probe received much media attention, mainly because the well-known CEO of German postal giant Deutsche Post AG was detained during the first hours of the police raid. According to an interim result published by the public prosecutor's office in March 2010, 244 out of 596 preliminary proceedings have been completed, generating proceeds of € 181 million. Moreover, the case triggered also voluntary declarations of tax liabilities amounting to € 626 million, € 222 million thereof related to LGT Bank (Leyendecker, Hans 2010. "Liechtenstein-CD bringt 626 Millionen Euro." Süddeutsche Zeitung, April 7). In 2010, German authorities bought three other CDs with i.a. data on customers of Swiss banks Credit Suisse and Julius Baer. Another CD with data on customers of Luxembourgian banks was purchased in 2011. In July 2012, North Rhine-Westphalia confirmed that it had bought, yet again, CDs containing information about tax evasion by customers of Swiss banks. It is estimated that the acquisition of the above-mentioned CDs will generate proceeds of several billion Euro in total.

state are subject of a vigorous debate. The highly-charged atmosphere of the debate can be attributed to the fact that the purchase of "tax CDs" conflates two tricky subjects: tax evasion and whistleblowing. In this paper, we do not immerse into the broader issues of law and justice but focus on a preliminary question: Are monetary incentives for whistleblowing effective in fostering tax compliance? In particular, we ask the following three questions: How do monetary incentives for whistleblowing

I. influence the decision to report tax evasion?

II. influence taxpayers' decision to declare taxable income?

III. influence the revenues of the state?

We find that monetary rewards for whistleblowing lead to a significant increase in the reporting of tax evasion: the larger the reward the more pronounced the increase in whistleblowing. This result contradicts previous research reporting that external whistleblowing rewards may cause crowding-out effects. However, since some subjects do not follow the incentives, the results are also contrary to standard economic theory. Interestingly, this supposedly surprising whistleblowing behavior is, by and large, correctly anticipated by taxpayers. Thus, even without monetary incentives, whistleblowing proves to be fairly effective in curbing tax evasion. But incentivizing whistleblowing makes whistleblowing an even more powerful tool against tax evasion.

The remainder of the paper is organized as follows: Providing the background for our study, the following section discusses the complex phenomenon of whistleblowing. Then we will give a review of the related literature. After we have presented the design of the experiment, we will present our hypotheses. The results of the experiment will be reported in several sections and the last section concludes.

Whistleblowing – Loyalty and Betrayal

Though there is no generally accepted definition of whistleblowing, it is usually described as the deliberate disclosure of information about illegal, immoral, or illegitimate practices by current or former organization members to persons or organizations that may be able to effect action (Near and Miceli, 1985). The term ”whistleblower” became popular in the early 1970s since it allowed to avoid existing expressions (”informer”, ”snitch”, ”denouncer”) that all have more or less negative connotations. The struggle about words and definitions points to the profound ambiguity in the way whistleblowing is ethically perceived by the public. Essentially, this ambiguity can be attributed to the fact that social enforcement raises conflicts of loyalties which are often hard to resolve. On the one hand, it is sometimes suggested that all citizens have an ethical obligation to aid in the arrest and prosecution of criminal actors (Asbury, 2010). This notion corresponds to the ideas of loyalty to the state, loyalty to principles, or loyalty to actual or potential victims of the observed wrongdoing. However, the whistleblower also belongs to the group whose members are – at least partially – involved in the observed wrongdoing. Since the afore-mentioned loyalties collide with group loyalty, particularly the group loyalty to coworkers or employers, potential whistleblowers face a complex ethical dilemma (Bok, 1983; Hersh, 2002). In view of the intricate conflict of loyalties, it does not come as a surprise that whistleblowing is sometimes condemned as a treacherous breach of trust, while others glorify it as a heroic act of civic duty. The social acceptance of whistleblowing apparently depends on several factors. First, not every legal norm is backed by a corresponding social norm, and the intensity of underlying social norms may vary, too. Broadly speaking, the weaker the underlying social norm, the stronger the social rejection of reporting behavior. Second, whistleblowing always involves cutting some ties of loyalty, but the strength of these ties varies: The stronger

the social expectation that some form of trust should not be breached, the more pronounced the social rejection of whistleblowing (Koch, 2007). Third, the acceptance of whistleblowing hinges on the perceived motives of the whistleblower: The more selfish the motives, the lower the acceptance of whistleblowing. Considering the purchase of "tax CDs", social acceptance especially depends on the following questions:

I. How strong is the social norm that backs the legal prohibition of tax evasion?

II. To what degree are employees of banks expected to be loyal to their employers and to the customers of their banks?

III. Do financial rewards for whistleblowers reduce the social acceptance of whistleblowing?

Among the many strategies used to foster whistleblowing behavior[2], the offer of monetary rewards is particularly controversial. In most countries, it is a well-established practice for public authorities to offer monetary rewards for information that will help to solve crimes or enable the police to locate and apprehend criminals. In contrast, there are considerable differences between legal orders when it comes to rewards that transfer a part of the revenue collected by the state to the whistleblower.[3]

[2] Feldman and Lobel (2010) identify four main strategies: anti-retaliation protections, affirmative reporting duties, fines for the failure to report, and monetary rewards.

[3] Under current German law, no provisions allow for such a profit-sharing between public authorities and whistleblowers. But historically, some examples for such provisions can be found in German law. Moreover, other legal orders often resort to this instrument. Under US American law, employees of fraudulent government contractors can file qui tam suits on behalf of the government and receive a compensation of up to 30% of the recovery if they win the suit. Likewise, according to Section 7623(b) of the US American Internal Revenue Code, whistleblowers receive as an award at least 15% but not more than 30% of the collected proceeds. Another recent example is Dodd-Frank Act Section 922: under this provision, whistleblowers who provide the Securities and Exchange Commission (SEC) with original information about violations of the securities laws are entitled to rewards in the range of 10% to 30% of the monetary sanctions.

The mixed legal views on monetary rewards for whistleblowing correspond to an ongoing lively debate among behavioral economists about the relationship between intrinsic and extrinsic aspects of motivation. While extrinsic motivation "refers to the performance of an activity in order to attain some separable outcome" (Ryan and Deci, 2000), such as monetary rewards, intrinsic motivation means that people behave in a certain way because of the inherent satisfaction of the activity itself (Ryan and Deci, 2000). Some studies suggest that, for a range of activities, the introduction of monetary rewards can undermine intrinsic motivation (Frey and Jegen, 2011). This finding runs counter to classic economic predictions. However, monetary rewards do not always cause such "crowding-out effects" (for a meta-study see: Deci, Koestner, and Ryan, 1999, and sometimes outside intervention in form of rewards can even increase intrinsic motivation Osterloh, Frost, and Frey, 2002).

Hence, different attitudes towards whistleblower rewards in various legal orders may not only reflect different moral concepts, but also point to uncertainty about the effects of such rewards. The present study aims to provide valuable insight into the impact of whistleblower rewards on reporting of illegal activities, such as tax evasion. Thus, it contributes to the existing experimental literature on the interplay between intrinsic and extrinsic motivation.[4]

[4] Rewards for whistleblowers can be viewed through the lens of expressive law, too. Once monetary incentives for whistleblowing are implemented through legislation, the legal provisions may send out a signal about moral values. In recent years, legal scholars put forward the idea that individuals do not only react to rewards and sanctions laid down in the law, but also respond to expressive signals embodied within our legal system. Monetary rewards for cooperation may not just offer material gains but also signal that whistleblowing is a socially desired behavior. If this signal is able to shape people's attitude towards whistleblowing in that direction, monetary rewards and expressive signals may mutually reinforce one another. However, the signal can also backfire. For example, people may get the impression that blowing the whistle is generally considered an abhorrent breach of loyalty, since substantial monetary rewards are needed to induce at least some level of reporting. Thus, in the case of whistleblowing rewards, it is entirely unclear how the expressive function of law will work. The present study aims to shed some light on this topic, too.

Related Literature

The related experimental research is basically divided in two pillars: tax evasion experiments and whistleblowing experiments. To our knowledge, so far there are no laboratory experiments that directly combine both types of experiments.

Beginning with the first study of Friedland, Maital, and Rutenberg (1978), most tax evasion experiments rely on the standard model of tax evasion first proposed by Allingham and Sandmo (1972) and Yitzhaki (1974), in which the taxpayer acts as an isolated expected utility maximizer facing a gamble with detection probability and punishment. The most important finding of the early tax evasion experiments is that the level of income reporting is much higher than the expected utility model would predict (Torgler, 2002). In order to unravel the puzzle of tax compliance, some experiments investigate the effects of various additional aspects of tax administration. On the other hand, a growing number of experiments examine how various societal institutions and processes influence tax compliance. Basically, our experiment belongs to this second branch of recent tax evasion experiments.

In particular, audit rules can be designed so that a subject's chance of audit depends upon the behavior of other subjects. Alm and McKee (2004) conducted an experiment to test an endogenous audit rule based upon how far a taxpayer's tax report drops below the average of tax reports in the relevant cohort. Since audit selection is based upon relative reporting behavior, there is a coordination game among taxpayers with multiple equilibria. Those equilibria that involve low reporting are clearly preferable for the group, but they are also vulnerable to defection. The results presented by Alm and McKee show that subjects are unable to coordinate on the low compliance equilibrium if they are not permitted to discuss their strategies among themselves prior to reporting their income

("cheap talk").[5]

While tax evasion experiments flourished during the last decades, only recently economists have begun to include whistleblowing options in at least two well established types of laboratory experiments: antitrust experiments and experimental research on corruption.

A fast-growing body of antitrust experiments examines the effects of leniency programs. Leniency in antitrust enforcement can be defined as the reduction of penalties for firms or individuals that first confess to involvement in a cartel. The experimental research on leniency in antitrust enforcement has provided ambiguous results. If a leniency program is in place, subjects make more often use of the whistleblowing option and, thus, more cartels are detected (Bigoni, Fridolfsson, Le Coq, and Spagnolo, 2011). But often, cartel formation, cartel stability and market prices are not reduced, and the prices charged by cartel members may even increase (Apesteguia, Dufwenberg, and Selten, 2007; Bigoni, Fridolfsson, Le Coq, and Spagnolo, 2011; Dijkstra, Haan, and Schoonbeek, 2011). The most prominent explanation argues that the threat of whistleblowing deters cartel members from underbidding the agreed upon prices. This has motivated the investigation of whistleblower rewards. The theoretical prediction (Spagnolo, 2004; Aubert, Kovacic, and Rey, 2006) that these rewards would strengthen the effectiveness of leniency rules has received mixed support in recent experimental studies. Whistleblowing rewards did not perform well in a one-shot game (Apesteguia, Dufwenberg, and Selten, 2007), but, in a repeated game, they turned out to be the only welfare enhancing leniency scheme (Bigoni, Fridolfsson, Le Coq, and

[5] To some extent, this experiment neighbors on our study, since both whistleblowing-triggered audits and the endogenous audit rule described above can lead to coordination problems. But there are major differences. In particular, subjects facing an endogenous audit rule have to guess the tax compliance behavior of other subjects, whereas subjects facing a whistleblowing-triggered audit have to guess the whistleblowing behavior of other subjects. Since the social norms of tax compliance may substantially differ from the social norms of whistleblowing, the results of the experiments are not directly comparable.

Spagnolo, 2011).

Corruption experiments have recently been combined with whistleblowing mechanisms, too. Schikora (2011) and Engel, Goerg, and Yu (2012) investigate whether whistleblowing mechanisms with symmetric or asymmetric punishment are more effective in curbing corruption. Both studies find that whistleblowing mechanisms may serve as a threat that forces officials to reciprocate, thus stabilizing the corrupt relationship.[6]

However, there is a substantial difference between the whistleblowing behavior of cartel members, bribers or corrupt officials on the one hand and whistleblowing in the context of taxation on the other hand. The first type of whistleblowing serves to break up conspiracies from the inside and may be called "traitorous whistleblowing". The second type of whistleblowing covers the reporting of wrongdoing by a third person that observed the wrongdoing but did not take part in it. It seems fitting to label this as "watchdog whistleblowing".[7] Since the role of incentives for watchdog whistleblowing has not yet been investigated in laboratory experiments, our study adds a novel facet to the existing experimental research on whistleblowing.

This facet has recently received some attention in the theoretical literature and in vignette studies. Yaniv (2001) was the first to model the individual's decision to blow the whistle on tax evaders. But since Yaniv takes tax evasion as given, his model is unable to provide insight into the impact of whistleblowing

[6] In particular, punishing the briber less harshly than the bribee leads to more corruption (Engel, Goerg, and Yu, 2012). On the other hand, asymmetric leniency programs for officials who blow the whistle significantly reduce the level of corruption, since the official is protected from retaliation (Schikora, 2011).

[7] Traitorous whistleblowing is tantamount to self-indictment causing punishment for the whistleblower himself. However, the watchdog whistleblower can abide by the law and still be able to observe and report the wrongdoing of others. Thus, the problem that, without special mechanisms like leniency, whistleblowers are automatically punished for their own wrongdoing does not arise. Nevertheless, whistleblowers might face some risk of retaliation so that whistleblowing remains a costly activity. It should also be noted that due to these differences between the two types of whistleblowing the relevant social norms might vary, too.

on tax compliance. To our knowledge, so far, only Mealem, Tobol, and Yaniv (2010) model the taxpayer's decision to evade taxes under the threat of whistleblowing and the tax authority's optimal design of audit policy in the presence of whistleblowers. The authors' main findings can be summarized as follows: If audit costs are too high to audit so many taxpayers that everybody prefers to report his income truthfully, the tax authority might be better off running a second, whistleblowing-triggered audit round although all reports about tax evasion will stem from dishonest whistleblowers who falsely denounce taxpayers to harass them.[8]

A large part of the research on the characteristics of whistleblowers and the contextual variables of whistleblowing uses scenario-based approaches to ascertain when an observer of wrongdoing will be likely to blow the whistle. Most notably, Feldman and Lobel (2010) use a large vignette study to compare the effect of different regulatory mechanisms on individual whistleblowing motivation and behavior. The study produced several intriguing results. First, participants predict that they themselves would be more likely to report than others ("holier than thou effect"). Second, participants who have a low perception of misconduct severity and, thus, a low internal motivation to report, are less willing to blow the whistle if they are offered low monetary rewards compared to legal mechanisms that do not offer any whistleblowing rewards. Third, compared to the low reward scenario, participants with low internal motivation are much more likely to report if they are offered high monetary rewards. Forth, the reporting behavior of participants with high internal motivation is very similar across the various

[8] In the second round the tax authority threatens to audit a sufficiently high fraction of denounced taxpayers so that all taxpayers who fear a whistleblower refrain from evading taxes. Hence, no honest whistleblower will report a taxpayer. The results derived from the model are somewhat counterintuitive. Since the motivations that underpin the behavior of whistleblowers are very complex, it is questionable to focus exclusively on the desire to take revenge. In terms of external validity, one may also question the complete lack of penalties for false denounciations. Therefore, our experiment is not primarily designed to test the model developed by Mealem, Tobol, and Yaniv (2010).

legal mechanisms used to incentivize whistleblowing.[9] The authors conclude that these findings "indicate that often offering monetary rewards to whistleblowers will lead to less, rather than more, reporting of illegality" and hint at "a type of crowding effect in which the introduction of an external reward interferes with the moral dimension of reporting". However, fifth, participants think that the size of the whistleblowing reward is far more influential when it comes to others in comparison to themselves. Considering the general limitations of vignette studies, the last finding casts some doubts on the alleged crowding out effect. Therefore, we feel it desirable to experimentally test the results presented by Feldman and Lobel (2010).

The Experiment

The experiment was programmed and implemented using z-Tree (Fischbacher, 2007) version 3.3.11 and is subdivided into four treatments. We shall refer to these treatments as "BASE", "NO INC", "SM INC" and "LA INC". The experiment consisted of two sessions for each of the four treatments. All eight sessions were conducted at the Laboratory for Experimental Economics, at the University of Bonn, Germany. We had six sessions with 24 subjects and two sessions with 21 subjects. Hence, a total of 186 subjects, mostly students from various undergraduate courses at the University of Bonn, including law, economics, political science, etc., took part in the experiment. When participants arrived at the laboratory they were seated in a lecture room where the experimenter thanked them for coming and informed them that they take part in an experiment that allows them to earn money. The participants were informed that the money earned in the experiment would be paid anonymously at the end of the experiment, so that no participant could learn about the money earned by other participants. They

[9] The four legal mechanisms were: anti-retaliation protections, affirmative reporting duties, fines for the failure to report, and monetary rewards.

were also told not to communicate with fellow participants. Finally, the participants were randomly assigned to visually isolated cubicles equipped with computer terminals. The cubicles were numbered from 1 to 24. Each session lasted less than two hours including the time for payments. Earnings were recorded in an experimental currency named "Taler". Talers were convertible to Euros at the rate of 7000 Talers per Euro. Average earnings in the experiment were 10.75 Euro.

Each session comprises two parts. The first part is once more divided into two stages.

At the beginning of the experiment, the participants receive only the instructions for the first stage of the first part, a real effort task based on sliders (Gill and Prowse, 2011a,b). This stage consists of three rounds. The first round is a practice round and the two remaining rounds are paying rounds. In every round, the slider task consists of a single screen displaying 48 sliders. The screen does not vary across subjects or across rounds. Initially, all 48 sliders are positioned at 0. The subjects can use the mouse to position each slider at any integer location between 0 and 100 inclusive. The current position of each slider is displayed to the right of the slider. The subjects' task is to position as many sliders as possible exactly at 50 within the allotted time of 120 seconds. For every slider that is correctly positioned at 50 at the end of the allotted time, the subjects receive 1000 points. As the task proceeds, the subject's current points score is displayed above the 48 sliders and the remaining time is displayed in the right upper corner of the screen. At the end of the first stage, for each subject, the points scored in the first and the second paying round are added up to a total score. The subjects receive a payoff of 1 Taler for every point of their total score. Additionally, each subject gets a fixed payoff of 4000 Taler.

After the first stage is completed, the instructions for the second stage of the first

part are handed out to the participants. At the beginning of the second stage, the subjects receive the payoff that they have earned in the previous stage and are asked to "pay taxes" on their income. For all subjects, the tax rate is set at 30% and the taxes paid by the subjects are donated to a charity with the DZI donation seal.[10] This is common knowledge, but the name of the charity is only revealed to the participants at the end of the experiment. The subject must decide how much income to report, and must pay taxes on all reported income. The subject pays no taxes on underreported income. To simplify decision-making, the subject can choose the percentage of income that he wants to report with the help of a slider on the computer screen. This slider can be positioned at any integer location between 0 and 100 inclusive. For each possible position of the slider, the corresponding percentage of income reported, the value of income reported (in Taler), the resulting tax payment (in Taler), and the remaining income after taxes (in Taler) are displayed to the subject. Before the subject decides how much income to report, he is informed about the following steps.

The subject is told that his "tax declaration" will be randomly audited with a probability of $\frac{1}{24}$. To make the randomness of the audit more salient, an old fashioned hand turning bingo cage containing 24 balls numbered from 1 to 24 is put up in the middle of the room and the subjects are allowed to watch the drawing. Once the balls are mixed, one ball is drawn from the bingo cage, and the number on the ball is displayed to all subjects. Afterwards, the subject that is seated in the cubicle labeled with the number on the drawn ball is audited.[11] Based on a comparison between the actual and the reported income of the subject, the audit entails the payment of any unpaid taxes and the additional payment of a "fine" if the reported income is less than the actual income. The fine is fixed

[10] In all eight sessions, the taxes paid by the subjects were donated to UNICEF.

[11] If less than 24 subjects take part in one session, and the ball drawn from the bingo cage is labeled with one of those numbers that belong to the unoccupied cubicles, no subject is selected for audit. This was made clear to all subjects taking part in sessions with less than 24 subjects.

at 200% of unpaid taxes. The paid tax arrears are also donated to the charity whereas the fine is not donated. To facilitate the decision-making, a small table depicting the subject's payoff depending on the percentage of income reported and the occurrence of an audit is presented to the subject. In "BASE", stage 2 of part 1 ends after the random audit. In the other treatments, stage 2 of part 1 continues with the third step.

Once the random audit has been conducted, every subject observes the tax declaration of one other subject. Accordingly, the tax declaration of every subject is observed by exactly one other subject. It is common knowledge that two subjects cannot mutually observe each other. In order to have as many independent observations as possible, in every session, there were eight random matching groups consisting of three subjects that observe each other in circles: A observes B, B observes C, and C observes A.[12]

Since it is well possible that many subjects will declare the same percentage of their income, for many other tax declaration values there may not be enough observations. To circumvent this problem, we adopt the strategy method (Selten 1967). The subject is presented with six different tax declarations: five fictive tax declarations and the actual tax declaration of the observed subject. The subject is told that five of the six tax declarations are fictive but he cannot figure out which of the six tax declaration is the actual tax declaration. For all six tax declarations, the subject has to decide whether or not to report the tax declaration. In the further course of the experiment, only the decision to or not to report the actual tax declaration leads to consequences for the subjects.

If the actual tax declaration is reported, the reported tax declaration is audited. If the declared income is less than the actual income (i.e. the percentage of income reported ranges between 0% and 99%), the audited subject must pay any

[12] The subjects were not explicitly informed about these random matching groups. The participants did not ask any questions with regard to the exact observation pattern.

unpaid taxes and an additional fine fixed at 200% of unpaid taxes. Again, the fine is not donated but the paid tax arrears are donated to the charity. If the declared income equals the actual income (i.e. the percentage of income reported amounts to 100%), the audit does not entail consequences for the audited subject. If a subject that did not fully declare his income has already been randomly audited, the actual tax declaration observed by the other subject is set at 100% and, thus, a possible second audit would not entail further consequences for the audited subject.

Reporting the actual tax declaration costs the whistleblower 2000 Taler.[13] In "NO INC", the whistleblower is not entitled to any financial rewards. In contrast, "SM INC" provides an exiguous financial incentive for reporting: The whistleblower gets a bonus of 2200 Taler if the audit leads to the detection of some level of tax evasion, i.e. if the actual tax declaration ranges between 0% and 99%. In "LA INC", the bonus for whistleblowing is raised to 12000 Taler, thus strengthening the incentives for reporting. The conditions for the award of the bonus do not vary between the treatments. To prevent the dissemination of information about the actual tax compliance of other subjects, the subjects that report the tax declaration in at least one of the six decision situations learn only at the end of the experiment (i.e. after part 2) whether they are eligible to get the whistleblowing bonus. Stage 2 of part 1 ends after the whistleblowing procedure and the experiment continues with part 2. It is common knowledge that the decisions taken in part 2 cannot influence the subject's payoff from the first part of the experiment. After the first part of the experiment is completed, the instructions for the second part are handed out to the participants. The second part is based on a "surprise restart", i.e. all stages and steps of the first part are repeated without any substantial changes to the rules of the game. We only omitted the practice round of

[13] Since only the decision to or not to report the actual tax declaration leads to consequences for the subjects, it is not costly to report one of the five fictive tax declarations.

the slider task since the subjects were already able to sufficiently practice the task during the first part of the experiment. The matching groups remain unchanged and the subjects are explicitly told that the observation mechanism stays the same.

Hypotheses

Standard economic theory is known for poorly predicting the behavior of subjects taking part in tax evasion experiments. Hence in order to formulate our hypotheses, we mainly use the results provided by a vignette study of Feldman and Lobel (2010) to predict subjects' tax compliance and whistleblowing behavior. According to this study, some subjects predict that they would likely report wrongdoing even if there are no whistleblowing rewards and no fines for failure to report. This behavior can be attributed to intrinsic motivation. Some antitrust experiments using one-shot games also find whistleblowing in treatments without incentives for reporting. Given this evidence, we formulate our first hypothesis about subject's whistleblowing behavior:

Hypothesis 1. *In "NO INC", some subjects will blow the whistle on tax evaders.*

With regard to subjects with low internal motivation, Feldman and Lobel (2010) find that low rewards for whistleblowing reduce the willingness to report compared to situations where no incentive is present. This reaction to the introduction of external rewards may be interpreted as a crowding-out effect. Along these lines, we formulate our second hypothesis:

Hypothesis 2. *In "SM INC", the probability that subjects blow the whistle on tax evaders will be smaller than in "NO INC".*

However, the results presented by Feldman and Lobel (2010) suggest that high whistleblowing rewards are able to overcome the crowding-out effect. Thus, we

expect that an increase in whistleblowing rewards leads to more reporting:

Hypothesis 3. *In "LA INC", the probability that subjects blow the whistle on tax evaders will be higher than in "SM INC".*

The following predictions concern the tax evasion behavior. Since the tax evasion behavior in "BASE" does not depend on any whistleblowing mechanism, we derive our forth hypothesis from the numerous tax evasion experiments that have been used to test the standard model of tax evasion. In line with these experiments, we expect to observe a substantial amount of tax evasion, but we also assume that the level of income reporting will be higher than the expected utility model would predict for risk-neutral or moderately risk-averse subjects:

Hypothesis 4. *In "BASE", some subjects will declare more than 0% of their income, but only few subjects will fully declare their income.*

With regard to the three other treatments, we assume that tax evasion behavior depends on what subjects expect other subjects to do under the various incentive schemes for whistleblowing. Data from vignette studies demonstrates that the perceived effect of whistleblowing mechanisms differs from the effect that subjects report for their own actions. In general, individuals predict that others are less likely to blow the whistle than they themselves are (Feldman and Lobel, 2010). Despite "this holier than thou" effect, subjects predict that others will to some degree report wrongdoing even if there are no whistleblowing rewards and no fines for failure to report. Considering these results, we expect subjects to believe that some other subjects with high intrinsic motivation will make use of costly whistleblowing options even if there are no monetary rewards. Put another way, we assume that, in "NO INC", subjects expect a higher audit probability than in "BASE". Thus, we formulate our fifth hypothesis as follows:

Hypothesis 5. *In "NO INC", subjects will declare a larger share of their income*

compared to "BASE".

Feldman and Lobel (2010) also report that subjects believe that the average person's whistleblowing behavior is much more externally motivated than their own actions. Therefore, we expect that subjects do not to anticipate the crowding-out effect described in hypothesis 2 but rather believe that the introduction of small rewards does not have a significant impact on whistleblowing behavior. This expectation results in our sixth hypothesis:

Hypothesis 6. *In "SM INC", subjects will declare a share of income that does not significantly differ from the share of income declared in "NO INC".*

Moreover, following the above-mentioned results, we expect that subjects believe that the introduction of large rewards will substantially increase the willingness of others to blow the whistle on tax evaders. Hence we assume that, in "LA INC", subjects expect a higher audit probability than in all other treatments. As a result, we formulate our seventh hypothesis as follows:

Hypothesis 7. *In "LA INC", subjects will declare a larger share of their income compared to "NO INC" and "SM INC".*

Experimental Results

We present our results in three parts. At first we start with the investigation of whistleblowing behavior under the three different incentive schemes. After that we will turn to the differences in tax compliance between the four treatments. Finally, we take a closer look at the efficiency of whistleblowing incentives by evaluating the differences in tax payments and state revenues between the treatments.

Experimental Results – Whistleblowing

Figure 1 gives the frequencies of whistleblowing over all decisions for the two rounds in each of the three treatments that include a whistleblowing mechanism. The whistleblowing frequencies are calculated as follows: In the whistleblowing stage each subject observes six different tax declarations, one of them always being 100%. Since reporting an honest taxpayer is obviously not a sound decision[14], we excluded reactions to observed full tax compliance from the calculation.[15]

Whistleblowing in the Various Treatments

We start with a quick look at the whistleblowing behavior in "NO INC". In the first round, in 27.1% of all cases, subjects decided to blow the whistle on tax evaders. In the second round, the whistleblowing frequency slightly rises to 31.6%, but this increase is not significant ($p = 0.269$, two-sided Wilcoxon signed-rank test).

The introduction of small rewards changes the picture significantly. In the first round of "SM INC", subjects blew the whistle on tax evaders in 48.4% of all cases. This whistleblowing frequency is significantly higher than the whistleblowing frequency of 27.1% observed in the first round of "NO INC" ($p = 0.013$, two sided Mann-Whitney u-test). For the second round of "SM INC", whistleblowing frequency even amounts to 62.2%, again resulting in a significant increase compared to the whistleblowing frequency of 31.6% in the second round of "NO INC" ($p = 0.001$, two sided Mann-Whitney u-test). It is also worth noting that the whistleblowing frequency in the second round of "SM INC" is significantly higher than

[14] Reporting an honest taxpayer does not entail monetary consequences for the audited taxpayer, but for the reporting subject it is a costly activity that does not yield any monetary profits. Moreover, it seems implausible that subjects may be intrinsically motivated to report honest taxpayers. The inclusion of the "100%"-scenario mainly allows to control whether subjects understood the rules of the experiment.

[15] With regard to the remaining five observed tax declarations, for every subject the number of decisions to blow the whistle is divided by the number of observed tax declarations, i.e. 5.

the whistleblowing frequency in the first round of "SM INC" ($p = 0.017$, two-sided Wilcoxon signed-rank test).

Increasing the size of rewards has a significant impact on whistleblowing behavior, too. In "LA INC", subjects were even more inclined to blow the whistle on tax evaders. In the first round, tax evasion was reported in 79.2% of all cases, and, in the second round, subjects even decided to blow the whistle in 85.8% of all cases. The increase from the first to the second round is only weakly significant ($p = 0.054$, two-sided Wilcoxon signed-rank test), but in both rounds the whistle was blown much more often than in the other treatments: In the first round of "LA INC", the whistleblowing frequency is significantly higher than in the first round of "NO INC" ($p < 0.001$, two sided Mann-Whitney u-test) and in the first round of "SM INC" ($p < 0.001$, two sided Mann-Whitney u-test). Likewise, in the second round of "LA INC", tax evasion was reported significantly more often than in the second round of "NO INC" ($p < 0.001$, two sided Mann-Whitney u-test) and in the second round of "SM INC" ($p = 0.001$, two sided Mann-Whitney u-test).

Explaining Whistleblowing Behavior

In this subsection, we investigate the effects of monetary rewards for whistleblowing on the willingness to report tax evaders in greater detail, i.e. we account for individual characteristics when comparing treatments. The variable that we seek to explain in the analysis is the whistleblowing frequency for each of the two rounds. Whistleblowing frequency is a left- and right-censored dependent variable: Whistleblowing frequency cannot be smaller than 0 and it cannot exceed 1. Therefore, we run a Tobit regression that censors the observations at both minimum and maximum whistleblowing frequency.

Table 1 gives the results of the Tobit regression models used to explain whistleblowing frequency in the first round. Model 1 of Table 1 basically repeats the nonparametric tests. The difference between "NO INC" and the other treatments is striking: Both "SM INC" and "LA INC" have a significantly positive effect ("SM INC": $p < 0.025$; "LA INC": $p < 0.01$) on the whistleblowing frequency. The picture does not change if we use subjects' effort in round 1 as an additional explanatory variable (Model 2 of Table 1). Because subjects earn 1000 Taler for every correctly positioned slider, this additional variable does not just show how much effort subjects put in[16], but it also depicts the income level of subjects. Thus, no statistically significant dependence of whistleblowing frequency on either effort or income level was detected. Adding subjects' attitude towards charities in general (with higher values indicating a more positive attitude) as an additional explanatory variable (Model 3 of Table 1) yields a similar result. The coefficient of attitude towards charities is positive but not statistically significant. Finally, Model 4 of Table 1 extends the explanatory variables by including age and gender (gender is set to 1 for females) of subjects. Again, the coefficients for both variables are not statistically significant. In contrast, the coefficients for the treatments remain virtually unchanged compared to Model 1, Model 2 and Model 3: In all four models they are positive and statistically significant ("SM INC": $p < 0.025$; "LA INC": $p < 0.01$).

The results of the Tobit regression models used to explain whistleblowing frequency in the second round are shown in Table 2. The overall picture is virtually unchanged compared to the results for the first round: In all four models, both "SM INC" and "LA INC" have a significantly positive effect ($p < 0.01$) on the whistleblowing frequency, while the coefficients of all further explanatory vari-

[16] The variable "Effort (round1)" shows the number of correctly positioned sliders plus 4 (taking into account that subjects received a fixed payoff equal to the earnings for four correctly positioned sliders). Thus, strictly speaking, the variable depicts the income level rather than the effort level.

ables are not statistically significant.

Summing up – Results on Whistleblowing

Given the results of the Mann-Whitney u-tests and the regression analyses, we do not reject Hypothesis 1 and Hypothesis 3, but we reject Hypothesis 2:

Result 1. *Even if, as in "NO INC", whistleblowing is costly and no monetary incentives for whistleblowing are provided, some subjects still blow the whistle on tax evaders.*

Result 2. *If, as in "SM INC", small monetary incentives for whistleblowing are introduced, subjects decide significantly more often to blow the whistle on tax evaders.*

Result 3. *If, as in "LA INC", substantial monetary incentives for whistleblowing are provided, the whistleblowing frequency is significantly higher compared to treatments where only small rewards or even no rewards are provided.*

Summing up, monetary rewards lead to a significant increase in whistleblowing frequency, and the larger the reward the more pronounced the increase in whistleblowing and the resulting detection probability of tax evasion. These findings do not lend support to the hypothesized crowding-out effect of external rewards. The results also run counter to predictions based on standard economic theory: Though, in general, incentives for whistleblowing do not backfire, there are some subjects who report tax evasion even if whistleblowing is costly, while other subjects refuse to blow the whistle even if substantial rewards are provided. Thus, the whistleblowing behavior does not go from one extreme to the other but rather steadily shifts, thereby cushioning the effects of incentives.

Experimental Results – Tax Compliance

Figure 2 gives the arithmetic mean of tax declarations for the two rounds in each of the four treatments. Figure 3 provides a supplementary box plot diagram for tax declaration.[17] The term "tax declaration" is here used to designate the percentage of income reported.[18]

Tax Compliance in the Various Treatments

We observe a substantial amount of tax evasion in "BASE". In the first round subjects declare on average only 36.38% of their income, and in the second round tax declaration even declines to a meager 25.21% of income, significantly less than in the first round ($p = 0.039$, two-sided Wilcoxon signed-rank test). Only few subjects fully declare their income (16.67% in round 1, and 6.25% in round 2).

When a non-incentivized whistleblowing mechanism is introduced, a completely different picture emerges. In the first round of "NO INC", tax declaration amounts to 69.82% of income. This value significantly drops to 59.62% in the second round ($p = 0.042$, two-sided Wilcoxon signed-rank test). A comparison between treatments shows that, in both rounds, tax declaration in "NO INC" is significantly higher than in "BASE" (for round 1: $p < 0.001$, two sided Mann-Whitney u-test; for round 2: $p < 0.001$, two sided Mann-Whitney u-test).

Compared to "NO INC", introducing small whistleblowing rewards does not have a major impact on tax declaration. In "SM INC", tax declaration amounts to 68.49% in the first round, and 73.47% in the second round. The small increase from the first to the second round is not statistically significant ($p = 0.271$, two-

[17] The boxplot also includes subjects' expectations about average tax declaration of all subjects in the respective treatment.

[18] Since a flat tax is applied, the percentage of income reported is naturally equal to the share of taxes due voluntarily paid by a subject.

sided Wilcoxon signed-rank test). There is also no significant difference in tax declaration between "NO INC" and "SM INC" (for round 1: $p = 0.99$, two sided Mann-Whitney u-test; for round 2: $p = 0.109$, two sided Mann-Whitney u-test). Consequentially, we find that tax declaration in "SM INC" is significantly higher than in "BASE" (for round 1: $p < 0.001$, two sided Mann-Whitney u-test; for round 2: $p < 0.001$, two sided Mann-Whitney u-test).

However, providing large whistleblowing rewards strongly affects tax declaration. In the first round of "LA INC", subjects declare on average 85.67% of their income, and in the second round tax declaration even increases to 90.21%. The increase in tax declaration from round 1 to round 2 is only weakly significant ($p = 0.086$, two-sided Wilcoxon signed-rank test), but in both rounds tax declaration is much higher than in the other treatments: In the first round of "LA INC", tax declaration is significantly higher than in the first round of "BASE" ($p < 0.001$, two sided Mann-Whitney u-test), "NO INC" ($p = 0.006$, two sided Mann-Whitney u-test), and "SM INC" ($p = 0.02$, two sided Mann-Whitney u-test). Likewise, in the second round of "LA INC", subjects declared a significantly higher share of income than in the second round of "BASE" ($p < 0.001$, two sided Mann-Whitney u-test), "NO INC" ($p < 0.001$, two sided Mann-Whitney u-test), and "SM INC' ($p = 0.007$, two sided Mann-Whitney u-test).

Explaining Tax Compliance in Round 1

In order to explore the effects of whistleblowing rewards on tax compliance more thoroughly, the following analysis takes account of individual characteristics when comparing treatments. The variable that we seek to explain is the tax declaration for round 1. Since tax declaration is a left- and right-censored dependent variable[19], we run a Tobit regression.

[19] Tax declaration is expressed as a percentage of income declared and, thus, ranges from 0 to 100.

Table 3 shows the results of the Tobit regression models used to explain tax compliance behavior in the first round. In line with the results of the nonparametric tests, Model 1 of Table 3 reveals clear-cut treatment effects. Compared to "BASE", all treatments with whistleblowing mechanisms have a significantly positive effect ($p < 0.01$) on tax declaration. Model 2 of Table 5 adds subjects' effort – and thereby subjects' income – as explanatory variable. The coefficient of effort and income on tax declaration is negative but not statistically significant. Model 3 of Table 5 also uses subjects' attitude towards charities in general as an explanatory variable. The attitude towards charities has a significantly positive effect ($p < 0.025$) on tax declaration, i.e. subjects who have a positive view of charities declare a higher percentage of their income. However, the treatment effects in Model 3 do not substantially differ from the treatment effects in Model 1 and 2. The same applies to Model 4 of Table 5. Here, age and gender are added as further explanatory variables. Neither of them has a significant effect on tax declaration.

Put simply, it seems that subjects react to a perceived higher risk of whistleblowing with a higher degree of tax compliance. But this interpretation might be premature: Figure 4 gives subjects' predictions of the whistleblowing behavior of other subjects depending on the observed percentage of income declared (0%, 20%, 40%, 60%, 80%, and 100%). These predictions are remarkably similar for "SM INC" and "LA INC", casting some doubt on the notion that the higher level of tax compliance observed in "LA INC" can be explained with a higher expected whistleblowing frequency.[20] Therefore, we run further linear and Tobit regression

[20] However, the data on the expected whistleblowing frequency should be treated with caution. Subjects' predictions of other subjects' whistleblowing behavior were only made after round 2 and may, thus, be shaped by subjects' experiences with whistleblowing in the two previous rounds of the experiment. Since, in "LA INC", only few subjects attempted to evade taxes at all, most subjects had no actual information about other subjects' willingness to report tax evaders. In "SM INC", tax compliance was lower and therefore more subjects got actual feedback on the whistleblowing behavior of other subjects. If we assume that tax evaders believe that other subjects are not very likely to blow the whistle, and if we also take into ac-

models that use the average whistleblowing expectation as an additional explanatory variable for tax compliance in round 1. Table 4 shows the results of the linear regression models. "SM INC" and "LA INC" are compared to "NO INC", since there are no whistleblowing expectations for "BASE". Model 1 of Table 4 shows that, compared to "NO INC", "LA INC" has a significantly positive effect ($p < 0.05$) on tax declaration, whereas Model 2 and 3 of Table 4 demonstrate that effort and attitude towards charities do not have a significant effect on tax declaration. The significant effect of "LA INC" remains largely unchanged ($p < 0.05$ and $p < 0.025$, respectively). However, the picture changes with regard to Model 4 of Table 4: Once the whistleblowing expectation is added as an explanatory variable for tax compliance, "LA INC" has no longer a significantly positive effect on tax declaration, whereas we now observe that the coefficient of whistleblowing expectation is positive and statistically significant. This is strong evidence that it is indeed the difference in the whistleblowing expectation that causes the different levels of tax compliance in the treatments. However, we must also admit that the corresponding Tobit regression models (Table 5) yield slightly different results, thus calling for future statistical in-depth analysis.

Explaining Tax Compliance – Adaptions in Round 2

Finally, in this subsection we take a closer look at tax compliance in the second round. In order to find out how subjects adapt their tax declarations to the experiences from the first round, we run another linear regression, seeking to explain tax declaration in round 2. We used five linear regression models, the results of which are shown in Table 6. Model 1 of Table 6 uses only one explanatory variable: "Controlled and Punished" is a dummy variable that takes the value

count that whistleblowing frequency in "SM INC" is rather high, many tax evaders experienced whistleblowing-triggered audits and may have adjusted upwards their beliefs about whistleblowing frequency. This would explain why the clear-cut treatment effect of higher whistleblowing rewards on tax declaration is not mirrored by a much higher expected whistleblowing frequency.

1 for every subject that was audited[21] and punished for tax evasion in the first round. The coefficient of this variable is positive and statistically significant ($p < 0.01$). However, it would be premature to conclude that tax declarations in round 2 are mainly driven by previous experiences of audits and punishments. As we have seen, whistleblowing frequency strongly depends on the treatments and, thus, the number of subjects that were audited and punished in round 1 depends on the treatments, too. Therefore, Model 2 of Table 6 includes the three treatments with whistleblowing mechanisms as additional explanatory variables. Once again the treatments have a significantly positive effect ($p < 0.01$) on tax declaration. The coefficient of "Controlled and Punished" remains positive but it is not longer statistically significant. Models 3, 4, and 5 of Table 6 successively add effort, attitude towards charities, and, finally, age and gender as further explanatory variables. None of these variables has a statistically significant effect on tax declarations in round 2. It is noteworthy that the attitude towards charities had a significantly positive effect ($p < 0.025$) on tax declaration in round 1, but in the second round this effect is no longer detectable. This finding suggests that tax compliance in round 2 is so predominantly driven by the various whistleblowing mechanisms that the influence of other, rather intrinsic motivations is harder to identify.

Summing up – Results on Tax Compliance

Based on the results of the Mann-Whitney u-tests and the regression analyses, we do not reject Hypotheses 4, 5, 6, and 7:

Result 4. *Even in "BASE", subjects declare on average more than 0% of their income, but there is a high level of tax evasion and full tax compliance is rare.*

[21] Audits include random audits as well as whistleblowing-triggered audits.

Result 5. *If, as in "NO INC", taxpayers face a whistleblowing mechanism without monetary incentives for whistleblowing, subjects declare a significantly larger share of their income compared to "BASE".*

Result 6. *If, as in "SM INC", taxpayers face a whistleblowing mechanism with small monetary incentives for whistleblowing, subjects declare a share of their income that does not significantly differ from the share of income declared in "NO INC".*

Result 7. *If, as in "LA INC", taxpayers face a whistleblowing mechanism with large monetary incentives for whistleblowing, subjects declare a significantly larger share of their income compared to "NO INC" and "SM INC".*

In sum, these results nicely fit the hypotheses that were based on the perceived effect of whistleblowing reported by Feldman and Lobel (2010). Broadly speaking, it seems that subjects react to a perceived higher risk of whistleblowing with a higher degree of tax compliance.

Efficiency – Tax Payments and State Revenues

Finally, we take a closer look at the efficiency of tax regimes under the various whistleblowing mechanisms. The efficiency of a tax regime can be defined in a number of different ways. Here we focus on three distinct yardsticks: voluntary tax compliance (C_V), total tax compliance (C_T), and state revenues (R). Denoting all taxes due as T_D, all taxes paid voluntarily as T_V, all taxes paid due to audits as T_A, all fines collected as F, and all rewards paid to whistleblowers as B, we define:

$$C_V = \frac{T_V}{T_D},\ C_T = \frac{T_V + T_A}{T_D} \text{ and } R = \frac{T_V + T_A + F - B}{T_D}$$

Table 7 gives the values of C_V, C_T, and R for both rounds in each of the four treatments. The values for C_V are almost identical to the values for tax declaration reported in Section 6.2. The values for C_T show that, in "BASE", the tax authority is only able to collect 38.4% of taxes due in the first round, while all treatments with whistleblowing mechanisms perform much better ("NO INC": $C_T = 77.7\%$, "SM INC": $C_T = 85.1\%$, "LA INC": $C_T = 96.9\%$). With regard to this yardstick for efficiency, we can also conclude: The larger the whistleblowing reward the higher the efficiency of the tax regime. However, the values for R tell a different story. While "BASE" still performs poorly ($R = 42.2\%$), "SM INC" is now the treatment with the highest efficiency score: Since R equals 114.6%, the bottom line is that the state actually profits from tax evasion attempts. As we have seen, in "SM INC", medium-scale tax evasion collides with medium-scale whistleblowing frequency, leading to a comparably high number of whistleblowing triggered audits (Figure 5) and correspondingly high fines. Therefore, the fines overcompensate the losses from undetected tax evasion and the costs of paying whistleblowing rewards. Compared to "SM INC", the number of reported and punished tax evaders in "NO INC" is too small to generate a similar-sized effect ($R = 97.2\%$). In "LA INC", the number of reported and punished tax evaders is also very high, but the amount of evaded taxes and, consequently, fines is smaller. Moreover, the costs of paying whistleblowing rewards are much higher. As a result, the "overcompensation effect" is less pronounced than in "SM INC" ($R = 103.1\%$). In the second round, the difference between "SM INC" ($R = 121.6\%$) and "LA INC" ($R = 97.4\%$) becomes even larger.

Conclusion

In this paper, the effect of monetary incentives on whistleblowing behavior and the effect of different whistleblowing regimes on tax compliance were experimen-

tally investigated. A tax regime without any whistleblowing mechanism was compared to three tax regimes that allow subjects to blow the whistle on tax evaders. Those three tax regimes differ from each other in terms of the incentives for whistleblowing: In one treatment, whistleblowing is costly, while in the other two treatments whistleblowers receive a reward that exceeds the costs in one case marginally and in the other case substantially.

We observe that monetary rewards for whistleblowing lead to a significant increase in the reporting of tax evasion. Moreover, we find a distinct pattern: the larger the reward the more pronounced the increase in whistleblowing. These findings do not lend support to a crowding-out effect of external whistleblowing rewards that was hypothesized in the literature. In the present experiment, subjects' behavior seems to be much more motivated by external rewards than the actions that participants of vignette studies predicted to take. However, our results also contradict predictions based on standard economic theory, since subjects do not always follow the incentives. While some subjects blow the whistle even if this behavior results in certain financial losses, other subjects refuse to report tax evasion even if they have to forego substantial whistleblowing rewards. Thus, the results suggest that some whistleblowers as well as some people who refuse to blow the whistle are motivated by intrinsic ethical concerns. The whistleblowing behavior is, by and large, correctly anticipated by taxpayers. Only the increase in whistleblowing behavior due to the introduction of small monetary incentives is underestimated by subjects. These results nicely fit the perceived general whistleblowing behavior predicted by participants of vignette studies. It thus seems that subjects have a fairly realistic view of others when it comes to assessing the influence of external rewards on whistleblowing behavior. From an efficiency point of view, especially large monetary incentives for whistleblowers are an effective way to maximize tax compliance. This makes whistleblowing

rewards a very powerful tool against tax evasion. But even without monetary incentives, whistleblowing proves to be fairly effective in curbing tax evasion.

References

Allingham, M. G., and A. Sandmo (1972): "Income tax evasion: a theoretical analysis," *Journal of Public Economics*, 1, 323–338.

Alm, J., and M. McKee (2004): "Tax Compliance as a Coordination Game," *Journal of Economic Behavior & Organization*, 54, 297 – 312.

Apesteguia, J., M. Dufwenberg, and R. Selten (2007): "Blowing the Whistle," *Economic Theory*, 31, 143–166.

Asbury, B. D. (2010): "Anti-Snitching Normas and Loyalty," *Oregon Law Review*, 89, 1257 – 1312.

Aubert, C., W. Kovacic, and P. Rey (2006): "The Impact of Leniency and Whistleblowing Programs on Cartels," *International Journal of Industrial Organization*, 24, 1241 – 1266.

Bigoni, M., S.-O. Fridolfsson, C. Le Coq, and G. Spagnolo (2011): "Fines, Leniency and Rewards in Antitrust," *IFN Working Paper*, 738, 1 – 33.

Bok, S. (1983): "Whistleblowing and Professional Responsibility," in *Ethical Theory and Business*, ed. by T. L. Beauchamp, and N. E. Bowie, chap. Whistleblowing and Professional Responsibility, p. 261. Prentice Hall, New Jersey, 2nd edn.

Deci, E. L., R. Koestner, and R. M. Ryan (1999): "A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation," *Psychological Bulletin*, 125, 627 – 668.

Dijkstra, P., M. Haan, and L. Schoonbeek (2011): "Leniency Programs and the Design of Antitrust: Experimental Evidence with Unrestricted Communication," .

Engel, C., S. J. Goerg, and G. Yu (2012): "Symmetric vs. Asymmetric Punishment Regimes for Bribery," Max Planck Institute for Research on Collective Goods Preprints.

Feldman, Y., and O. Lobel (2010): "The Incentives Matrix: A Study of the Comparative Effectiveness of Monetary Rewards as Compliance Systems," *Texas Law Review*, 88, 1151 – 1212.

Fischbacher, U. (2007): "z-Tree: Zurich Toolbox for Ready-made Economic Experiments," *Experimental Economics*, 10, 171 – 178.

Frey, B., and R. Jegen (2011): "Motivation Crowding Theory," *Journal of Economic Surveys*, 15, 589 – 611.

Friedland, N., S. Maital, and A. Rutenberg (1978): "A Simulation Study of Income Tax Evasion," *Journal of Public Economics*, 10, 107 – 117.

Gill, D., and V. Prowse (2011a): "A Novel Computerized Real Effort Task Based on Sliders," IZA Discussion Paper No. 5801.

——— (2011b): "A Structural Analysis of Disappointment Aversion in a Real Effort Competition," *American Economic Review*, 102, 469 – 503.

Hersh, M. (2002): "Whistleblower – Heroes or Traitors?: Individual and Collective Responsibility for Ethical Bevahiour," *Annual Reviews of Control*, 26, 243 – 262.

Koch, A. (2007): "Denunziation als Straftat nach 1945 und 1989," in *Der willkommene Verrat*, ed. by M. Schröter, pp. 185 – 202. Velbrück Wissenschaft, Weilerswist.

Mealem, Y., Y. Tobol, and G. Yaniv (2010): "Whistle-blowers as a Deterrent to Tax Evasion," *Public Finance Review*, 38, 306 – 320.

Near, J., and M. Miceli (1985): "Organizational Dissidence: The Case of Whistle-Blowing," *Journal of Business Ethics*, 5, 1 – 16.

Osterloh, M., J. Frost, and B. Frey (2002): "The Dynamics of Motivation in New Organizational Forms," *International Journal of the Economics of Business*, 9, 61 – 77.

Ryan, R., and E. L. Deci (2000): "Self-determination Theory and the Facilitation of Intrinsic Motivation, Social Development and Well-Being," *American Psychologist*, 55, 68 – 78.

Schikora, J. (2011): "Bringing good and bad whistle-blowers to the lab," Munich discussion paper.

Selten, R. (1967): "Die Strategiemethode zur Erfoschung des eingeschränkt rationalem Verhaltens im Rahmen eines Oligopolexperimentes," in *Beiträge zur experimentellen Wirtschaftsforschung*, ed. by H. Sauermann, pp. 136 – 168. J.C.B. Mohr (Paul Siebeck), Tübingen.

Spagnolo, G. (2004): "Divide et Impera: Optimal Leniency Programmes," CEPR Discussion Papers 4506.

Torgler, B. (2002): "Speaking to Theorists and Searching for Facts: Tax Morale and Tax Compliance in Experiments," *Journal of Economic Surveys*, 16, 657 – 683.

Yaniv, G. (2001): "Revenge tax informing, and the optimal bounty," *Journal of Public Economic Theory*, 3, 225 – 233.

Yitzhaki, S. (1974): "A Note on Income Tax Evasion: A Theoretical Analysis," *Journal of Public Economics*, 3, 201 – 202.

Appendix

Figures

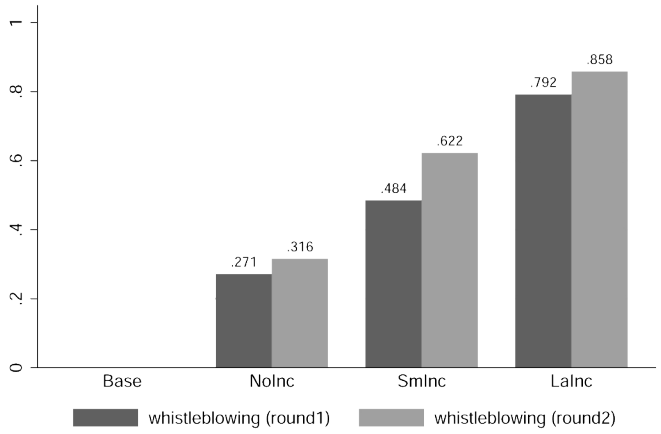


Figure 1: Whistleblowing (Probability of Whistleblowing)

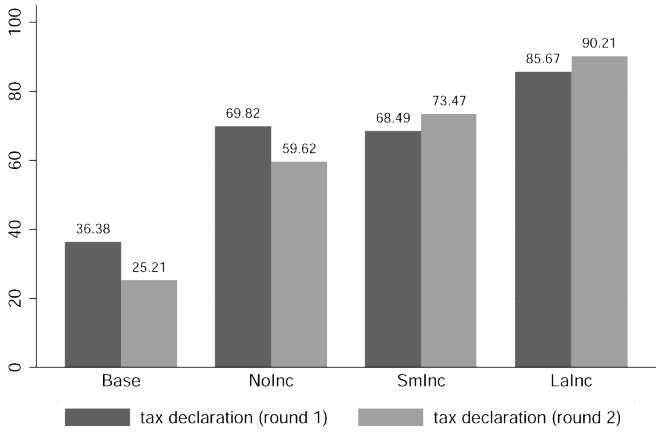


Figure 2: Tax Declaration (Percentage of Income Reported)

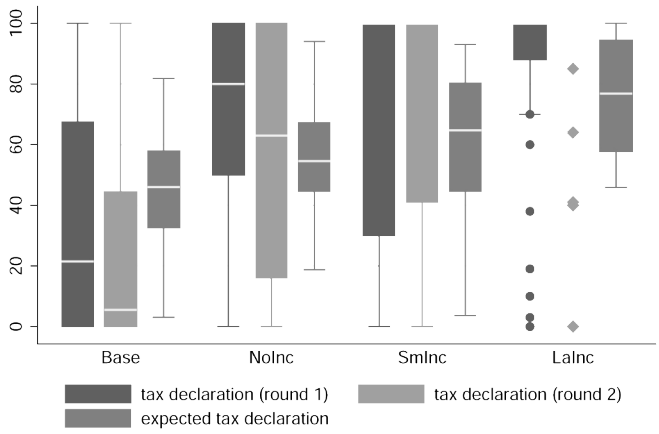


Figure 3: Box Plot for Tax Declaration

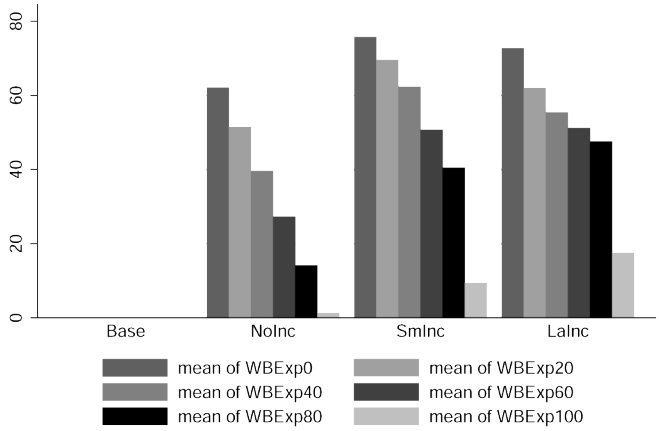


Figure 4: Whistleblowing Expectations

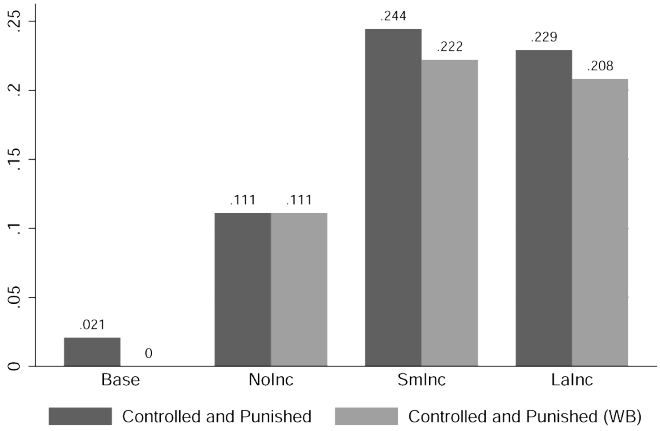


Figure 5: Share of Controlled and Punished Subjects (Round 1)

Tables

Whistleblowing (round 1)	(1)	(2)	(3)	(4)
Small Incentives	0.522**	0.537**	0.514**	0.518**
	(0.2098)	(0.2112)	(0.2102)	(0.2069)
Large Incentives	1.394***	1.394***	1.564***	1.588***
	(0.2463)	(0.2465)	(0.3232)	(0.3224)
Effort (round 1)		-0.0045	0.0001	-0.0029
		(0.0068)	(0.0076)	(0.0079)
Charity			0.0255	0.0332
			(0.0486)	(0.0483)
Age				0.0110
				(0.0203)
Gender				-0.2660
				(0.1973)
Constant	-0.0546	0.1684	-0.2314	-0.2661
	(0.1539)	(0.3662)	(0.5358)	(0.7371)
Observations	138	138	114	114
$Prob > Chi^2$	0.0000	0.0000	0.0000	0.0000

Standard errors in parentheses
*** $p < 0.01$, ** $p < 0.025$, * $p < 0.05$

Table 1: Whistleblowing (Round 1), Tobit Regression Models

Whistleblowing (round 2)	(1)	(2)	(3)	(4)
Small Incentives	0.956***	0.956***	0.935***	0.950***
	(0.3139)	(0.3143)	(0.3053)	(0.3056)
Large Incentives	2.047***	2.047***	2.038***	2.045***
	(0.4071)	(0.4081)	(0.4843)	(0.4833)
Effort (round 1)		0.0002	-0.0016	-0.0050
		(0.0091)	(0.0098)	(0.0104)
Charity			0.0090	0.0006
			(0.0683)	(0.0687)
Age				-0.0124
				(0.0288)
Gender				-0.2375
				(0.2826)
Constant	-0.0636	-0.0734	0.1007	0.6250
	(0.2158)	(0.5554)	(0.7719)	(1.0825)
Observations	138	138	114	114
$Prob > Chi^2$	0.0000	0.0000	0.0000	0.0000

Standard errors in parentheses

*** $p < 0.01$, ** $p < 0.025$, * $p < 0.05$

Table 2: Whistleblowing (Round 2), Tobit Regression Models

Tax Compliance (round 1)	(1)	(2)	(3)	(4)
No Incentices	65.97***	67.82***	59.31***	62.48***
	(16.87)	(16.98)	(15.78)	(15.89)
Small Incentives	66.08***	69.48***	57.86***	59.74***
	(17.09)	(17.44)	(16.27)	(16.33)
Large Incentives	112.3***	113.8***	123.5***	125.3***
	(18.78)	(18.90)	(24.49)	(24.48)
Effort (round 1)		-0.553	-0.346	-0.0925
		(0.510)	(0.525)	(0.534)
Charity			7.730**	7.229**
			(2.985)	(2.989)
Age				1.504
				(1.497)
Gender				16.66
				(12.70)
Constant	27.29**	52.96*	-3.680	-58.80
	(11.37)	(26.17)	(32.45)	(53.79)
Observations	186	186	162	162
$Prob > Chi^2$	3.20e-09	7.61e-09	4.28e-10	1.47e-09

Standard errors in parentheses

*** $p < 0.01$, ** $p < 0.025$, * $p < 0.05$

Table 3: Tax Compliance (Round 1), Tobit Regression Models

Tax Declaration (round 1)	(1)	(2)	(3)	(4)
Small Incentives	-1.333	-0.750	-1.498	-6.708
	(7.219)	(7.256)	(6.999)	(7.301)
Large Incentives	15.84*	15.58*	20.32**	15.74
	(7.105)	(7.118)	(8.600)	(8.726)
Effort (round 1)		-0.201	-0.171	-0.237
		(0.229)	(0.250)	(0.248)
Charity			2.976	2.568
			(1.592)	(1.578)
Whistleblowing Expectation				0.623*
				(0.122)
Constant	69.82**	79.83***	57.64***	53.56***
	(5.104)	(12.52)	(17.44)	(17.27)
Observations	138	138	114	114
R^2	0.051	0.057	0.121	0.157
$Prob > Chi^2$	0.0284	0.0490	0.00659	0.00215

Standard errors in parentheses

*** $p < 0.01$, ** $p < 0.025$, * $p < 0.05$

Table 4: Tax Compliance and Whistleblowing Expectation (Round 1), Linear Regression Models

Tax Declaration (round 1)	(1)	(2)	(3)	(4)
Small Incentives	0.188	1.966	-1.365	-11.59
	(17.85)	(17.87)	(16.27)	(16.96)
Large Incentives	47.73**	47.33**	64.67***	54.31*
	(19.14)	(19.10)	(24.55)	(24.28)
Effort (round 1)		-0.631	-0.409	-0.533
		(0.605)	(0.633)	(0.618)
Charity			8.154*	7.336
			(3.878)	(3.794)
Whistleblowing Expectation				0.534
				(0.305)
Constant	94.49***	125.8***	56.43	47.55
	(12.96)	(33.13)	(42.00)	(41.15)
Observations	138	138	114	114
$Prob > Chi^2$	0.0149	0.0233	0.00119	0.000801

Standard errors in parentheses

*** $p < 0.01$, ** $p < 0.025$, * $p < 0.05$

Table 5: Tax Compliance and Whistleblowing Expectation (Round 1), Tobit Regression Models

Tax Declaration (round 2)	(1)	(2)	(3)	(4)	(5)
Controlled and	24.57***	8.835	8.616	12.78	11.88
punished	(8.680)	(7.613)	(7.723)	(9.042)	(9.068)
No Incentives		33.62***	33.54***	32.54***	31.10***
		(7.477)	(7.507)	(7.739)	(7.802)
Small Incentives		46.28***	46.16***	44.45***	43.59***
		(7.638)	(7.687)	(8.074)	(8.117)
Large Incentives		63.16***	63.20***	66.81***	66.14***
		(7.494)	(7.517)	(9.760)	(9.769)
Effort (round 1)			0.0394	0.00728	-0.109
			(0.211)	(0.239)	(0.253)
Charity				0.971	1.247
				(1.437)	(1.451)
Age					-0.378
					(0.687)
Gender					-8.575
					(6.188)
Constant	58.28***	25.02***	22.93	18.51	37.38
	(3.368)	(5.182)	(12.36)	(16.42)	(26.27)
Observations	186	186	162	162	162
R^2	0.042	0.323	0.323	0.325	0.334
$Prob > Chi^2$	0.00516	0	0	0	1.04e-10

Standard errors in parentheses

*** $p < 0.01$, ** $p < 0.025$, * $p < 0.05$

Table 6: Tax Compliance (Round 2), Linear Regression Models

	BASE	NO INC	SM INC	LA INC
Voluntary Tax	36.463%	67.939%	68.805%	85.235%
Compliance (C_V)	24.818%	60.477%	73.296%	90.808%
Total Taxes paid (C_T)	38.379%	77.692%	85.090%	96.907%
(voluntary or due to controls)	29.314%	76.881%	90.513%	95.595%
State Revenues (R)	42.212%	97.197%	114.573%	103.091%
	38.305%	109.689%	121.557%	97.370%

in % of taxes due for Round 1 (upper entry) and 2 (lower entry)

Table 7: Efficiency

The Influence of Regret on Decision Making: Theory and Experiment

Felix Schran*

Introduction

Most people experience regret when realizing that an alternative decision would have been the better one, retrospectively. They do not only care about what they actually get but also what they might have gotten had they chosen differently. So their utility is anything but independent from results of other previously possible decisions. Furthermore, people are able to anticipate regret (Zeelenberg, M., 1999). Therefore, regret is not only an emotion felt ex post, but is also able to influence decisions ex ante. The necessity to compare the present situation to a hypothetical situation makes regret a more complex emotion than the basic emotions like anger, fear or happiness. As a consequence, regret is developed relatively late in childhood namely at an age of approximately five to seven years (Guttentag, R. E. and J. M. Ferrell, 2004). In addition, regret is often influenced by culture and morality (Zeelenberg, M. and R. Pieters, 2007). Expected utility theory (EUT) ignores that decision makers usually compare what they could have gotten had they chosen differently to what they get. The standard model

* Felix Schran received his degree in Economics (B. Sc.) from the University of Bonn in 2013. The present article refers to his bachelor thesis under the supervision of Dr. S. Ebert and Prof. Dr. H. Hakenes, which was submitted in December 2012.

proposes that only the actually received outcome matters, and people will choose the alternative with the highest utility in expectation. However, several violations to EUT like the paradox of Allais, M. (1953) were observed in empirical studies casting doubt on the underlying assumptions. As a consequence, more descriptive theories like prospect theory (Kahneman, D. and A. Tversky, 1979; Tversky, A. and D. Kahneman, 1992) or regret theory (Loomes, G. and R. Sugden, 1982; Bell, D. E., 1982) were developed, able to explain a lot of the observed phenomena.

Regret theory takes into account that the well-being of a person is not independent of the results of decisions which were also possible. If another decision would have been better ex post people regret, otherwise they rejoice. So regret is a negative emotion whereby large intensities of regret are weighted disproportionally heavier than small ones (Zeelenberg, M. and R. Pieters, 2007). Decision makers try to avoid regret, especially large amounts. Hence, they are regret averse. The here discussed version of regret theory (Loomes, G. and R. Sugden, 1982) is based on two functions only - a common utility function and a regret function capturing the impact of regret. To answer the question whether this concept is in accordance with empirical evidence, an experiment was conducted in this work based on the parameter free approach of Bleichrodt, H., A. Cillo and E. Diecidue (2010). The measurement approach is parameter free in the sense that no assumptions on neither the utility nor the regret function are necessary. Furthermore, heterogeneity of preferences is considered, and so the method can be performed on an individual level.

Theory and Examples

Regret theory suggests that the utility of a decision maker depends on both, what he actually receives and what he could have received had he chosen in a

different way. The outcome of the unchosen alternative now serves as a reference point to which the result of the chosen alternative is relatively evaluated. The following notation refers to Loomes, G. and R. Sugden (1982) and Loomes, G. and R. Sugden (1987). Let $A = \{A_1, ..., A_m\}$ be a set of possible actions which can be chosen. If the outcome of the selected action, say A_i, is smaller than the outcome of at least one unchosen action, say A_k, one regrets. Therefore, regret theory adjusts EUT regarding the influence of regret by proposing a two-dimensional, skew-symmetric and real-valued utility function $\Psi(x_{ij}, x_{kj})$. x_{ij} denotes the outcome of action A_i in state S_j whereby state $S_j \in \{S_1, ..., S_n\}$ occurs with probability p_j. Similarly, x_{kj} denotes the outcome of action A_k in state S_j. One regrets his decision for A_i if $x_{kj} > x_{ij}$ because he could have chosen A_k instead. Thus, $\Psi(\cdot, \cdot)$ should be decreasing in the second and increasing in the first argument. Since no regret can occur if both possible actions yield the same outcomes, $\Psi(\xi, \xi) = 0$ should follow. As a result of skew-symmetry ($\Psi(\xi, \phi) = -\Psi(\phi, \xi)$), an action A_i is preferred to A_k if the expected adjusted utility of choosing A_i and simultaneously rejecting A_k is positive:

$$A_i \succ A_k \Leftrightarrow \sum_{j=1}^{n} p_j \Psi(x_{ij}, x_{kj}) > 0$$

To make the estimation easier, a restricted form of $\Psi(\cdot, \cdot)$ was used with $\Psi(x_{ij}, x_{kj})$ $= Q(u(x_{ij}) - u(x_{kj}))$ where $Q(\cdot)$ is a strictly increasing function with the property of symmetry: $Q(-\xi) = -Q(\xi)$. $u(\cdot)$ denotes a concave Bernoulli utility function. Therefore, the intensity of regret only depends on the utility difference between chosen and rejected action. As mentioned earlier, most people are regret averse; they want to avoid large regrets in particular (Zeelenberg, M., 1999). This characteristic can be considered if $Q(\xi)$ is convex for all $\xi > 0$. If this holds, large differences between $u(x_{ij})$ and $u(x_{kj})$ are weighted heavier than small ones. It

follows that large amounts of regret decrease the expected adjusted utility disproportionally more than small amounts due to $Q(-\xi) = -Q(\xi)$.

A vivid example when regret has influence on decisions is the reluctance to exchange lottery tickets observed by van de Ven, N. and M. Zeelenberg (2011). In their experiment participants were endowed with lottery tickets all of them having an equal chance to win a voucher. One group obtained the tickets in sealed envelopes not knowing their ticket number whereas participants of another group knew their ticket numbers. The subjects were then asked to exchange their tickets with a member of their group. As an incentive they received a pen when exchanging. Now it appeared that significantly more participants from the sealed envelope group were willing to exchange their ticket. If one assumes that the pen exceeds the transaction costs of the exchange, this behavior is not explainable by EUT. Every participant should be willing to trade. If regret about their decisions is considered, such a behavior is understandable, though. Subjects knowing their number are able to anticipate possible regret in case they had exchanged and their former ticket number is drawn. As a consequence to regret aversion, they are reluctant to exchange. In contrast, the attendants of the control group will never gain information about their former number after having exchanged. They will not be able to regret their decisions, and so they are consistently more willing to trade their tickets.

A further example is the experimentally observed tendency of bidders in first-price sealed-bid auctions to bid more than predicted by theory (Cox, J. C., V. L. Smith and J. M. Walker, 1988). In a first-price auction the subject with the highest bid wins and has to pay his own bid. The other bidders pay nothing. To gain profit, bidders have to provide bids lying under their true valuations for the auctioned good. Solving the implied game an unambiguous Bayesian-Nash-Equilibrium is found. Nevertheless, empirical evidence shows that offered bids

often lie above the predicted ones if it is common knowledge that the winning bid will be revealed afterwards. If so, a bidder may regret his decision having offered a bid lying way under his valuation in case the winning bid was only slightly higher than his own (Filiz-Ozbay, E. and E. Y. Ozbay, 2007). If he had only bidden little more, he would have won while still making profit. The possible regret can be anticipated if it is known that the winning bid will be common knowledge afterwards. Consistently, observed bids in first-price auctions, with ex post information about the winning bid, lie slightly above the predicted ones.

Experimental Design

The problem in quantitatively measuring regret theory is the composition of regret and utility function. A first feasible method was developed by Bleichrodt, H., A. Cillo and E. Diecidue (2010). The following refers to their work. The foundation of their parameter free approach is the so called trade-off method originally proposed by Wakker, P. P. and D. Deneffe (1996) to measure the value function of cumulative prospect theory (Tversky, A. and D. Kahneman, 1992). In a first step this method is applied to elicit a standard sequence of outcomes $x_0, ..., x_k$ to infer a participant's utility function u. The second step consists of constructing a second sequence $z_1, ..., z_l$ to measure the regret function Q. Consistency with regret theory would imply that the obtained utility functions are concave whereas the elicited regret functions are convex.

To determine the standard sequence of outcomes, the participants were asked which amount of money x_i would lead to indifference between the two binary actions $A_{i-1} = x_{i-1_p}G_{1-p}$ and $B_i = x_{i_p}g_{1-p}$ for $i = 1, ..., k$. The elicitation of x_i was accomplished by an iterative algorithm. x_{i-1}, obtained with probability p, denotes the outcome determined in the previous stage. Therefore, the elicitation of each x_i depends on the prior determination of x_{i-1} resulting in a chained

structure of the standard sequence. G and g denote two gauge outcomes obtained with probability $1-p$ each. They can be arbitrarily chosen as long as $x_0 > G > g \geq 0$ holds to ensure that the sequence increases. x_0 denotes the starting value of the sequence. It can be assumed that a participant will choose $x_i > x_{i-1}$ due to $g < G$. Expressed by regret theory the indifference between A_{i-1} and B_i and between A_i and B_{i+1} yields:

$$B_i \sim A_{i-1} \Leftrightarrow pQ(u(x_i) - u(x_{i-1})) + (1-p)Q(u(g) - u(G)) = 0$$
$$B_{i+1} \sim A_i \Leftrightarrow pQ(u(x_{i+1}) - u(x_i)) + (1-p)Q(u(g) - u(G)) = 0$$

Since the regret function is assumed to be strictly increasing, Q is invertible. Thus, the two foregoing equations put together result in:

$$u(x_{i+1}) - u(x_i) = u(x_i) - u(x_{i-1}), i = 1, ..., k-1$$

By setting $u(x_0) = 0$ and $u(x_k) = 1$, it follows: $u(x_{i+1}) - u(x_i) = 1/k$. Therefore, the points $(x_0, 0), ..., (x_j, \frac{j}{k}), ..., (x_k, 1)$ can be obtained. In the conducted experiment the gauge outcomes $G = 16, g = 11$, the starting value $x_0 = 20$ and the probabilities $p = 1/3, 1-p = 2/3$ were used. Overall, six points were elicited, hence $k = 5$. The experiment was programmed and conducted with 'z-Tree' (Fischbacher 2007). To determine x_1, the participants were iteratively asked which outcome would make them indifferent between $A = 20_{1/3}16_{2/3}$ and $B = x_{1_{1/3}}11_{2/3}$ beginning with $x_1 = 45$ and then slightly modifying x_1 to find the true value. Therefore, the subjects always just had to choose a lottery A or B instead of directly indicating the sought value because a direct method might result in massive response errors due to the large cognitive demands. A screenshot of the first elicitation part can be found in Figure 1. Overall, 27 subjects, mainly economics undergraduates, participated. Two of them were randomly chosen playing for real

money in two of their choices to motivate truthful answers. One experimental unit was worth 0.10 euros resulting in an average payment of 13.20 euros. The subjects knew that the outcomes of the unchosen alternatives would be revealed afterwards. Therefore, they were able to anticipate regret.

In the second step the elicited standard sequence of outcomes was used to determine a further sequence $z_1, ..., z_l$. In contrast to the first step, now the probabilities were altered for each outcome. Firstly, the subjects were asked which outcome z_1 would make them indifferent between $A_1 = x_{4_{p_1}} x_{0_{1-p_1}}$ and $B_{z_1} = x_{3_{p_1}} z_{1_{1-p_1}}$ with x_4, x_0, x_3 as outcomes belonging to the standard sequence and $p_1 = 1/4, 1 - p_1 = 3/4$ denoting the corresponding probabilities. The elicitation of z_1 was done by an iterative algorithm, so the subjects only had to choose one out of two lotteries again. Generally the subjects were asked for outcomes z_j making them indifferent between $A_j = x_{4_{p_j}} x_{0_{1-p_j}}$ and $B_{z_j} = x_{3_{p_j}} z_{j_{1-p_j}}$. Expressing this indifference via regret theory gives:

$$p_j Q(u(x_4) - u(x_3)) + (1 - p_j) Q(u(x_0) - u(z_j)) = 0, j = 1, ..., l$$

Scaling the regret function via $Q(u(x_4) - u(x_3)) = Q(1/k) = 1$ and because of $u(x_0) = 0$ together with $Q(0) = 0$, this results in:

$$Q(u(z_j)) = \frac{p_j}{1 - p_j}, j = 1, ..., l$$

By using the probabilities $p_1 = 1/4, p_2 = 2/5, p_3 = 3/5, p_4 = 3/4$ the six points $(0, 0), (u(z_1), 1/3), (u(z_2), 2/3), (1/5, 1), (u(z_3), 3/2), (u(z_4), 3)$ were obtained.

Since only the outcomes $z_1, ..., z_4$ were determined in the procedure, the corresponding utilities had to be inferred to elicit the function $Q(\cdot)$ instead of eliciting the composition function $Q(u(\cdot))$. This was done by linear interpolation using the previously estimated utility functions.

Results

The obtained utility functions u of the first five subjects and the utility function based on the mean data can be found in Figure 2. On the horizontal axis the estimated outcomes are located. The vertical axis shows the corresponding utilities. The slightly dotted line is drawn for comparison to the case of a linear utility function. Consistency with regret theory would imply concave functions. To test for concavity, two classifications were used which were also applied by Bleichrodt, H., A. Cillo and E. Diecidue (2010) – a parametrical and a non-parametrical one. In the parametrical classification non-linear regressions were conducted for every participant estimating the coefficients α, β of the power function $y = \alpha(x - 20)^{\beta}$ by non-linear least squares. Afterwards, it was tested whether β was significantly smaller (larger) than 1 corresponding to a concave (convex) function. Overall, 17 concave and 3 convex functions were found. Thus, a one-tailed binomial test showed that there were significantly more concave than convex utility functions ($p = 0.001$). The estimated power coefficient based on the mean data was $\hat{\beta} = 0.832$ ($SE = 0.025$). Therefore, the function based on the mean was classified concave as well (see Figure 3). The non-parametrical classifications were more technical. For each function twenty differences $\Delta_{gh,lm} = (x_g - x_h) - (x_l - x_m),\ g > h,\ g > l,\ g - h = l - m$ for all outcomes x_i with $g,\ h,\ l,\ m \in \{0, ..., 5\}$ were calculated. Since $u(x_g) - u(x_h) = u(x_l) - u(x_m)$ holds if $g - h = l - m$, it follows that the corresponding part is concave if $\Delta_{gh,lm}$ is positive. If $\Delta_{gh,lm}$ is negative, however, this corresponds to a convex part. Larger getting distances between the determined outcomes imply concavity because the difference between two contiguous utilities is always the same: $1/k$. A function was already classified concave if only 50% of the calculated values were positive due to response errors. By this method, 23 concave and 3 convex functions were found. Again, significant evidence for concavity showed up ($p = 0.000$, one-tailed

binomial test). The utility function based on the mean was equally classified concave.

Figure 4 shows the elicited regret functions Q of the first five participants and the one based on the mean data. On the horizontal axis the utility difference between chosen and unchosen action can be found. On the vertical axis the value of Q is shown. The dotted line is drawn for comparison representing a linear function. If the subjects behaved according to regret theory, the estimated regret functions should be convex because large differences between chosen and unchosen actions carry disproportionally more weight than small ones. Again a parametrical and a non-parametrical classification were used. For the first one the coefficients α, β of the power function $y = \alpha x^{\beta}$ were calculated. Afterwards, it was tested whether β was significantly greater (smaller) than 1 corresponding to a convex (concave) shape of the regret function. As can already be supposed by Figure 4, the estimation of Q revealed more noise than the elicitation of u probably due to the more complex lotteries faced in the second part. Thus, the standard errors were generally greater. Only 10 functions exhibited a significant convex and 2 functions a significant concave shape. For the regret function based on the mean data a power coefficient $\hat{\beta} = 1.376$ ($SE = 0.187$) was computed being only weakly significant greater than 1 (see Figure 5). However, the non-parametrical method revealed more evidence for convexity. For each subject twenty values $\nabla_{gh,lm} = (Q(g/5) - Q(h/5)) - (Q(l/5) - Q(m/5)),\ g > h,\ g > l,\ g - h = l - m,\ g,\ h,\ l,\ m \in \{0, ..., 5\}$ were calculated whereby positive values now corresponded to convex parts. A regret function was said to be convex if only 50% of the differences were positive. Hereby 22 convex and 3 concave functions were found. Therefore, the proportion of convex to concave shapes was highly significant in support of convex functions ($p = 0.000$, one-tailed binomial test). Likewise, the regret function based on the mean data was classified convex.

Thus, despite more noise in the estimation of the regret functions, evidence for regret aversion was found.

Summarizing, the results of Bleichrodt, H., A. Cillo and E. Diecidue (2010) as well as the results presented in this work showed that most subjects take the feeling of regret into account when making a decision. They compare retrospectively what they could have gotten to what they actually get, but they are also able to anticipate regret ex ante. Regret occurs if an alternative decision would have been the better one whereas subjects are disproportionally averse to large regrets. The data revealed evidence for this because mainly convex regret functions were found corresponding to regret aversion. Furthermore, evidence for concave utility functions showed up. So overall the assumptions of regret theory could be experimentally verified.

References

Allais, M. (1953): "Le comportement de l'homme rationel devant le risque, critique des postulates et axiomes de l'ecole americaine," *Econometrica*, 21, 503–546.

Bell, D. E. (1982): "Regret in Decision Making Under Uncertainty," *Operations Research*, 30, 961–981.

Bleichrodt, H., A. Cillo and E. Diecidue (2010): "A Quantitative Measurement of Regret Theory," *Management Science*, 56, 161–175.

Cox, J. C., V. L. Smith and J. M. Walker (1988): "Theory and Individual Behavior of First-Price Auctions," *Journal of Risk and Uncertainty*, 1, 61–99.

Filiz-Ozbay, E. and E. Y. Ozbay (2007): "Auctions with Anticipated Regret: Theory and Experiment," *American Economic Review*, 97, 1407–1418.

Fischbacher, U. (2007): "z-Tree: Zurich Toolbox for Ready-made Economic Experiments," *Experimental Economics*, 10, 171–178.

Guttentag, R. E. and J. M. Ferrell (2004): "Reality Compared With Its Alternatives: Age Differences in Judgments of Regret and Relief," *Developmental Psychology*, 40, 764–775.

Kahneman, D. and A. Tversky (1979): "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47, 263–291.

Loomes, G. and R. Sugden (1982): "Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty," *The Economic Journal*, 92, 805–824.

——— (1987): "Some Implications of a More General Form of Regret Theory," *Journal of Economic Theory*, 41, 270–287.

Tversky, A. and D. Kahneman (1992): "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5, 297–323.

van de Ven, N. and M. Zeelenberg (2011): "Regret Aversion and the Reluctance to Exchange Lottery Tickets," *Journal of Economic Psychology*, 32, 194–200.

Wakker, P. P. and D. Deneffe (1996): "Eliciting von Neumann-Morgenstern Utilities when Probabilities are Distorted or Unknown," *Management Science*, 42, 1131–1150.

Zeelenberg, M. (1999): "Anticipated Regret, Expected Feedback and Behavioral Decision Making," *Journal of Behavioral Decision Making*, 12, 93–106.

Zeelenberg, M. and R. Pieters (2007): "A Theory of Regret Regulation 1.0," *Journal of Consumer Psychology*, 17, 3–18.

Appendix

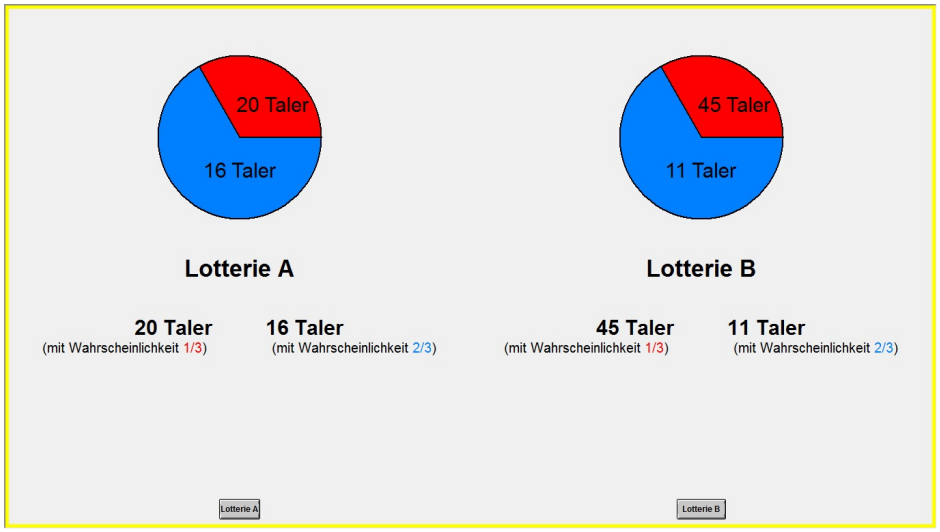


Figure 1: Example of a screen faced in the first part

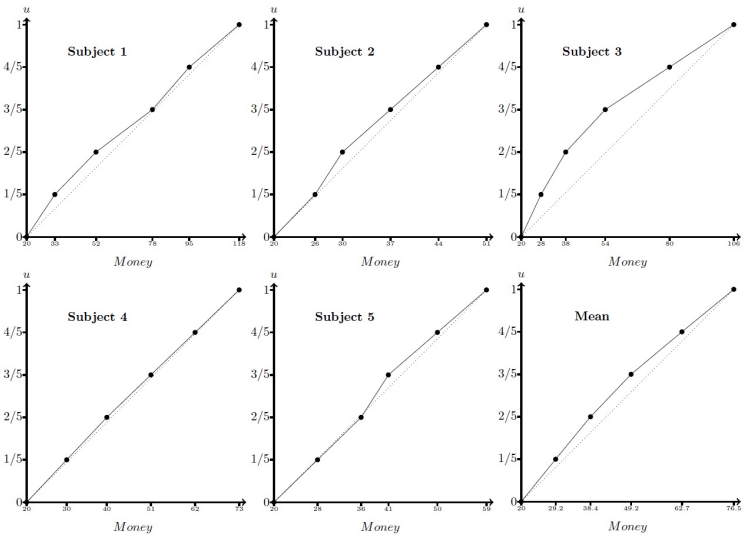


Figure 2: Utility functions of the first five subjects and based on the mean data

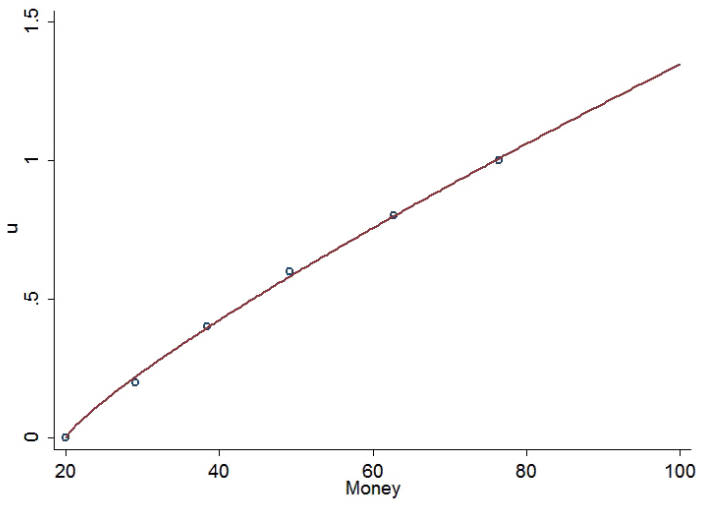


Figure 3: Non-linear regression of the utility function based on the mean data

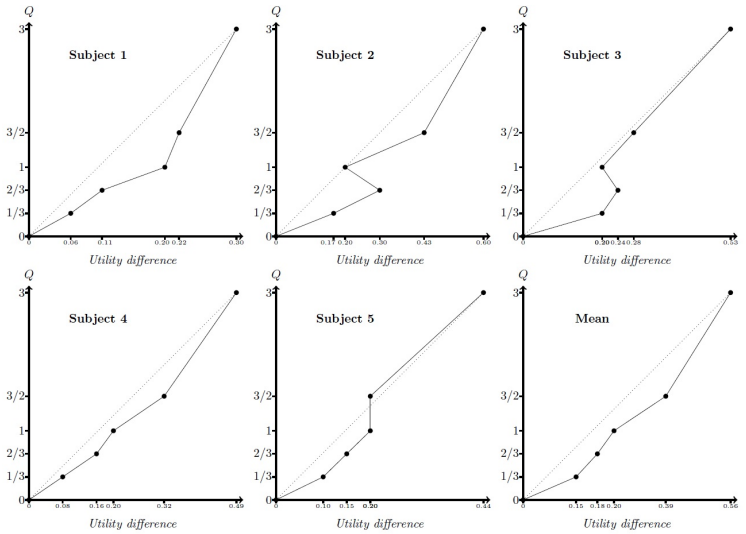


Figure 4: Regret functions of the first five subjects and based on the mean data

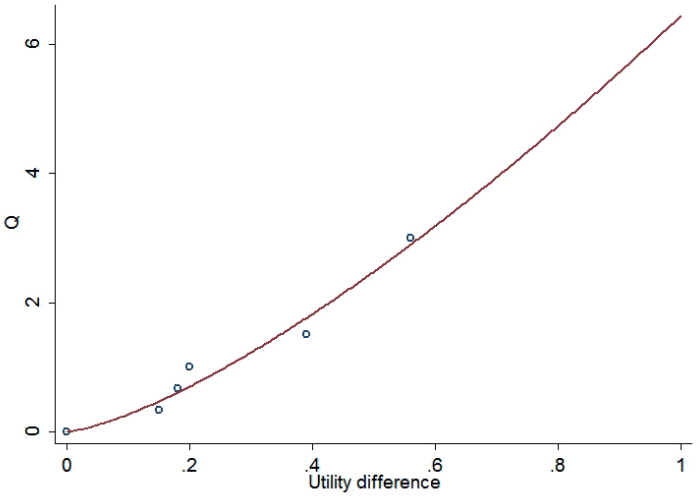


Figure 5: Non-linear regression of the regret function based on the mean data

The Determinants of Migration Flows in Europe

Ulrike Steins*

Introduction

The Treaty of Rome, which was signed in 1957, laid the foundation for a European common market. The member states, i.e. France, Germany, Italy and the Benelux countries, set themselves the target to create an economic area with free movement of goods, services and workers across borders (European Union, 2012). This union of states was expanded in the following years, so that in 2012 the European Union comprises 27 member states. However, the current European sovereign debt crisis has questioned the construction of the European Union. A public debate is underway on which structural reforms are needed to ensure the cohesion of the EU in the future.

One fundamental problem is the lack in geographic labor mobility. The simultaneous presence of high unemployment rates and labor shortages indicates an inefficient allocation of resources. Bonin et al. (2008) argue that the economic gain of higher labor mobility would overweight negative externalities such as potential downward pressure on wages. Thus, by improving the adjustment of labor market imbalances, labor mobility within the European Union would have a pos-

* Ulrike Steins received her degree in Economics (M. Sc.) from the University of Bonn in 2012. The present article refers to her master thesis under the supervision of Prof. Dr. Klaus F. Zimmermann, which was submitted in August 2012.

itive welfare effect in total.

Although migration flows are too small, there are already people who move within the EU. The determination of forces that have driven these migration flows may give indication which policies could efficiently increase labor mobility. This work analyzes bilateral migration within the European area from a macroeconomic perspective. Beneath economic and demographic factors, such as income differences, the influence of past political events is also analyzed. This accounts for the fact that the identification of migration determinants may serve as a basis for policies intended on higher labor mobility. For instance, if the Maastricht Treaty, which eased free movement within the EU, affected migration flows in a positive way, this would suggest that a change in institutional settings could be a way to increase mobility. It is focused on the migration flows between Germany and a panel of the remaining first eleven member states (EU-11) for the following reasons. First, definitions of migration flows differ within the EU. By limiting the analysis to Germany, data from one source can be used so that comparability issues are avoided. Second, Bonin et al. (2008) find that there is a great disparity among EU countries regarding migration patterns so that pooling countries for the analysis on migration in Europe might not be reasonable.

The question which factors drive migration flows has been discussed in various empirical studies. Those studies that use macro panel data can be broadly distinguished into two groups. On the one hand, there is a group of empirical papers that study bilateral migration flows with underlying data of both various origins and destination countries (Pedersen, Pytlikova, and Smith, 2008; Ortega and Peri, 2009; Mayda, 2010). On the other hand, there are studies that examine the migration flows from and to one specific country (Karemera, Oguledo, and Davis, 2000; Bertoli and Moraga, 2011; Brücker, Siliverstovs, and Trübswetter, 2003).

One issue that arises when bilateral migration flows are analyzed is that potential

migrants have to choose between various destination countries. This means the attractiveness of other countries influences the migration decision. Several studies do not consider this aspect in the empirical specification, which may bias results. Based on Bertoli and Moraga (2011), it is controlled for the presence of alternative destinations by applying the so-called Augmented Mean Group estimator (Eberhardt and Bond, 2009). It includes an additional regressor during the estimation procedure which measures the average dynamic evolution of explanatory variables across countries. This estimation method is also an appropriate choice as it is robust to cross-sectional dependence across countries and non-stationary variables, which is likely to be an issue in this sample (Brücker, Siliverstovs, and Trübswetter, 2003).

Determinants of Migration

Most of the empirical studies refer to a model of migration in which an individual maximizes its utility by comparing the outcome when staying in the source country with the one in potential destination countries. Thereby, the net utility of moving takes into accounts the costs and risks of migration. If the utility of moving is greater than the one of staying, the individual decides to migrate. The sum of these individual decisions causes migration flows from a sending country i to a destination country j (Clark, Hatton, and Williamson, 2007).

The empirical model is based on Pedersen, Pytlikova, and Smith (2008). A migration flow from country i to country j at time t is modeled in the following

way:

$$\ln(m_{ijt}) = \beta_{1ij}\ln(s_{ijt-1}) + \underbrace{\beta_{2ij}\ln(Y_{it-1}) + \beta_{3ij}\ln(U_{it-1}) + \beta_{4ij}\ln(P_{it-1})}_{Push factors}$$
$$+ \underbrace{\beta_{5ij}\ln(Y_{jt-1}) + \beta_{6ij}\ln(U_{jt-1})}_{Pull factors} + \beta_{7ij}T_{ijt} + c_{ij} + u_{ijt} \quad (1)$$

where m_{ij} represents the migration flow from i to j and s_{ijt-1} stands for the stock of country i's immigrants already residing in country j, which accounts for network effects. In order to achieve comparability across countries, m_{ij} and s_{ijt-1} are expressed in relative terms. Push and pull factors of the source and destination country are accounted for by real GDP per capita as a proxy for income (Y_{it-1} and Y_{jt-1}) and the unemployment rate of the sending and destination country (U_{it-1} and U_{jt-1}). Besides, the share of young people in the source country's population (P_{it-1}) is taken as a demographic push factor. The explanatory variables are lagged one period, which accounts for the fact that potential migrants need time to gather information before they decide to move to another country (Pedersen, Pytlikova, and Smith, 2008). In addition, issues of reverse causality are avoided. c_{ij} captures country fixed effects that are constant over time and $\beta_{1ij}, ..., \beta_{4ij}$ are the country-specific slope parameters. T_{ijt} is a linear trend variable and u_{ijt} is an error term. Except for the linear trend, all variables that vary over time are taken in logs. By doing so, the coefficients represent the partial elasticities of migration flows with respect to the corresponding explanatory variable (see Wooldridge, 2002, pp.15–18).

The following hypotheses regarding the explanatory variables are stated (see Table 1): β_{1ij}, the coefficient of s_{ijt-1} in (1), should have a positive sign. The more immigrants already live in a destination country, the higher the migration inflows. The higher the income in the destination country or the lower the in-

come in the source country, the more likely does an individual decide to move. Hence, β_{2ij}, the coefficient of Y_{it-1} in (1), is assumed to have a negative sign, whereas β_{5ij}, the coefficient of Y_{jt-1}, is expected to have a positive sign. A high unemployment rate decreases the probability of finding a job, so that earning opportunities are lowered. As a result, the push factor U_{it-1} should have a positive coefficient, whereas the corresponding pull factor U_{jt-1} should have a negative one, i.e. $\beta_{3ij} > 0$ and $\beta_{6ij} < 0$ (see Table 1).

Clark, Hatton, and Williamson (2007) argue that the individual's utility can be understood as the future income stream. They conclude that if an individual expects a higher income in the destination country, the utility gain is the greatest for young individuals due to their longer working life. To put it differently, the higher the share of young people in the population is, the more people decide to move. Hence, the effect of the push factor P_{it-1} should be positive, i.e. $\beta_{4ij} > 0$. As mentioned before, the empirical analysis focuses on the migration flows from and to Germany. Thus, regarding migration inflows, the sample consists of one single destination country, Germany, and multiple source countries, the member states of the EU. On the contrary, there is a single country of origin and multiple destination countries when outflows from Germany to the EU member states are analyzed. Applying this to equation (1), induces the following two equations for inflows to and outflows from Germany, respectively:

$$\ln(m_{it}) = \beta_{1i}\ln(s_{it-1}) + \underbrace{\beta_{2i}\ln(Y_{it-1}) + \beta_{3i}\ln(U_{it-1}) + \beta_{4i}\ln(P_{it-1})}_{Push factors}$$
$$+ \underbrace{\beta_{5i}\ln(Y_{Gt-1}) + \beta_{6i}\ln(U_{Gt-1})}_{Pull factors} + \beta_{7i}T_{it} + c_i + u_{it} \qquad (2)$$

and

$$\ln(m_{jt}) = \beta_{1j}\ln(s_{jt-1}) + \underbrace{\beta_{2j}\ln(Y_{Gt-1}) + \beta_{3j}\ln(U_{Gt-1}) + \beta_{4j}\ln(P_{Gt-1})}_{Push factors}$$
$$+ \underbrace{\beta_{5j}\ln(Y_{jt-1}) + \beta_{6j}\ln(U_{jt-1})}_{Pull factors} + \beta_{7j}T_{jt} + c_j + u_{jt} \quad (3)$$

In the following, equations (2) and (3) will be refered to as the basic versions with respect to inflows and outflows, respectively.

Beneath economic and demographic factors, it is also accounted for political events which might have restrained or eased bilateral migration between Germany and the European Union. This is done by including dummy variables, denoted as d_i, which are zero until including the year of a political event and one afterwards. This implies two assumptions: first, a political event will impact migration flows with a one year delay as individuals need time to gather information before they decide to migrate. Second, the event's effects will be lasting. Beneath the direct impact of political events their interaction with push and pull factors is measured. In particular, it is analyzed whether a political occasion has strengthened or reduced the effects of these factors on migration. Following Mayda (2010), this interaction is measured by multiplying the dummy variables with push and pull factors, respectively.

Four important political changes, concerning Germany and the EU-15, are considered: The German Reunification in 1990, the Maastricht Treaty coming into force in 1993, the Schengen Agreement in 1995 and 2000, respectively, and the Introduction of the Euro in 1999. It is expected that the reunification had negative effects both on migration inflows directly and indirectly through a reduction of push and pull effects. The same effect is assumed to be true for the outflows. The Maastricht Treaty, the Schengen Agreement and the Introduction of the Euro are

expected to have positively influenced migration flows between Germany and the EU-11.

The corresponding dummy variables that account for a specific event are denoted as $d_i^{\text{Reunification}}$, $d_i^{\text{Maastricht}}$, d_i^{Schengen} and d_i^{Euro}, respectively. Table 2 lists the different political variables that are added to the basic version. For the sake of clarity, only the extensions regarding the basic version of inflows are listed. The basic version of outflows is extended analogously. In total there are four modifications of the basic version: first, dummies of the political events are added solely. In modifications two, three and four, the respective push and pull interaction terms of the GDP per capita, the unemployment rates and the share of the young population are added.

Estimation Strategy

DESTATIS (2012a) and DESTATIS (2012b), the German Federal Statistical Office, provides times series on migration flows between Germany and the EU-11 over the period from 1986 until 2011 so that there are 286 observations in total. The economic explanatory variables, i.e. real GDP per capita and the unemployment rate, are taken from OECD (2012a) and OECD (2012b). The real GDP per capita is GDP per capita in US-dollars adjusted for purchasing power parities and at constant prices with base year 2005 so that it is comparable both across countries and across time. Up to 1990, all used data refers to West Germany. Afterwards, the unified Germany is considered. One problem with the data is that the nationalities of migrants are not distinguished. Everyone who moves to and from Germany is accounted as an immigrant and emigrant, irrespectively of the citizenship. In 2011, 20% of the inflows to Germany from the EU-11 were people of German nationality. At the same time, only 31% of the outflows from Germany to a state of the EU-11 were German citizens. These numbers suggest

that the data on migration flows includes a substantial amount of people with a nationality different to that of the sending country. This might influence the coefficient estimates of the stock of immigrants residing in the destination country. The data structure is likely to be both autorcorrelated due to the long time period and cross-sectional dependent. The latter issue may arise due to the small sample of EU countries which have close economic ties. In addition, non-stationarity of the macroeconomic variables might be a problem. Using Monte Carlo simulations, **??** find that so called Common Correlated Effects mean group estimator (CCE) developed by Pesaran (2006) is robust to non-stationarity and spatial correlations. Besides, Bertoli and Moraga (2011) show that the estimation procedure of the CCE can account for multilateral resistance to migration. The CCE assumes heterogeneous slope parameters across units. In addition, both the explanatory variables and the error term of the unit specific regression equations are assumed to be influenced by unobservable common factors, whereby the impact of these factors differs across units. To control for the unobservable common factors, i.e. for the spatial dependence, panel averages of the explanatory and dependent variables are included as additional independent variables in the regression equation. This equation is estimated for each unit separately. Afterwards, the individual coefficient estimates are averaged over all panel units. However, one issue of the CCE is that the number of estimated parameters is rather high which means that the results may become imprecise in small samples. Eberhardt and Bond (2009) developed an estimator, the Augmented Mean Group estimator (AMG), which assumes the same econometric model as the CCE. Their Monte Carlo simulations indicate that the AMG performs equally well as the CCE in the presence of cross-sectional dependence and non-stationarity. The estimation procedure differs, though. First, the regression equations are set up in first differences and year dummies are included. At that stage, the observations are pooled, i.e. slope

parameters are assumed to be homogeneous across units. The estimated coefficients of the year dummies represent a common dynamic process of the variables. In a second step, this common dynamic process is included as an explicit variable in the originally regression equation. Again, the regression equation is estimated for each unit individually and the coefficients averaged across units. The AMG estimator has the advantage that just one variable, the common dynamic process, is added to the regression equation. Hence, for the estimation of migration determinants this latter estimator is employed instead of the CCE. In order to analyze if and how the "common dynamic process" of the AMG influences estimation results, the estimates of the AMG are compared to those of OLS with Newey-West standard errors.

Results

Table 3 in the appendix displays results of both the averaged and individual AMG coefficients plus the estimates of the OLS regression with Newey-West standard errors. The common dynamic process that is included in the regression equation during the AMG estimation procedure seems to impact migration flows substantially: its parameter estimates are significant on the individual level for the majority of countries. Comparing the AMG and OLS outcomes shows a considerable difference with regard to coefficient estimates. This is particularly true for those countries whose coefficient of the common dynamic process is significant. The considerable influence of the common dynamic process on estimation results underpins the arguments that estimation results become indeed biased if an estimator is used that is not robust to issues of cross-sectional dependence and non-stationarity. Hence, the following analysis concentrates on the estimation results of the AMG.

The stock of immigrants, which accounts for network effects, is either insignifi-

cant or even seems to have a negative effect on inflows to Germany regarding the majority of countries. These results may be due to the data structure, as the nationalities of migrants are not distinguished. The push effect of income, i.e. GDP per capita, is significant for the majority of countries. As predicted by theory the coefficients are negative. The pull effect of the German GDP per capita only has a significant influence on inflows from a few countries. However, the respective parameter estimates are all positive. This suggests that a high income in the destination country indeed pulls migrants. When the individual parameter estimates are averaged, the push and pull effect of sending and destination country's GDP per capita, respectively, are still significant. This indicates that inflows from the EU-11 react similarly to changes in income. The unemployment rate does not seem to influence migration flows to Germany. This result supports the point discussed above that labor is not efficiently adjusted within the European Area. The share of young people affects migration inflows only with respect to a few countries.

Beneath economic and demographic factors, the influence of political variables is analyzed. Thereby, both the direct effects of political events and the indirect effects via a change in push and pull effects are evaluated (see Table 4 and 5). However, the results show that none of the political events seem to have clearly changed migration patterns regarding inflows to Germany from the EU-11. The direct effect of political events measured via the inclusion of dummy variables is close to zero. In addition, the indirect effect measured by the interaction variables does also not allow for an unambiguous interpretation, either. Tables 6 to 8 display estimates regarding outflows. The comparison between estimation results of outflows with those of inflows reveals considerable differences regarding the forces that drive these flows. First, the push and pull effects of GDP per capita that are predicted by theory are clearly identified for inflows to Germany.

However, this does not apply to outflows. The coefficients of GDP have different signs across countries. Estimation results with respect to unemployment rates and the share of young people are similar for in- and outflows. The effects of unemployment rates are rather low. The parameters with respect to the share of young people are significant for several countries, their directions differ though. The main difference with respect to political events is the influence of GDP per capita interaction variables. There was no clear effect for the inflows, while estimation results of outflows suggest that the Maastricht Treaty and, particularly, the Schengen Agreement strengthened push and pull effects assumed by theory. Again, the interpretations regarding the influence of political events have to be seen with caution, as it is not controlled for events in the EU-11 countries.

Conclusion

Within the European Union labor mobility is too low, which induces an inefficient allocation of resources. Particularly with regard to the demographic change and the resulting shortage of labor and skills, the improvement of labor adjustment is an important issue. This thesis aimed at the determination of forces of previous migration flows in order to identify potentials to increase mobility of European citizens. To obtain consistent results, the AMG estimator was applied to the empirical specification of migration flows. The comparison between OLS and AMG estimation results indicated substantial differences among estimation results, which confirmed that estimation results are indeed biased if the issues discussed above are not considered. In addition, the results affirmed that the adjustment of labor does not work well within the European area. Both migration inflows to Germany and outflows from Germany were not driven by the evolution of unemployment rates. In contrast, results of inflows to Germany suggest that incomes of the sending and destination country do seem to influence migration

flows. This is in line with the findings of Bonin et al. (2008, p. 8) that the prospect of a higher income is a key factor of the migration decision. There is no clear evidence that either the Maastricht Treaty or the Schengen Agreement of the Introduction of the Euro has strengthened mobility.

One aspect that has been left out in the analysis is the influence of education on migration flows. The country's level of education could not be included as there was no consistent data source. However, particularly with regard to the imminent skill shortage in the European Union it would be interesting to see whether a high level of education induces high mobility of people and consequently a more effective allocation of labor. A further intriguing topic would be the question if and how the European debt crisis has changed labor mobility within the EU. Several member states have to make massive cuts which also concern to social welfare systems. This might increase the pressure to migrate to another member state of the EU if it offers better employment opportunities.

References

Bertoli, S., and J. F.-H. Moraga (2011): "Multilateral Resistance to Migration," IZA Discussion Papers 5958, Institute for the Study of Labor (IZA).

Bonin, H., W. Eichhorst, C. Florman, M. O. Hansen, L. Skiöld, J. Stuhler, K. Tatsiramos, H. Thomasen, and K. F. Zimmermann (2008): "Report No. 19: Geographic Mobility in the European Union: Optimising its Economic and Social Benefits," IZA Research Reports 19, Institute for the Study of Labor (IZA).

Brücker, H., B. Siliverstovs, and P. Trübswetter (2003): "International Migration to Germany: Estimation of a Time-Series Model and Inference in Panel Cointegration," Discussion Papers of DIW Berlin 391, DIW Berlin, German Institute for Economic Research.

Clark, X., T. J. Hatton, and J. G. Williamson (2007): "Explaining U.S. Immigration, 1971-1998," *The Review of Economics and Statistics*, 89(2), 359–373.

Coakley, J., A.-M. Fuertes, and R. P. Smith (2006): "Unobserved heterogeneity in panel time series models," *Computational Statistics & Data Analysis*, 50(9), 2361–2380.

DESTATIS (2012a): "Bevölkerung und Erwerbstätigkeit. Vorläufige Wanderungsergebnisse. 2011," https://www.destatis.de/DE/Publikationen/Thematisch/Bevoelkerung/Wanderungen/vorlaeufigeWanderungen5127101117004.pdf?__blob=publicationFile, Accessed: August 21, 2012.

——— (2012b): "Wanderungen zwischen Deutschland und dem Ausland: Jahre, Staaten der Europäischen Union," https://www-genesis.destatis.de/genesis/online/logon, Accessed: August 21, 2012.

Eberhardt, M., and S. Bond (2009): "Cross-section dependence in nonstationary panel models: a novel estimator," MPRA Paper 17692, University Library of Munich, Germany.

Karemera, D., V. I. Oguledo, and B. Davis (2000): "A gravity model analysis of international migration to North America," *Applied Economics*, 32(13), 1745–1755.

Mayda, A. (2010): "International migration: a panel data analysis of the determinants of bilateral flows," *Journal of Population Economics*, 23(4), 1249–1274.

OECD (2012a): "Gross Domestic Product," http://stats.oecd.org/index.aspx?queryid=556#, Accessed: August 21, 2012.

——— (2012b): "Labour Force Statistics by Sex and Age," http://stats.oecd.org/Index.aspx?DataSetCode=LFS_SEXAGE_I_R, Accessed: August 21, 2012.

Ortega, F., and G. Peri (2009): "The Causes and Effects of International Migrations: Evidence from OECD Countries 1980-2005," Working Paper 14833, National Bureau of Economic Research.

Pedersen, P. J., M. Pytlikova, and N. Smith (2008): "Selection and network effects–Migration flows into OECD countries 1990-2000," *European Economic Review*, 52(7), 1160–1186.

Pesaran, M. H. (2006): "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure," *Econometrica*, 74(4), 967–1012.

Wooldridge, J. (2002): *Econometric Analysis of Cross Section and Panel Data*. Mit Press.

Appendix

Explanatory variables	s_{ij}	Push			Pull	
		Y_i	U_i	P_i	Y_j	U_j
Expected effect	+	−	+	+	+	−

Table 1: Expected effects of explanatory variables on migration flows

Modification	1	2	3	4
Dummy	d_i^{Event}			
Push factors		$d_i^{\text{Event}} \times \ln(Y_{it-1})$	$d_i^{\text{Event}} \times \ln(U_{it-1})$	$d_i^{\text{Event}} \times \ln(P_{it-1})$
Pull factors		$d_i^{\text{Event}} \times \ln(Y_{Gt-1})$	$d_i^{\text{Event}} \times \ln(Y_{Gt-1})$	

d_i^{Event} represents $d_i^{\text{Reunification}}$, $d_i^{\text{Maastricht}}$, d_i^{Schengen} and d_i^{Euro}, respectively

Table 2: Political variables added to the basic version of inflows

Countries		Independent variables						
		$\ln(s_{it-1})$	$\ln(Y_{it-1})$	$\ln(U_{it-1})$	$\ln(P_{it-1})$	$\ln(Y_{Gt-1})$	$\ln(U_{Gt-1})$	Common dynamic process
EU-11	AMG	0.17	**−2.14****	0.15	−0.34	**1.34***	0.02	**0.85*****
		(0.19)	**(0.87)**	(0.12)	(0.86)	**(0.76)**	(0.14)	**(0.28)**
BE	OLS	0.36	**−1.40****	**0.17*****	−0.77	0.17	**−0.25*****	-
		(0.28)	**(0.51)**	**(0.05)**	(0.48)	(0.57)	**(0.06)**	
	AMG	0.44	**−1.24****	**0.15***	−0.55	−0.12	**−0.23*****	0.05
		(0.29)	**(0.58)**	**(0.09)**	(0.68)	(0.79)	**(0.08)**	(0.10)
DK	OLS	−0.06	−7.71	−0.43	−0.78	**9.44*****	**(1.81)**	**0.64***
		(0.21)	(6.07)	(0.59)	(1.55)		**(0.33)**	
	AMG	−0.37	**−8.88****	−0.52	−2.13	**7.64*****	**0.65****	**0.73***
		(0.43)	**(3.87)**	(0.35)	(1.43)	**(2.5)**	**(0.32)**	**(0.43)**
FR	OLS	0.5 (1.16)	**−3.46****	**0.54*****	−1.39	**4.00***	**−0.60****	-
			(1.47)	**(0.17)**	(2.33)	**(1.94)**	**(0.25)**	
	AMG	0.63	**−3.64***	−0.13	−2.77	0.81	−0.33	0.84
		(1.00)	**(2.04)**	(0.63)	(3.86)	(3.32)	(0.32)	(0.64)
GR	OLS	**−3.23*****	**−5.10*****	0.18	**4.02****	1.67	**−1.13*****	-
		(1.08)	**(0.84)**	(0.49)	**(1.87)**	(1.96)	**(0.21)**	
	AMG	**−2.69*****	**−3.03****	0.49	**3.66*****	−1.64	**−1.13*****	**0.94****
		(0.79)	**(1.19)**	(0.35)	**(1.34)**	(2.38)	**(0.19)**	**(0.44)**
IR	OLS	0.03	−1.64	**0.56***	**4.12****	**7.28*****	−0.52	-
		(0.17)	(1.67)	**(0.30)**	**(1.88)**	**(1.37)**	(0.35)	
	AMG	0.15	−0.4	0.30	−0.52	**3.26****	0.04	**1.47*****
		(0.19)	(1.10)	(0.23)	(2.04)	**(1.43)**	(0.27)	**0.33**
IT	OLS	1.07	**−4.01*****	0.41	2.38	1.13	0.26	-
		(0.78)	**(1.38)**	(0.34)	(1.48)	(4.34)	(0.17)	
	AMG	0.74	**−5.95*****	0.37	**−3.62*****	−0.16	**0.31***	**2.06*****
		(0.47)	**(0.87)**	(0.24)	**(1.28)**	(1.65)	**(0.12)**	**(0.33)**
LU	OLS	−0.04	0.56	0.03	**3.78*****	**2.4****	**0.45****	-
		(0.28)	(0.49)	(0.06)	**(0.56)**	**(1.07)**	**(0.17)**	
	AMG	−0.01	0.56	0.03	**3.62*****	**2.52***	**0.42****	−0.08
		(0.28)	(0.49)	(0.09)	**(0.62)**	**(1.29)**	**(0.17)**	(0.17)
NL	OLS	**−0.74****	**−1.85****	−0.06	0.80	**3.78*****	**0.32*****	-
		(0.32)	**(0.82)**	(0.10)	(0.53)	**(0.72)**	**(0.09)**	
	AMG	**−0.76****	**−2.05***	−0.06	0.73	**4.04*****	**0.30***	−0.05
		(0.37)	**(1.24)**	(0.12)	(0.62)	**(1.25)**	**(0.18)**	(0.20)
PT	OLS	**−3.75*****	−4.22	**1.44****	**32.12*****	**9.98*****	**−1.52*****	-
		(1.24)	(4.27)	**(0.57)**	**(6.54)**	**(3.28)**	**(0.37)**	
	AMG	−0.16	−2.13	0.76	11.58	1.61	−0.21	**2.94*****
		(1.54)	(2.71)	(0.49)	(8.15)	(3.94)	(0.57)	**(1.01)**
SP	OLS	**0.59****	0.53	0.34	0.05	0.61	−0.21	-
		(0.25)	(2.12)	(0.30)	(1.26)	(0.77)	(0.16)	
	AMG	**1.04*****	1.51	0.39	−1.2	−0.55	−0.07	**0.48***
		(0.35)	(1.83)	(0.26)	(1.39)	(1.14)	(0.16)	**(0.28)**
UK	OLS	**−0.77*****	**−2.35*****	0.11	**−1.34*****	**3.53*****	**−0.24****	-
		(0.22)	**(0.69)**	(0.17)	**(0.31)**	**(0.63)**	**(0.09)**	
	AMG	0.02	**−2.41*****	−0.15	0.42	0.94	−0.03	**0.71*****
		(0.29)	**(0.63)**	(0.13)	(0.53)	(0.71)	(0.08)	**(0.14)**

Dependent variable: $\ln(m_{it})$. *Additional controls:* T_{it}, c_i. *'Common dynamic process' included as additional independent variable during AMG estimation procedure. Estimates of EU-11 are the averages of country specific coefficients. Numbers refer to the period 1986–2011. Standard errors are in brackets. For the OLS regression Newey-West standard errors are computed. * significant at 10%, ** significant at 5%, *** significant at 1%.*

Table 3: Influence of economic and demographic factors on inflows to Germany (results of OLS and AMG estimation)

Countries	Independent variables			
	$d_i^{\text{Reunification}}$	$d_i^{\text{Maastricht}}$	d_i^{Schengen}	d_i^{Euro}
EU-11	0.03	−0.06	0.05	0.00
	(0.07)	(0.07)	(0.03)	(0.04)
BE	−0.02	−0.04	**0.08****	0.03
	(0.05)	(0.05)	**(0.04)**	(0.04)
FR	**−0.54****	−0.17	0.12	0.13
	(0.21)	(0.31)	(0.18)	(0.23)
GR	−0.08	**−0.29***	−0.17	−0.14
	(0.17)	**(0.18)**	(0.12)	(0.16)
IR	**0.28****	−0.05	-	−0.09
	(0.12)	(0.14)		(0.13)
IT	−0.14	**0.34****	0.12	−0.08
	(0.09)	**(0.13)**	(0.07)	(0.09)
LU	0.08	−0.07	**0.18****	0.03
	(0.09)	(0.09)	**(0.07)**	(0.08)
NL	0.09	0.17	0.03	**0.19****
	(0.11)	(0.11)	(0.10)	**(0.08)**
PT	0.31	0.36	−0.28	**−0.38***
	(0.26)	(0.32)	(0.23)	**(0.21)**
SP	**−0.14****	−0.12	**0.09***	0.10
	(0.07)	(0.08)	**(0.05)**	(0.06)

Dependent variable: $\ln(m_{it})$. *Additional controls:* $\ln(s_{it-1})$, $\ln(Y_{it-1})$, $\ln(U_{it-1})$, $\ln(P_{it-1})$, $\ln(Y_{Gt-1})$, $\ln(U_{Gt-1})$, T_{it}, c_i. *Estimates of EU-11 are the averages of country specific coefficients. Numbers refer to the period 1986–2011. Standard errors are in brackets. * significant at 10%, ** significant at 5%, *** significant at 1%.*

Table 4: Influence of political events on inflows to Germany (results of AMG estimation)

Countries	Independent variables				
	$\ln(Y_{it-1})$	$\ln(Y_{it-1})\times$			
		$d_i^{\text{Reunification}}$	$d_i^{\text{Maastricht}}$	d_i^{Schengen}	d_i^{Euro}
EU-11	$\mathbf{-5.67}^{**}$	2.62	2.23	−0.80	0.46
	(2.51)	(4.42)	(1.71)	(2.79)	(1.2)
DK	$\mathbf{-9.74}^{**}$	−9.96	6.84	$\mathbf{15.44}^{***}$	-
	(3.82)	(6.79)	(8.27)	**(4.24)**	
GR	$\mathbf{-15.56}^{***}$	$\mathbf{23.19}^{***}$	−29.57	19.38	−0.31
	(2.88)	**5.95**	(18.19)	(20.51)	(3.29)
IT	$\mathbf{-15.44}^{***}$	$\mathbf{18.96}^{***}$	-	−12.75	6.24
	(5.47)	**(6.02)**		(11.33)	(8.65)
NL	1.58	1.03	−7.43	−2.89	$\mathbf{7.58}^{**}$
	(12.74)	(13.63)	(47.02)	(45.3)	**(3.24)**
PT	0.80	$\mathbf{-70.30}^{*}$	$\mathbf{84.16}^{*}$	$\mathbf{-29.65}^{*}$	2.85
	(2.35)	**(40.18)**	**(50.17)**	**(15.83)**	(4.92)
SP	2.50	−4.22	0.00 (-)	2.82	$\mathbf{-3.93}^{***}$
	(2.55)	(5.22)		(4.66)	**(1.37)**
UK	$\mathbf{-5.44}^{***}$	0.59	$\mathbf{3.73}^{**}$	-	-
	(1.20)	(1.42)	**(1.48)**		

Countries	Independent variables				
	$\ln(Y_{Gt-1})$	$\ln(Y_{Gt-1})\times$			
		$d_i^{\text{Reunification}}$	$d_i^{\text{Maastricht}}$	d_i^{Schengen}	d_i^{Euro}
EU-11	1.66	−2.49	−2.24	0.72	−0.45
	(1.72)	(4.4)	(1.7)	(2.87)	(1.18)
DK	1.85	9.94	−6.79	$\mathbf{-15.52}^{***}$	-
	(3.61)	(6.78)	(8.23)	**(4.26)**	
GR	$\mathbf{13.01}^{***}$	$\mathbf{-22.28}^{***}$	28.24	−18.52	0.29
	(3.64)	**(5.72)**	(17.38)	(19.58)	(3.16)
IT	$\mathbf{15.42}^{***}$	$\mathbf{-18.79}^{***}$	$\mathbf{0.03}^{***}$	12.62	−6.19
	(5.41)	**(5.98)**	**(0.01)**	(11.22)	(8.57)
NL	0.34	−1.04	7.43	2.89	$\mathbf{-7.63}^{**}$
	(12.93)	(13.66)	(47.06)	(45.35)	**(3.27)**
PT	$\mathbf{-6.67}^{*}$	$\mathbf{67.09}^{*}$	$\mathbf{-80.26}^{*}$	$\mathbf{28.26}^{*}$	-2.76
	(3.69)	**(38.31)**	**(47.88)**	**(15.08)**	(4.75)
SP	−2.29	4.12	0.00	−2.73	$\mathbf{3.87}^{***}$
	(2.37)	(5.08)	(0.01)	(4.52)	**(1.34)**
UK	$\mathbf{2.34}^{**}$	−0.54	$\mathbf{-3.65}^{**}$	-	-
	(1.11)	(1.4)	**(1.44)**		

Dependent variable: $\ln(m_{it})$. *Additional controls:* $\ln(s_{it-1})$, $\ln(U_{it-1})$, $\ln(P_{it-1})$, $\ln(U_{Gt-1})$, T_{it}, c_i. *Estimates of EU-11 are the averages of country specific coefficients. Numbers refer to the period 1986–2011. Standard errors are in brackets. * significant at 10%, ** significant at 5%, *** significant at 1%.*

Table 5: Influence of political events on inflows' push and pull effects of GDP per capita (results of AMG estimation)

Countries	Independent variables					
	$\ln(Y_{Gt-1})$	$\ln(U_{Gt-1})$	$\ln(P_{Gt-1})$	$\ln(Y_{it-1})$	$\ln(U_{it-1})$	Common dynamic process
EU-11	0.32	−0.08	0.13	−0.20	−0.01	**0.62*****
	(0.87)	(0.06)	(0.51)	(0.61)	(0.09)	**(0.15)**
BE	−0.23	0.00	**0.47****	0.12	−0.15	**0.68*****
	(0.97)	(0.09)	**(0.21)**	(0.69)	(0.12)	**(0.12)**
DK	**5.20****	**0.88*****	**2.83*****	−3.67	−0.37	**1.94*****
	(2.21)	**(0.34)**	**(0.63)**	(2.81)	(0.30)	**(0.33)**
FR	−0.16	**−0.12****	**−0.56****	0.39	**−0.21***	**0.62*****
	(0.64)	**(0.06)**	**(0.26)**	(0.52)	**(0.12)**	**(0.09)**
GR	−0.25	−0.11	**−1.35***	**2.49*****	0.34	**0.70*****
	(1.10)	(0.16)	**(0.71)**	**(0.70)**	(0.23)	**(0.23)**
IR	2.38	−0.38	0.15	1.55	0.45	**1.81*****
	(2.17)	(0.30)	(0.91)	(1.75)	(0.37)	**(0.39)**
IT	**−3.53***	−0.17	−0.67	**1.96****	−0.05	0.23
	(1.88)	(0.22)	(1.19)	**(0.78)**	(0.41)	(0.30)
LU	**2.30***	0.02	**1.78*****	**−0.95****	0.02	**0.56*****
	(1.18)	(0.13)	**(0.39)**	**(0.43)**	(0.09)	**(0.17)**
NL	**2.80*****	**−0.22****	**0.55***	**−2.22*****	0.03	**0.77*****
	(0.83)	**(0.10)**	**(0.29)**	**(0.78)**	(0.08)	**(0.14)**
PT	**−3.90****	−0.16	**−5.34*****	−0.23	0.22	**1.87*****
	(1.83)	(0.16)	**(0.69)**	(1.23)	(0.24)	**(0.33)**
SP	−1.59	0.09	**−1.76****	−1.47	−0.40	0.11
	(1.48)	(0.19)	**(0.86)**	(1.51)	(0.26)	(0.19)
UK	0.54	**0.23*****	**0.62*****	−0.63	0.01	**0.90*****
	(0.47)	**(0.06)**	**(0.18)**	(0.51)	(0.10)	**(0.09)**

Dependent variable: $\ln(m_{jt})$. *Additional controls:* T_{jt}, c_j. *'Common dynamic process' included as additional independent variable during AMG estimation procedure. Estimates of EU-11 are the averages of country specific coefficients. Numbers refer to the period 1986–2011. Standard errors are in brackets. * significant at 10%, ** significant at 5%, *** significant at 1%.*

Table 6: Influence of economic and demographic factors on outflows from Germany (results AMG estimation)

Countries	Independent variables			
	$d_i^{\text{Reunification}}$	$d_i^{\text{Maastricht}}$	d_i^{Schengen}	d_i^{Euro}
EU-11	−0.02	0.02	0.00	0.04
	(0.04)	(0.03)	(0.04)	(0.04)
GR	−0.08	0.02	0.03	**0.20****
	(0.08)	(0.12)	(0.06)	**(0.09)**
IR	**0.36***	−0.03	-	−0.05
	(0.19)	(0.28)		(0.24)
IT	0.00	0.14	**0.20*****	**0.17***
	(0.10)	(0.13)	**(0.07)**	**(0.10)**
NL	−0.04	**−0.18****	**−0.18*****	0.00
	(0.06)	**(0.08)**	**(0.06)**	(0.06)
PT	**−0.21***	**0.37****	−0.09	0.14
	(0.11)	**(0.14)**	(0.09)	(0.12)
SP	**−0.14***	0.07	**0.15****	0.08
	(0.08)	(0.10)	**(0.07)**	(0.08)

Dependent variable: $\ln(m_{jt})$. *Additional controls:* $\ln(Y_{Gt-1})$, $\ln(U_{Gt-1})$, $\ln(P_{Gt-1})$, $\ln(Y_{jt-1})$, $\ln(U_{jt-1})$, T_{jt}, c_j. *Estimates of EU-11 are the averages of country specific coefficients. Numbers refer to the period 1986–2011. Standard errors are in brackets. * significant at 10%, ** significant at 5%, *** significant at 1%.*

Table 7: Influence of political events on outflows from Germany (results of AMG estimation)

Countries	Independent variables				
	$\ln(Y_{Gt-1})$	$\ln(Y_{Gt-1})\times$			
		$d_i^{\text{Reunification}}$	$d_i^{\text{Maastricht}}$	d_i^{Schengen}	d_i^{Euro}
EU-11	0.57	3.49	3.09	**−6.28*****	0.10
	(1.42)	(2.36)	(2.27)	**(2.00)**	(0.99)
FR	−1.25	1.02	1.61	−11.17	**9.60*****
	(3.21)	(3.99)	(12.50)	(13.59)	**(2.36)**
IR	4.98	−14.07	12.46	-	**−3.83***
	(5.16)	(9.08)	(7.70)		**(1.97)**
IT	**8.43****	−5.56	**−18.92***	−0.02	**14.29***
	(3.27)	(4.82)	**(9.74)**	(0.02)	**(7.87)**
NL	**−10.54***	**11.74****	**72.19*****	**−73.06*****	−0.36
	(5.82)	**(5.80)**	**(24.12)**	**(23.70)**	(1.70)
PT	−2.71	14.34	5.20	**−13.88***	−1.30
	(2.08)	(23.22)	(28.19)	**(7.49)**	(2.64)
SP	0.05	5.77	**0.02*****	**−9.76*****	**4.72*****
	(1.56)	(3.87)	**(0.01)**	**(3.68)**	**(1.30)**
UK	−0.96	**3.44*****	**−2.43****	-	-
	(0.95)	**(1.27)**	**(1.08)**		

Countries	Independent variables				
	$\ln(Y_{it-1})$	$\ln(Y_{it-1})\times$			
		$d_i^{\text{Reunification}}$	$d_i^{\text{Maastricht}}$	d_i^{Schengen}	d_i^{Euro}
EU-11	0.25	−3.55	−2.95	**6.45*****	−0.11
	(0.57)	(2.38)	(2.39)	**(2.08)**	(1.00)
FR	1.01	−0.98	−1.61	11.27	**−9.67*****
	(3.09)	(4.01)	(12.62)	(13.72)	**(2.38)**
IR	−1.48	14.62	−12.86	-	**3.80***
	(3.11)	(9.38)	(7.96)		**(1.96)**
IT	**−6.97****	5.63	**19.17***	-	**−14.41***
	(2.84)	(4.85)	**(9.86)**		**(7.94)**
NL	**11.22****	**−11.68****	**−72.12*****	**72.97*****	0.35
	(5.72)	**(5.79)**	**(24.10)**	**(23.67)**	(1.68)
PT	−0.16	−14.98	−5.38	**14.59***	1.36
	(1.26)	(24.36)	(29.54)	**(7.86)**	(2.74)
SP	−0.06	−5.91	0 (-)	**10.05*****	**−4.81*****
	(1.08)	(3.97)		**(3.79)**	**(1.33)**
UK	0.61	**−3.43*****	**2.51****	-	-
	(0.74)	**(1.28)**	**(1.11)**		

Dependent variable: $\ln(m_{jt})$. *Additional controls:* $\ln(U_{Gt-1})$, $\ln(P_{Gt-1})$, $\ln(U_{jt-1})$, T_{jt}, c_j. *Estimates of EU-11 are the averages of country specific coefficients. Numbers refer to the period 1986–2011. Standard errors are in brackets. * significant at 10%, ** significant at 5%, *** significant at 1%.*

Table 8: Influence of political events on outflows' push and pull effects of GDP per capita (results of AMG estimation)

The Advantages of Mergers for Market Entry - A Study Case for the Brewing Industry

Benjamin Winter*

Introduction

Over the last decade, the global industry for brewing has witnessed a great number of mergers involving breweries from different parts of the world. This fact can be seen in the increasing global market share of the five leading breweries. It has increased from 25.4 percent in 2000 to 46.3 percent in 2009[1]. Today the market is being dominated by four major breweries. These are, listed by size, Anheuser-Busch Inbev, SABMiller, Heineken and Carlsberg.

The most common explanation for the merger-wave is the trend in beer consumption. For western industrialized nations, demand has been decreasing since the late 1990s, whereas developing countries such as Brazil, China or Russia have been experiencing strong growth rates in their beer consumption[2]. Facing this development, beer suppliers of the West have had to search for new markets in order to maintain growth.

This essay assumes that most breweries use mergers as a tool to gain access to

* Benjamin Winter received his degree in Economics (B. Sc.) from the University of Bonn in 2013. The present article refers to his bachelor thesis under the supervision of JProf. Dr. Eugen Kováč, which was submitted in October 2012.

[1] source: Euromonitor-International (2010)

[2] see Colen and Swinnen (2011)

new and growing markets. The gist of this essay lies in the attempt to explain why mergers seem to be the preferred tool. This attempt is important because Greenfield-Investments as well as exports might suffice to fulfil this purpose[3]. This essay will explain the advantage of mergers over the other two choices.
The basis will be given by a model introduced by Qiu and Zhou (2006). The model combines product differentiation and information assymetry with Cournot competition (simultaneous competition in quantities). This demonstrates in which way a high degree of product diffentiation creates incentives to merge and how these incentives can be increased by the presence of information assymetry. The underlying problem of the model is the merger-paradox, which was introduced by Salant, Switzer, and Reynolds (1983).
The following chapter roughly presents the merger paradox, gives an explanation for product differentiation and the way it can overcome the merger paradox. Afterwards the relevant parts of the model will be introduced so that finally the results can be applied to the brewing industry. In the end it will be possible to give an explanation for why mergers on the brewing market are the preferred tool to aquire market entry.

The merger paradox and product differentiation

As described in the introduction, the underlying problem of the model is the merger paradox, which was introduced by Salant, Switzer, and Reynolds (1983). The authors examine mergers on a Cournot market with linear demand and homogenous firms. They come to the conclusion that mergers lead to losses for the involved firms, if the market share after the merger is below 80 percent. The reason for this is the fact, that Cournot competition results in declining reaction functions. If a merger between two or more firms takes place, the involved firms

[3] for Greenfield-Investment see Raff, M., and Stähler (2009)

(in the following refered to as insiders) will decrease joint outputs to increase the market price. The output reduction leads to an increase in profits because it internalizes negative externalities, which were created through former competition. Unfortunately for the insiders, in the next step uninvolved firms (outsiders) will increase output, thereby also increasing their profits. This output increase has a negative effect on the profits of the insiders which might prevail. The merger leads to increasing profits for the outsiders, whereas the insiders might suffer from losses. This result is known as the merger paradox[4].

The described scenario assumes that all firms involved are homogenous. This homogeneity is abolished when offered products are differentiated. Product differentiation captures how strong products differ in the eyes of the consumers. Therefore it describes the willingness of a consumer to substitute a product for that of another firm. If the consumers express a high willigness to substitute products, the firms on this market find themselves in a situation of intense competition. If, however, the consumers have strong preferences towards a certain product, competition between firms is weakened. The willingness of consumers to substitute goods and the following degree of product differentiation has a direct influence on the reaction functions of the firms. The higher the degree of product differentiation, the lesser the influence of a firm's supply on the demand of its competitors. A high degree of product differentiation therefore weakens the merger paradox. The output reduction by the insiders that follows a merger leads to a smaller output increase of the outsiders if the products are differentiated. This, however, allows the positive effects of the merger to prevail. The model by Qiu and Zhou describes when this is the case.

[4] see Pepall, Richards, and Norman (1999).

Introducing the model

The model examines a market where n domestic firms compete á la Cournot and offer differentiated products. A foreign firm tries to enter the market and has two choices of how to acquire entry. It can either merge with one of the domestic firms or supply the market via exports. An assumption of the model is that the domestic firms are informed about local market demands but the foreign firm is not (information assymetry). If the foreign firm merges with one of the domestic firms, the domestic firm shares its information concerning the demand with the foreigner and they adjust joint outputs. They will, however, continue to offer differentiated products. If the foreign firm does not merge, it will lack information regarding the demand. In this case it will not be able to accurately optimize its profits. The degree of product differentiation within the model determines the market's equilibrium. Varying the degree of product differentiation, therefore, has a crucial influence on whether a merger takes place or not.

Worth mentioning is the fact that the authors present three different scenarios. The first scenario does not take information assymetry into account, meaning that the foreigner has the same information regarding local market demands as the domestic firms. The second scenario studies a case of information assymetry, the difference being that a merger is only followed by information sharing, but not by an adjustment of joint outputs. The third scenario takes all aspects into account. The firms are faced with information assymetry and a merger is followed by information sharing as well as the adjustment of joint outputs. This third case will be presented in the following chapter.

The model

As already mentioned, n domestic firms and a foreign firm compete á la Cournot. The foreign firm's index is 0, that of the domestic firms $i \in N = \{1, ..., n\}$. The set of all firms is therefore defined as $M = \{0\} \cup N$. As the firms offer differentiated products, each firm is confronted with the inverse demand function $p_i = a + \theta - q_i - bQ_{-i}$, $i \in M$. In this equation p_i and q_i represent both the price and the quantity of product i and $Q_{-i} = \sum_{j \in m,\, j \neq i} q_j$ the aggregated supply of all firms with the exception of the supply of firm i. a is presumed to be large enough so that every firm offers a positive supply. θ is a random variable with the expected value $E[\theta] = 0$ and a variance of $\sigma^2 \equiv Var[\theta] = E\left[\theta^2\right]$. θ captures the fluctuations of the demand, known only to the domestic, not, however, to the foreign firm. Finally $b \in [0, 1]$ describes the degree of product differentiation. Obviously, the products offered by the firms are homogeneous and, therefore, completely substitutable if $b = 1$. If, however, $b = 0$, the products are differentiated to a maximum, therefore allowing each firm to behave as a monopolist. The marginal costs of each firm equal zero ($c = 0$). Given the inverse demand function each firm optimizes its profits according to the following profit function:

$$\pi_i = q_i p_i = q_i(a + \theta - q_i - bQ_{-i}),\ i \in M$$

The game of the model takes place on two levels. At level one a native firm (F1) decides whether or not to merge with the foreigner (F0). In between the two levels the domestic firms observe the true value of θ. At level two all firms simultaniously compete in quantities. Level one includes two possible scenarios leading to different outcomes on level two. Whereas in *scenario 1* F1 decides not to merge, in *scenario 2* F0 and F1 merge. In the following the equlilibrium on level two in both scenarios will be calculated.

Scenario 1: First the quantity which all firms offer is calculated, if $\theta = 0$ and all firms, including the foreign firm, are aware of this. In this case all firms maximize the following optimization problem:

$$\max_{q_i} \pi_i = q_i \left(a - q_i - bQ_{-i}\right), \; i \in M$$

Since all firms are faced with the same optimization problem, solving the first order condition provides the following equilibrium:

$$q^0 = \frac{a}{2 + bn}$$

q^0 is identical to the quantity which F0 will supply under assymetric information without having merged with F1 ($q^0 = q_0^u$), where q_0^u describes this quantity. The reason for this is that F0 does not know the actual value of θ and therefore has to optimize its expected profits $E\left[\pi_0\right] \equiv \pi_0^e$ with $E\left[\theta\right] = 0$ by using the expected quantities of all domestic firms. The domestic firms anticipate this behaviour and optimize:

$$\max_{q_i} \pi_i = q_i \left(a + \theta - q_i - bQ_{-i} - bq_0^u\right), \; i \in N$$

As above, solving the first order conditions provides the domestic firm's supply:

$$q_i^u = \frac{(2 + bn - b)a + (2 + bn)\theta}{(2 + bn)(2 + bn - b)} = q^0 + \frac{\theta}{2 + bn - b}$$

Inserting q_0^u and q_i^u into the profit function provides the equilibrium profits:

$$\pi_0^u = \left(q_0^u\right)^2 + \frac{(2 - b)a\theta}{(2 + bn)(2 + bn - b)} \;\; and \;\; \pi_i^u = \left(q_i^u\right)^2, \; i \in N \tag{1}$$

Scenario 2: In this scenario F0 and F1 merge. F1 reveals its information regard-

ing the demand and the merging parties adjust joint outputs by maximizing the joint profit function:

$$\max_{q_0,q_1} \pi_{0,1} = q_0(a+\theta-bq_0-bQ_{-0})+q_1(a+\theta-q_1-bQ_{-1})$$

$$\max_{q_i} \pi_i = q_i\left[a+\theta-q_i-b(q_0+q_1+Q_{-i})\right],\ i\in\{2,...,n\}$$

The second optimization problem is that of the outsiders. Again, the equelibrium is calculated by solving the first order conditions:

$$q_0^m = q_1^m = \frac{(2-b)(a+\theta)}{2(2+bn-b^2)}\ \ and\ \ q_i^m = \frac{a+\theta}{2+bn-b^2},\ i\in\{2,...,n\}$$

$$\pi_0^m = \pi_1^m = (1+b)(q_0^m)^2,\ \ and\ \ \pi_i^m = (q_i^m)^2 \tag{2}$$

At this point comparing the quantities in the two scenarios is possible. The merger paradox roots from the fact that $q_0^m < q_0^u$ and $q_i^u < q_i^m$. The insiders reduce, the outsiders increase their output. As explained in chapter two, if products are homogenous the possibility arises that the negative effect on the insiders' profits prevails. In this case, by using the results of (1) and (2), it is possible to calculate the precice degree of product differentiation at which a merger is profitable. In order to do so, the difference $\Delta\pi \equiv (\pi_0^m+\pi_1^m)-(\pi_0^u+\pi_1^u)$ is calculated. After several steps of calculation, $\Delta\pi$ can be simplified to:

$$\Delta\pi = \frac{\theta^2 Z(n,b)}{(2+bn)^2(2+bn-b)^2} + \frac{a\theta(nb^2-2b)}{(2+bn)^2(2+bn-b)} + \frac{b^2(a+\theta)^2 Y(n,b)}{2(2+bn)^2(2+bn-b^2)^2}$$

$Y(n,b)$ is defined as $Y(n,b)\equiv n^2b^3-(3n^2-4n+4)b^2-4(n-1)b+4$ and $Z(n,b)$ as $Z(n,b)\equiv(2+bn-2b)^2-2b^2$. Taking expectation of $\Delta\pi$ leads to the following

equation:

$$E\left[\Delta\pi\right] = \frac{1}{(2+bn)^2}\left[\frac{\sigma^2 Z(n,b)}{(2+bn-b)^2} + \frac{b^2(a^2+\sigma^2)Y(n,b)}{2(2+bn-b^2)^2}\right]$$

This equation describes the expected differences in profits with and without merger. Therefore, a merger is advantageous if, and only if $E\left[\Delta\pi\right] > 0$ is the case. In this case the sum of the profits of the merging firms is larger than the sum of the expected profits of the same firms without merging. This crucially depends on the degree of product differentiation b. Indeed the existence of $b(n) \in (0,1)$ can be proved[5], fulfilling following criteria:

$$E\left[\Delta\pi\right]\begin{cases} > 0, \; for \quad b \in (0, b(n)) \\ = 0, \; for \quad b = b(n) \\ < 0, \; for \quad b \in (b(n), 1] \end{cases} \qquad and \quad \frac{db(n)}{dn} < 0$$

Put in words, $b(n)$ describes the exact degree of product differentiation at which the gainings of a merger are equal to the losses. The underlying reaction chain has already been explained in chapter two. The more competitors there are, the higher the degree of product differentiation needs to be in order to maintain the gainings of a merger.

These results allow the definition of a sub game perfect Nash-Equilibrium, which the authors capture under proposition 4. For every value of n there exists exactly one $b(n) \in (0,1)$. If $b > b(n)$ there will be no merger between F0 and F1 on level one, resulting in the equilibrium quantities of $\{q_0^u, q_1^u, ..., q_n^u\}$ in level two. If, on the other hand, $b < b(n)$, F0 and F1 merge on level one, leading to the equilibrium quantities of $\{q_0^m, q_1^m, ..., q_n^m\}$ in level two.

As was pointed out at the beginning, there are three scenarios to the model,

[5] for proof see Qiu and Zhou (2006) page 23ff.

the first being defined by symmetric information (all firms, including F0, know the exact value of θ). As a consequence, the products need to be differentiated to a higher degree, in order to keep up the gainings of a merger. Asymmetric information, therefore, creates an additional incentive to merge.

Informative value of the model regarding the brewing industry

Having introduced the model, let us focus on the market for beer. The question is, to which degree the products on the beer market are differentiated and in how far assymetric information plays a role. Most important, however, is the question, whether or not competition in quantities is suitable for the beer market.

Allison and Uhl (1964) come to the conclusion that product differentiation plays a major role for the beer market. The authors analyze how subjects evaluate several brands of beer (e.g. the taste). They pay attention to the fact that the subjects evaluate, among others, their favourite brand. In the first of two treatments, the subjects are not aware of the brand they are evaluating. The results show that the subjects' evaluation of their favourite brand in this treatment is much more moderate than in the second treatment (when they are aware of the brand). This proves that the subjects have clear preferences towards their favourite brand, which, as mentioned in chapter two, is a strong criterion for product differentiation.

In order to show the existance of asymmetric information, it is necessary to prove that the demand is subject to fluctuations. Having a look at festive seasons and bank holidays of different nations will of course help. One can safely assume that the demand for alcohol in the USA on Independence Day is higher than on other days. The same can be said of the carneval season in Germany and Brazil. In fact every nation has its special occasions leading to a higher demand in alcohol,

including beer. Domestic breweries are naturally better aquainted with these fluctuations in demand than foreign breweries.

In the next step, the role of competition in quantities for the brewing market needs to be estimated. A look at the 27 EU-countries (see graph 1 in the appendix) interestingly shows that, for example, Belgium registers an increase in beer production despite the fact that demand is decreasing[6]. This increased production might be used as exports. However, as Madsen, Pedersen, and Lund-Thompson (2010) observe, the transportation of beer comes at very high costs, making local brewing more efficient. Therefore, one can assume that the increase of production serves the purpose of quantity adjustment in order to optimize profits.

With these three results in mind (high degree of product differentiation, presence of asymmetric information, competition in quantities), it is possible to explain the merger wave on the beer market according to the model. The product differentiation of beer allows the breweries to maintain profits through merging, because the output decrease of the insiders is followed by a smaller output increase of the outsiders. Due to the presence of asymmetric information, the products do not need to be differentiated to a particularly high degree. Therefore, even though Greenfield-Investments or exports migh suffice as serving the purpose of a market entrance (as mentioned in the introduction), the advantage of a merger lies exactly in the additional profits maintained by the high degree of product differentiation.

In conclusion, the prognosis of the model is consistent with the recent developments on the beer market. A high degree of product differentiation combined with asymmetric information will stimulate merger movements on a market.

As we have seen, this is also the case for the beer market.

[6] again, see Colen and Swinnen (2011)

References

Allison, R., and K. P. Uhl (1964): "Influences of Beer Brand Identification on Taste Perception," *Journal of Marketing Research*, 2 (August), 36–9.

Beer Statistics 2010 Edition, The brewers of Europe (2013): "`http://www.brauer-bund.de/download/Archiv/PDF/statistiken/\101126%20The%20Brewers%20of%20Europe%20-%20Beer%20Statistics%202010.pdf`," *Date of request: 14.04.2013*.

Colen, L., and J. Swinnen (2011): "Beer Drinking Nations The Determinants of Global Beer Consumption," *American Association of Wine Economists*, (79), 1–39.

Madsen, E. S., K. Pedersen, and L. Lund-Thompson (2010): "M&A as a Driver of Global Competition in the Brewing Industry," *Working paper No. 11-10, ISBN 9788778825438 (print) ISBN 9788778825452 (online)*.

Pepall, L., D. Richards, and G. Norman (1999): "Industrial Organization: Contemporary Theory and Practice," *Mason: South-Western*.

Qiu, L. D., and W. Zhou (2006): "International Mergers: Incentives and Welfare," *Journal of International Economics*, 68, 38–58.

Raff, H., R. M., and F. Ståhler (2009): "The Choice of Market Entry Mode: Greenfield Investment, M and A and Joint Venture," *International Review of Economics and Finance*, 18, 3–10.

Salant, S. W., S. Switzer, and R. J. Reynolds (1983): "Losses from Horizontal Merger: The Effects of an Exogenous Change in Industry Structure on Cournot-Nash Equilibrium," *Quarterly Journal of Economics*, 98, 185–199.

Appendix

Country	2003	2004	2005	2006	2007	2008	2009
Austria	8891.0	8670.0	8785.0	8818.0	9044.0	8937.0	8728.0
Belgium	15650.0	17270.0	17274.0	18311.0	18480.0	18044.0	18008.8
Bulgaria	4388.0	4312.0	4323.0	4841.0	5298.0	5358.0	4824.0
Croatia	3680.0	3593.0	3460.0	3586.0	3696.0	3887.0	3673.0
Cyprus	376.0	378.0	381.0	368.0	386.0	399.0	355.0
Czech Republic	18548.0	18753.0	19069.0	19787.0	19897.0	19806.0	18187.0
Denmark	8352.0	8550.0	8560.0	8105.0	7604.0	6474.0	6046.0
Estonia	1037.0	1097.0	1352.0	1418.0	1413.0	1275.0	1234.0
Finland	4564.0	4617.0	4587.0	4548.0	4547.0	4470.0	4491.0
France	18132.0	16801.0	16394.0	16029.0	15094.0	14777.0	14731.0
Germany	105990.0	108366.0	107678.0	104315.0	100628.0	99910.0	98078.0
Greece	4133.0	3952.0	3978.0	4028.0	4340.0	4374.0	4177.0
Hungary	7475.5	6611.2	6842.6	7482.7	7584.0	7102.1	6347.7
Ireland	8023.0	8142.0	8969.0	9337.0	9270.0	8846.0	8041.5
Italy	13673.0	13170.0	12798.0	12818.0	13462.0	13343.0	12776.0
Latvia	1363.8	1314.9	1290.2	1414.0	1410.1	1306.8	1357.4
Lithuania	2520.0	2782.4	2915.8	2966.0	3225.0	3074.0	2794.0
Luxembourg	391.0	377.0	374.0	338.2	322.0	312.0	324.6
Malta	78.0	80.0	76.7	105.0	110.0	112.0	103.7
Netherlands	25123.7	23828.4	24560.3	26478.9	27258.8	27180.9	25376.3
Norway	2250.0	2399.0	2398.0	2496.0	2553.0	2560.0	2516.0
Poland	28622.0	30108.0	31572.0	33953.0	36895.0	37108.0	35992.0
Portugal	7349.8	7436.3	7442.6	8358.6	8191.2	8208.4	7833.0
Romania	13086.9	14535.0	15172.0	17656.0	19554.0	20640.0	17600.0
Slovakia	4670.0	4218.0	3963.0	3794.0	3683.0	3558.0	3264.0
Slovenia	2272.0	1929.0	1867.0	1892.0	1545.5	1553.2	1443.2
Spain	30670.8	31335.8	32231.7	33590.6	34343.0	33402.3	33825.1
Sweden	4192.0	4495.0	3955.7	4377.4	4427.7	4287.6	4455.4
Switzerland	3666.0	3561.0	3417.0	3494.0	3532.0	3625.0	3555.0
Turkey	8360.0	8120.0	8188.0	8020.0	8430.0	9244.0	9231.0
United Kingdom	58014.0	57449.0	56255.0	53763.0	51341.0	49469.0	45141.0
Total	389493.7	391192.7	394684.8	395095.3	403971.4	398950.2	382650.8
Total EU 27	371537.7	373519.7	377221.8	377499.3	385760.4	379634.2	363675.8

see BEER STATISTICS 2010 EDITION, The Brewers of Europe

Table 1: Beer production (1000 HL)

Economic Inequality in Germany and the Role of Household Context

Nico Pestel*

Introduction

Economic inequality has increased considerably in many Western countries over the past decades. The growing gap between rich and poor is now one of the main issues on the policy agendas around the world. The recent period of economic crisis in the aftermath of the 2008 financial market collapse in the United States has rendered issues concerning the distribution of economic resources, in general, and questions of the appropriateness of extremely high earnings, in particular, even more urgent (OECD, 2011, p. 17). Austerity measures in the context of the euro crisis have recently triggered social unrest in countries like Greece and Spain where these measures are perceived to affect the poor disproportionally. The "*Occupy Wallstreet*" movement, which presses policy makers for steps against growing social and economic inequality, has popularized the catchphrase "*We are the 99%*". Interestingly, this slogan directly refers to academic research on the increasing income share of the richest 1% of the US population, which is nowadays

* Nico Pestel received his Ph.D. in Economics from the University of Cologne in 2013. The present article refers to his dissertation (Pestel, 2013b) under supervision of Prof. Dr. Clemens Fuest, which was submitted in November 2012 and defended in January 2013. He works as a Research Associate at the Institute for the Study of Labor (IZA) in Bonn.

back to historically high levels.[1] The latter example especially shows that the *distribution* of economic resources across the population is not just a matter for public debate and policy making. On the contrary, the analysis of distribution is long since "*back in from the cold*" (Atkinson, 1997) and has turned from "*watching the grass grow*" (Aaron, 1978) to an active and relevant area of research in (public) economics.

In this paper, I will first give a brief and more general overview of recent contributions to (empirical) research on economic inequality. I will first deal with the question why economists should care about inequality in the first place. This will be followed by a brief discussion of conceptual issues regarding the concept of economic resources and an overview of the development of economic inequality in Germany. I will then focus on some specific contributions dealing with the role of the household context for the distribution of economic resources with a special focus on Germany. The two papers, that are briefly summarized and discussed, are concerned with social changes, which have altered societies in many industrialized countries over the past decades. In many ways, the distribution of economic resources did not remain unaffected by these trends. I will finish by making some concluding remarks.

Economic Research on Inequality – Why should economists care?

The public as well as policy makers are very interested in issues concerning economic inequality. In addition, there is also *scientific* interest in this topic. Salverda, Nolan, and Smeeding (2009) argue that there are mainly three reasons

[1] Figures of this trend over the last century (e.g., Piketty and Saez, 2007, pp. 147 ff.) have become widespread not only in academic journals but also in leading newspapers (New York Times, Oct. 26, 2011: "*It's Official: The Rich Get Richer*", Frankfurter Allgemeine Sonntagszeitung, Oct. 14, 2012: "*Amerika entdeckt den Klassenkampf*").

for being interested in economic inequality. First, the distribution of economic resources and factors that influence this distribution "*were central concerns at the outset of market economics*" (p. 6). In addition, according to Musgrave (1959), income redistribution is one of three functions of government activity alongside the efficient allocation of resources and macroeconomic stabilization. Hence, the distribution of resources is a key component of economic research whereby the focus of the literature has shifted from functional to personal distributions over the last decades (Goldfarb and Leonard, 2005). Second, both citizens and policy makers have strong normative feelings about inequality. Economists should, therefore, be able to provide answers to economic phenomena that are of such vital concern for agents in the political process (Atkinson and Bourguignon, 2000, p. 4). Indeed, scientific interest in income distribution has increased alongside increases in inequality (Jenkins and Micklewright, 2007). Finally, even if inequality itself were not of great interest, there are a number of important implications that come with it. For example, many economists argue that inequality is not a bad thing per se. On the contrary, inequalities in relative factor prices are fundamental to the functioning of market economies. With a special focus on labor markets, Welch (1999) emphasizes that inequalities in wages are "good" since they signal scarcities, provide incentives for investments in human capital and compensate for different job attributes. However, Welch himself states that inequality becomes "destructive" when society does not view effort as worthwhile and upward mobility is perceived unlikely or even impossible. In general, public opinion in market economies shares economists' view that absolute equality in economic outcomes is not desirable and that inequalities are, to a certain degree, not only inevitable but even necessary (Salverda, Nolan, and Smeeding, 2009, p. 7). However, if income differences are viewed as insurmountable, social cohesion as well as acceptance of market economy and even democracy are challenged

(Stiglitz, 2012).

So, is inequality "good" or "bad" at the end of the day? Atkinson (1997) argues that the normative assessment of *equity* is rather concerned with *mobility* over the lifecycle or across generations and not with cross-sectional income differences. However, there is evidence that economies with greater levels of inequality also show lower levels of mobility (Björklund and Jäntti, 1997; OECD, 2008) which can hamper equality of opportunity (Roemer, 1998) and, hence, justice. In addition, a recent strand in the literature shows that relative income positions matter for subjective well-being of individuals (Luttmer, 2005; Senik, 2005; Clark, Frijters, and Shields, 2008). Moreover, Bartels (2008) and Gilens (2012) show for the US that increasing levels of economic polarization can lead to *political* polarization. Governments seem to become more responsive to preferences of the affluent population while preferences of low and middle income earners – the majority of the electorate – are less or even not at all represented when a small fraction of the population commands a large share of economic resources (Atkinson and Piketty, 2007).[2]

To sum up, in light of these direct and indirect effects of economic inequality on various dimensions, there are good reasons why economists should carry out sound analyses of the distribution of resources across households and individuals. This should serve as a basis for explaining causes and consequences to policy makers and the wider public. Given that there is no consensus on an "optimal level of inequality", it is very difficult for decision makers to judge whether a society experiences levels of inequality that actually harm public welfare or not. An objective basis for decision making should, therefore, consider the specific causes and implications of inequality, since it is usually the result of a complex

[2]Murray (2012) cites the former US Supreme Court Justice Louis D. Brandeis (1856–1941): "*We can have democracy in this country, or we can have great wealth concentrated in the hands of a few, but we cannot have both*" (p. 1).

interaction of various contributions and determinants that are discussed in the following. Speaking with Jenkins (1995), one should know "*whodunnit*" (p. 29).

Inequality of what among whom?

When dealing with economic inequality as a research subject the question "*inequality of what among whom*" arises (Goldfarb and Leonard, 2005). The answer to the part "*among whom*" is straightforward for economists. The term *economics* dates back to the ancient Greek word *oikos* which means *household*. Hence, the essence of the economics discipline is the study of the smallest unit of individuals within an economy jointly carrying out production and consumption activities. The question of "*what*" is related to the underlying concept of economic resources and is much more complex. Analyses of inequality are typically concerned with the distribution of wages, earnings or income. However, there are "*several steps between relative factor prices and [...] disposable income among households*" (Atkinson, 2003, p. 23). The most important steps in this process are the creation of *gross market income* from various sources and all household members, the design of the government's *tax and transfer system* as well as patterns of *household formation and composition*.

Firstly, gross labor earnings make up the largest share of total household incomes and are an important driver of income inequality (Atkinson, 2008). A vast literature in labor economics deals with rising wage and earnings dispersion. Common explanations are changes in the supply and demand for skills and tasks as well as changing labor market institutions and policies. In addition, differences in wages and earnings are affected by pay differentials across gender, race, occupations or sectors. Other market incomes, from self-employment and private pensions as well as from capital and property, have also gained importance as sources of both income and inequality (Frick and Grabka, 2003, 2010; OECD, 2011). A re-

cent strand in the literature deals with the contribution of top incomes to overall inequality and shows that large shares of total pre-tax income are increasingly concentrated among the rich (Piketty and Saez, 2003; Atkinson and Piketty, 2007; Atkinson, Piketty, and Saez, 2011).

Secondly, another important determinant of household income is the tax and transfer system. Governments redistribute market incomes via income and payroll taxes as well as cash and in-kind benefits. The level of redistribution, i.e., the difference between inequality of market and disposable income, is determined by the institutional setting of the economy as well as voters' and policy makers' perceptions and preferences for redistribution from rich to poor (see McCarty and Pontusson, 2009, for an overview of the political economy of redistribution).

Finally, total disposable household income depends on the household context, i.e., the number, composition and characteristics of individuals actually forming households. For given wages and labor market conditions, gross earnings depend on the number of hours worked, while taxes paid and cash benefits received are determined by the characteristics of and the family relationships within households. Hence, the household context, which has changed tremendously over the past decades, determines the distribution of resources both within and across households in the economy. The observed distribution of disposable income is not simply a matter of mechanically applying the tax and transfer schedule to gross incomes for a given household composition, but the result of complex interactions between the market production of gross income (joint decisions on labor supply and savings) and the formation of households (marriage, cohabitation and fertility decisions, ageing and retirement), which might, in turn, be affected by incentives from the tax and transfer system.

The case of Germany

Germany has long been a country with comparatively low levels of income inequality among the OECD world. Although still exhibiting average levels of inequality, the growth in the income gap has been considerably fast since the turn of the millennium (OECD, 2008, 2011). Therefore, empirical results in the academic literature relating to inequality in Germany differ substantially, depending on the specific period under consideration. After World War II, the distribution of income was quite stable until the 1980s, inequality started growing slowly in the 1990s and growth accelerated around 2000 (Dell, 2005, 2007; Atkinson, 2007). Drivers of this trend have mainly been gross incomes, especially at the top of the distribution (Bach, Corneo, and Steiner, 2009; Fuchs-Schündeln, Krüger, and Sommer, 2010).

Special attention has been paid to the development of wage inequality and the effects of globalization, technological change and changes in wage bargaining on the labor market (Dustmann, Ludsteck, and Schönberg, 2009; Antonczyk, Fitzenberger, and Sommerfeld, 2010). In addition and in line with similar experiences in other countries, capital and property have become more important income sources, which are very unequally distributed and increasingly contribute to overall inequality (Frick and Grabka, 2003; Frässdorf, Schwarze, and Grabka, 2011). This is also reflected in the growth of wealth inequality in Germany (Hauser, 2003; Frick, Grabka, and Hauser, 2010). The reunification of East and West Germany in 1990 and the transition process of the former East afterwards has also rendered the overall distribution more unequal (Schwarze, 1996; Grabka, Schwarze, and Wagner, 1999; Biewen, 2005; Fuchs-Schündeln, Krüger, and Sommer, 2010).

Moreover, household composition has changed considerably. For example, German household size is now the second lowest among OECD countries (OECD, 2008), which has important implications for the distribution of income. While

market income inequality is relatively high in Germany, inequality in disposable income after taxes and transfers is average in international comparisons (OECD, 2008, 2011). This is mainly due to the progressive system of income taxation. Although there is some evidence that a series of reforms after 2000 have reduced the redistributive effect of the income tax, it is still characterized by a high level of progression (Corneo, 2005; Bach, Corneo, and Steiner, 2011).

In a series of reports on *poverty and richness* in Germany (Bundesregierung, 2001–2012) the German federal government regularly monitors the development of inequality across various dimensions and gives an overview of the population's attitude to distributional issues. Moreover, the government states its general assessment of the current level and future development of inequality and how it intends to address this. According to these reports, policy makers and the public in Germany are, generally, very concerned with inequality and preferences for redistribution are quite high (Alesina and Angeletos, 2005).

Hence, Germany is an interesting case for the study of inequality, since every component determining the overall distribution of economic resources described above plays an important role in this country: Market income inequality has increased substantially over the past decades, the tax and transfer system is strongly redistributive and reduces market inequality. Moreover, the population structure has distinctly changed and inequality is an important issue on the policy agenda.

The Role of Household Context – Marital sorting and labor supply

Studies on pay differentials are mainly concerned with the adequacy of and inequalities in individual earnings. However, earned income is not only determined by a worker's productivity (the wage rate) but also by the number of hours

worked, which results from labor supply coordination within households. Therefore, Pestel (2013a) extends the analysis beyond the distribution of pay across individuals to the investigation of joint couple earnings. Increases in the correlation of spouses' earnings in couple households has been interpreted as an increasing similarity of spouses in terms of earnings-related characteristics (assortative mating), which has an amplifying effect on inequality since it reduces the level of redistribution within households. Previous studies on this issue can largely be classified as accounting approaches since observed earnings distributions are compared to counterfactuals by manipulating the correlation between male and female earnings. However, the role of labor supply behavior has so far not been taken into account.

Pestel (2013a) measures the effect of non-random sorting of spouses on inequality across couple households in West Germany from 1986 to 2010 by matching couples randomly to each other and predicting counterfactual labor supply choices. This allows me to quantify the pure effect of sorting in *earnings potential* rather than observed earnings. Using German microdata as well as a behavioral microsimulation model, I find that the impact of observed sorting on earnings inequality among couples turned from slightly equalizing to slightly disequalizing in recent years, but is generally rather neutral. However, after adjusting for the effect of labor supply choices, I find that sorting in productivity has a much stronger impact on earnings inequality. This is mainly due to positive correlation in earnings potential and increases in female employment that are more concentrated in the upper part of the distribution.

From a policy maker's perspective, this result implies a trade-off between policy measures promoting female labor force participation and redistributive policies. Achieving the objective of higher female employment apparently comes at the price of higher inequality. The policy implications are ambiguous. One could

argue that government intervention is not justified here, since this specific reason for increasing inequality is the result of couples' choices. However, the growing share of dual earner couples implies a declining importance of intra family redistribution, which could potentially be substituted by government redistribution. Policy advice on how to deal with this equity-efficiency trade-off can only be based on a theoretical framework of optimal taxation of couples. This should explicitly consider the role of market and non-market production of household goods and services affecting the distribution both within and across couple households as well as the selection into cohabitation and marriage.

Household Size and the Welfare State

Increasing correlation of spouses' earnings is only one aspect of changing living arrangements and household contexts in many Western countries. More generally, structural shifts in household composition are linked to rising inequality, since the number and characteristics of individuals living together affect the distribution of economic resources due to income sharing within households.

Peichl, Pestel, and Schneider (2012), therefore, address the effect of changing household compositions on inequality in pre and post government income (after subtracting income and payroll taxes and adding benefit payments to market incomes) and, hence, pay special attention to the role of the tax and transfer system in Germany. Moreover, while many contributions in the literature deal with the important role of gross income inequality, economic well-being depends on resources that are available for current and future consumption, i.e., disposable income. We quantify the effect of changes in household composition that are accompanied by changes in employment patterns on the income distribution. The case of Germany is of special interest in this respect since the demographic development is not only characterized by an ageing population, but also by a

sharp fall in average household size. This means that economies of scales in household consumption are more and more lost.

Using German microdata, we find that the growth of the income gap between 1991 and 2007 is indeed strongly related to changes in household composition. The result for income inequality before taxes and transfers is much larger than the result for inequality in disposable incomes. This means, that the welfare state has largely compensated for inequality induced by changes in household structure.

Concluding Remarks

Growing economic inequality has recently received increasing attention. The gap between rich and poor is potentially harmful for public welfare when it exceeds a certain threshold. That is why many policy makers are concerned with increasing levels of inequality. Economists should, therefore, provide an objective basis for decision making with regard to redistributive policies. Conducting analysis of economic inequality requires a decision on the exact research subject. This is concerned with the underlying concept of economic resources as well as the extent to which the household context is involved.

In this overview, I address building blocks in the literature on economic inequality that are not fully integrated. Therefore, it is an enormous challenge to formulate "*models of the household income distribution, incorporating not only models of labour market earnings [...] and the demographic factors affecting who lives with whom*" (Jenkins and Micklewright, 2007, p. 19). Making progress on the development of such a theoretical framework of the distribution of economic resources, comprising models of earnings and income from all sources as well as models of household formation processes is left for future research. As long as such a framework does not exist, one should instead combine single pieces of the puzzle

to get closer to the overall picture of economic inequality.

One important part of the puzzle, which deserves further study, is the interplay between social and demographic changes on the one hand and the distribution of economic resources on the other hand. Secular trends of changing living arrangements are related to serious demographic transitions many Western societies will face in coming years. These changes will fundamentally reshape the workforce and society more generally. This is particularly true for Germany. Economic inequality will not remain unaffected by these foreseeable changes, but our knowledge of this nexus is still limited and we do not exactly know which role policies (should) play. Hence, future research should further address this issue.

References

Aaron, H. (1978): *Politics and the Professors*. Brookings Institutions, Washington D.C.

Alesina, A., and G.-M. Angeletos (2005): "Fairness and Redistribution," *American Economic Review*, 95(4), 960–980.

Antonczyk, D., B. Fitzenberger, and K. Sommerfeld (2010): "Rising Wage Inequality, the Decline of Collective Bargining, and the Gender Wage Gap," *Labour Economics*, 17(5), 835–847.

Atkinson, A. B. (1997): "Bringing Income Distribution in from the Cold," *Economic Journal*, 107(44), 297–321.

——— (2003): "Income Inequality in OECD Countries: Data and Explanations," CESifo Working Paper No. 881.

——— (2007): "The Long Run Earnings Distribution in Five Countries: Remarkable Stability, U, V or W?," *Review of Income and Wealth*, 53(1), 1–24.

——— (2008): *The Changing Distribution of Earnings in OECD Countries*. Oxford University Press, Oxford.

Atkinson, A. B., and F. Bourguignon (2000): "Income Distribution and Economics," in *Handbook of Income Distribution*, ed. by A. B. Atkinson, and F. Bourguignon, vol. 1, pp. 1–86. Elsevier, Amsterdam.

Atkinson, A. B., and T. Piketty (2007): *Top Incomes over the Twentieth Century*. Oxford University Press, Oxford.

Atkinson, A. B., T. Piketty, and E. Saez (2011): "Top Incomes in the Long Run of History," *Journal of Economic Literature*, 49(1), 3–71.

Bach, S., G. Corneo, and V. Steiner (2009): "From Bottom to Top: The Entire Income Distribution in Germany, 1992–2003," *Review of Income and Wealth*, 55(2), 303–330.

——— (2011): "Effective Taxation of Top Incomes in Germany," FU Berlin School of Business & Economics Discussion Paper No. 2011/18.

Bartels, L. M. (2008): *Unequal Democracy. The Political Economy of the New Gilded Age*. Princeton University Press, Princeton and Oxford.

Biewen, M. (2005): "The Covariance Structure of East and West German Incomes and its Implications for the Persistence of Poverty and Inequality," *German Economic Review*, 6(4), 103–132.

Björklund, A., and M. Jäntti (1997): "Intergenerational Income Mobility in Sweden Compared to the United States," *American Economic Review*, 87(4), 1009–1018.

Clark, Andrew, E., P. Frijters, and M. A. Shields (2008): "Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles," *Journal of Economic Literature*, 46(1), 95–144.

Corneo, G. (2005): "The Rise and likely Fall of the German Income Tax, 1958–2005," *CESifo Economic Studies*, 51, 159–186.

Dell, F. (2005): "Top Incomes in Germany and Switzerland over the Twentieth Century," *Journal of the European Economic Association*, 3(2–3), 412–421.

——— (2007): "Top Incomes in Germany Throughout the Twentieth Century: 1891–1998," in *Top Incomes over the Twentieth Century*, ed. by A. B. Atkinson, and T. Piketty, pp. 365–425. Oxford University Press, Oxford.

Dustmann, C., J. Ludsteck, and U. Schönberg (2009): "Revisiting the German Wage Structure," *Quarterly Journal of Economics*, 124(2), 843–881.

Fräßdorf, A., J. Schwarze, and M. M. Grabka (2011): "The Impact of Household Capital Income on Income Inequality – A Factor Decomposition Analysis for Great Britain, Germany and the USA," *Journal of Economic Inequality*, 9(1), 35–56.

Frick, J. R., and M. M. Grabka (2003): "Imputed Rent and Income Inequality: A Decomposition Analysis for Great Britain, West Germany and the U.S.," *Review of Income and Wealth*, 49(4), 513–537.

——— (2010): "Accounting for Imputed and Capital Income Flows in Inequality Analysis," in *Measuring Poverty, Income Inequality, and Social Exclusion*, ed. by D. Besharov, and K. Couch. Oxford University Press, Oxford.

Frick, J. R., M. M. Grabka, and R. Hauser (2010): *Die Verteilung der Vermögen in Deutschland – Empirische Analysen für Personen und Haushalte*. edition sigma, Berlin.

Fuchs-Schündeln, N., D. Krüger, and M. Sommer (2010): "Inequality Trends for Germany in the Last Two Decades: A Tale of Two Countries," *Review of Economic Dynamics*, 13(1), 103–132.

Gilens, M. (2012): *Affluence & Influence. Economic Inequality and Political Power in America*. Princeton University Press, Princeton and Oxford.

Goldfarb, R. S., and T. C. Leonard (2005): "Inequality of What among Whom?: Rival Conceptions of Distribution in the 20th Century," *Research in The History of Economic Thought and Methodology*, 23a, 75–118.

Grabka, M. M., J. Schwarze, and G. G. Wagner (1999): "How Unification and Immigration Affected the German Income Distribution," *European Economic Review*, 43(4–6), 867–878.

Hauser, R. (2003): "The Development of the Distribution of Income and Wealth in Germany – An Overview," in *Reporting on Income Distribution and Poverty. Perspectives from a German and European Point of View*, ed. by R. Hauser, and I. Becker. Springer, Berlin and Heidelberg.

Jenkins, S. P. (1995): "Accounting for Inequality Trends: Decomposition Analysis for the UK, 1971–86," *Economica*, 62(245), 29–63.

Jenkins, S. P., and J. Micklewright (2007): *Inequality and Poverty Reexamined*. Oxford University Press, Oxford.

Luttmer, E. F. P. (2005): "Neighbors As Negatives: Relative Earnings And Well-Being," *Quarterly Journal of Economics*, 120(3), 963–1002.

McCarty, N., and J. Pontusson (2009): "The Political Economy of Inequality and Redistribution," in *The Oxford Handbook of Economic Inequality*, ed. by W. Salverda, B. Nolan, and T. M. Smeeding, pp. 665–692. Oxford University Press, Oxford.

Murray, C. (2012): *Coming Apart. The State of White America, 1960–2010*. Crown Forum, New York.

Musgrave, R. A. (1959): *The Theory of Public Finance – A Study in Public Economy*. McGraw-Hill, New York.

OECD (2008): "Growing Unequal? Income Distribution and Poverty in OECD Countries," Organisation for Economic Co-Operation and Development, Paris.

——— (2011): "Divided we stand. Why Inequality keeps rising," Organisation for Economic Co-Operation and Development, Paris.

Peichl, A., N. Pestel, and H. Schneider (2012): "Does Size Matter? The Impact of Changes in Household Structure on Income Distribution in Germany," *Review of Income and Wealth*, 58(1), 118–141.

Pestel, N. (2013a): "Beyond Inequality Accounting: Marital Sorting and Labor Supply in West Germany, mimeo," .

——— (2013b): "Inequality in Germany: The Role of Household Context and the Concept of Economic Resources," Dissertation, University of Cologne.

Piketty, T., and E. Saez (2003): "Income Inequality in the United States, 1913–1998," *Quarterly Journal of Economics*, 118(1), 1–39.

——— (2007): "Income and Wage Inequality in the United States 1913–2002," in *Top Incomes over the Twentieth Century*, ed. by A. Atkinson, and T. Piketty, pp. 141–225. Oxford University Press, Oxford.

Roemer, J. E. (1998): *Equality of Opportunity*. Harvard University Press, Cambridge.

Salverda, W., B. Nolan, and T. M. Smeeding (2009): "Introduction," in *The Oxford Handbook of Economic Inequality*, ed. by W. Salverda, B. Nolan, and T. M. Smeeding, pp. 3–22. Oxford University Press, Oxford.

Schwarze, J. (1996): "How Income Inequality Changed in Germany Following Reunification: An Empirical Analysis Using Decomposable Inequality Measures," *Review of Income and Wealth*, 42(1), 1–11.

Senik, C. (2005): "Income Distribution and Well-being: What can we learn from Subjective Data?," *Journal of Economic Surveys*, 19(1), 43–63.

Stiglitz, J. E. (2012): *The Price of Inequality: How Today's Divided Society Endangers Our Future*. W. W. Norton & Company, New York/London.

Welch, F. (1999): "In Defense of Inequality," *American Economic Review Papers and Proceedings*, 89(2), 1–17.

The Long Run View: Macroeconomic History

Prof. Dr. Moritz Schularick*

Introduction

Economic history is back en vogue, and rightly so. For Joseph Schumpeter, probably the most famous economist who taught in Bonn, the importance of economic history was obvious. Schumpeter had no doubt that economic history formed part of the canon that a good economist had to master. In his *History of Economic Analysis*, he put it this way: "What distinguishes the 'scientific' economist from all the other people who think, talk, and write about economic topics is a command of techniques that we class under three heads: history, statistics, and theory."[1] Today, roughly 60 years after Schumpeter's book, it is often the practitioners of economics at international institutions and central banks – people who have to make critical economic decisions at times when the uncertainty about the 'right' model runs high – who stress the importance of economic history and bemoan the absence of historical training in many universities. Take Stanley Fischer, the former Governor of the Central Bank of Israel and previous Chief Economist of the World Bank and Deputy Managing Director of the International

* Department of Economics, Institute of Macroeconomics and Econometrics, University of Bonn, Adenauerallee 24-42, D-53113 Bonn, Germany, e-mail: moritz.schularick@uni-bonn.de, phone: +49-228-737976.

[1] Schumpeter, 1954 p. 12. A paragraph later, he even added that "of these fundamental fields, economic history – which issues into and includes present day facts – is by far the most important."

Monetary Fund. Looking back at his long and successful career in public office he summarized candidly: "I think I've learned as much from studying the history of central banking as I have from knowing the theory of central banking and I advise all of you who want to be central bankers to read the history books."[2] Or take Randy Kroszner, a former Fed Governor and professor at the University of Chicago, who took part in the critical meetings of the Federal Reserve Board at the height of the global financial crisis. He reports back from these meetings that an economic history book – Milton Friedman's and Anna Schwartz' "Monetary History of the United States" – was "the single most important piece of economic research that provided guidance to Federal Reserve Board members during the crisis" (Kroszner, 2010).

With the benefit of hindsight, it seems all too obvious that eschewing Schumpeter's original advice and, by and large, purging undergraduate and graduate education in economics from economic history was a mistake that is now slowly being reversed. The complacency with which much of modern macroeconomics treated economic stability during the so-called *Great Moderation* as the normal state of affairs might have arguably been avoided. Kevin O'Rourke recently suggested that a better historical training would have forced the discipline to acknowledge that big breaks and discontinuities in the economic process have occurred frequently in the past, and may happen again: "Zoom out, and that swan may not seem so black after all." (O'Rourke, 2013)

This brings us to the reasons why economic history is back in fashion. I see two. First, economic history and economic historians have changed. Economic historians today study history to help build better economics; the discipline is no longer just about using economics to write better history. Second, economics itself has changed. The global financial crisis has led to some critical introspection and a

[2] Cited in O'Rourke (2013)

greater willingness to take the evidence from economic history on board. This is particularly true in the fields of macroeconomics and financial economics where the crisis exposed a number of, say, blind spots that made business as usual difficult. In this sense – and borrowing from Schumpeter's rich oeuvre one last time – one could say that crises are not only "a good, cold douche for capitalism", but the recent crisis was also a refreshing shower for the discipline of macroeconomics. In this paper, I want to present two examples that illustrate why and how economic history can play a useful role for better macroeconomic thinking. Relying on recent research in the field of long run macro-financial history, including my own, the aim is to highlight why economists in particular should take the long run view from economic history seriously. As the astute reader will quickly see, both examples also speak at least indirectly to the crisis in the Eurozone. For, even in the case of the Eurozone crisis – arguably an event that is closely linked to the specific institutional structure of the European Monetary Union – some of the problems do not seem quite as unique if viewed from a historical perspective.

Global Financial Cycles

"We, Gregor 1^{st}, sovereign Prince of the State of Poyais, to his most gracious majesty George IV" – thus opened the proclamation that appointed the new Chargé d'Affaires of the State of Poyais to the United Kingdom in the year 1823. The proclamation was written by no other than Gregor MacGregor, alongside Charles Ponzi one of the most famous fraudsters in financial history. MacGregor ambitiously invented an entire country, called Poyais, for which he then raised loans from investors in the London bond market in the 1820s.[3]

The fraud was professionally planned. MacGregor catered shrewdly to the commercial instincts of English investors. A 350-page guidebook described the Cen-

[3] Sinclair (2004).

tral American territory in great detail, marveling at its natural riches, its open and friendly people and explaining convincingly why and how it managed to escape from Spanish domination despite its geographic position between Honduras and Nicaragua. A British naval officer with good knowledge of the region and apparently impeccable reputation confirmed that the territory "is excelled by no other under the influence of the British dominion."[4] In audacious fashion, the fake guidebook left no doubt that an investment in Poyais was a once in a lifetime opportunity for those investors who were smart enough to connect the dots and think quickly on their feet to make the right investment decision. The investment case was clearly borne out by economic considerations and well supported by statistical evidence:

> It has been computed that, even in the uncivilized state of the country, and independent of the native consumption, manufactured goods to the value of upwards of fifty thousand pounds pass annually into the Spanish American provinces through this territory alone, yielding, under every disadvantageous contingency, a very large profit to the adventurers; and there is no doubt that this trade, protected by a wise and liberal policy on the part of the Government of Poyais, may be carried to an extent, much beyond any calculation which can at present be formed, and it will amply remunerate those who may become interested.[5]

The end of the story was that in 1822, Gregor MacGregor actually managed to raise 200,000 pounds in the London market for his country of Poyais. In today's money, this is equivalent to about 300 Million Euros.[6] The swindle became apparent when hopeful settlers arrived on two ships in the Bay of Honduras and found out that neither the country nor its port of St. Joseph actually existed.

These were the very first days of an organized market for overseas lending in London. British investors clearly lacked sound information to distinguish good

[4] Cited in Sinclair (2004, p. 6) .
[5] Cited in Sinclair (2004, p. 7).
[6] See http://www.measuringworth.com/

risks from bad at the time. Otherwise, nobody would have bought the Poyais bond issue. But the same lack of knowledge did not prevent investors them from buying when others did. The 1820s was a decade of abundant, risk-laden appetite for investment in emerging markets. The mood of the market in the 1820s was so exuberant that MacGregor's fictitious country could borrow at virtually the same interest rate as legitimate issuers such as Chile and Peru. The voraciousness of the London market was described as follows by the *Annual Register*:

> All the gambling propensities of human nature were constantly solicited into action and crowds of individuals of every description . . . hastened to venture some portion of their property in schemes of which scarcely anything was known except the name.[7]

What the Poyais episode demonstrated was that in phases of market optimism, virtually any country – even a fictitious one – could borrow at relatively cheap interest rates. Such boom phases of foreign lending, when the "risk appetite" of international investors is very healthy, have been a common feature of international financial markets throughout modern economic history. When times are good and optimism runs high, the market is wide open even to high-risk borrowers. In bad times, good borrowers too may find it hard or even impossible to raise funds in the market. From the 1820s, to the post-WWI imbalances of the 1920s, the Asian crisis in the late 1990s, global imbalances in the 2000s and the recent experience of the Eurozone – the list of such boom and bust episodes in capital flows in modern economic history is long. In the 19th century alone, there were no less than four big waves of capital flows to developing countries – and each of them ended in financial crises and government bankruptcies in at least a handful of countries. Arguably the largest boom and bust episode in the 20th century occurred in the 1920s – and ended in the Great Depression. Back then, Germany was on the recipient end of international capital flows and, by 1935,

[7] Cited in Marichal (1989, p. 24).

99% of foreign bonds issued by German borrowers in the 1920s were in default. Carmen and Vincent Reinhart and Kenneth Rogoff have recently told the story of these boom and bust in various papers and in a best-selling book "This Time is Different." (written by Carmen Reinhart and Kenneth Rogoff).[8] Based on meticulous historical work, they documented the historical regularity with which financial crises occurred and how banking crises and sovereign debt problems have often been closely linked. Some of the salient features of the Eurozone crisis have indeed striking historical precedents. Many of the historical episodes follow a relatively simple pattern that sounds all too familiar to European observers: asset prices in the countries at the receiving end of capital flows typically appreciate during the inflow boom and the real economy often grows briskly, supporting the original investment thesis by embellishing the underlying fundamentals. Yet economic fortunes change quickly when the inflows dry up. Now the positive feedback between financial flows and the real economy goes into reverse gear. The probability of default, banking crises or currency crashes is significantly higher in the wake of a capital inflow boom, in particular in emerging markets.[9]

Reinhart and Rogoff blame this repeated short-sightedness of investors on what they call the "this-time-is-different-syndrome." From the outside it would seem that investors are making the same mistakes over and over again and are not learning from past mistakes. Yet, as Reinhart and Rogoff point out, a narrative routinely develops in the market in boom times why the lessons of past boom and bust episodes do not apply in the specific case. In other words, investors convince themselves that the standard valuation metrics do not apply and that this time things are different: because dot.com stocks and the internet are different; because China is different; because the American housing market is different, etc. Eventually, economic logic and common sense catch-up with the market euphoria

[8] Reinhart and Rogoff (2009).
[9] See Reinhart and Reinhart (2008).

and the boom deflates.

It is a popular idea among economists that a key cause of the Eurozone crisis was the presence of moral hazard, i.e., an expectation of financial market participants that governments or the European Central Bank would ride to their rescue in a crisis. That the Eurozone crisis shares some features with the much older and regular economic phenomenon of capital flow bonanzas does not invalidate the moral hazard argument. But students of financial history and historically trained policymakers know that the risk of losing one's own money does not necessarily make the mood swings of financial markets less capricious. We do not yet understand the undercurrents of such waves of optimism and pessimism in markets very well. But the recent work of Reinhart and Rogoff demonstrates that cycles of overoptimistic lending and borrowing are part and parcel of the operation of international financial markets – and taking this historical evidence seriously might eventually help to build better economic models and design better policies. One recent example is provided by the work of Hélène Rey who, in a much debated contribution to the annual Jackson Hole conference, argued that economists have to rethink the so-called 'trilemma' or 'impossible trinity' – a central tenet of international macroeconomics. While the standard view is that under conditions of capital mobility, floating exchange rates provide room for independent monetary policy, Rey (2013) argues, partly on the basis of the historical evidence, that there is much less autonomy than commonly thought. This is because co-movements in capital flows, credit growth and asset prices – global financial cycles – effectively overwhelm the insulation provided by flexible exchange rates. In this view the trilemma is in fact a dilemma and only by managing the capital account can countries pursue monetary policies independent of financial conditions in international financial markets.

Public and Private Debt

Hardly any other topic is currently as voraciously debated in the media on both sides of the Atlantic as the issue of public debt. Two opinions in particular can be encountered frequently in the media. The first is that public debt has climbed to dangerously high levels and its reduction should be a key concern of policy; the second is that reckless public borrowing is closely linked to financial crises, as evidenced by recent events in the Eurozone. It's not hard to see the background of these concerns. A simple lesson that the wider public took from the financial crisis is that high debt harbors risks. This lesson holds for foreclosed homeowners in Florida as well as for governments in Southern Europe. When circumstances suddenly change, and they sometimes do in history, levels of indebtedness that seemed perfectly fine before may turn out to be highly problematic.

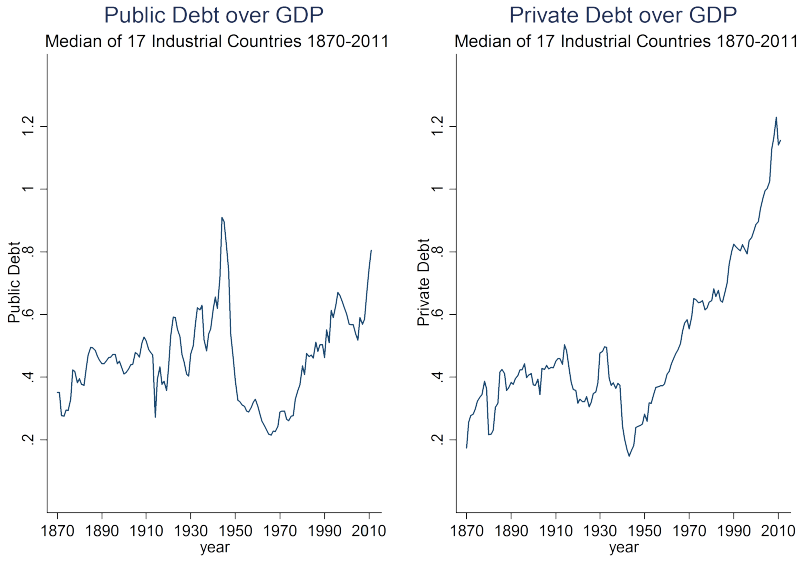
However, it is much less evident what debts one should worry about: public or private? *A priori*, most economists would intuitively point to the public sector where incentive problems of politicians might lead to reckless debt financing. Private households and companies, by contrast, are supposed to be acting in their enlightened self-interest. Many observers, often with great conviction, see the Eurozone crisis accordingly through the lens of public finance. The key source of financial fragility, in this view, was the inherent inability of governments in the so-called periphery to live within their means. It is true that in some countries, financial fragility was indeed located on the public sector balance sheet. When the economic outlook worsened, the sustainability of high public debts became an issue and doubts about the solvency of the sovereign quickly spread to banks with large holdings of government debt resulting in what Markus Brunnermeier has called a 'diabolic loop' (Brunnermeier et al., 2011).

However, on closer inspection it turns out that in 2007 Spain's public debt was below 36 percent of GDP, the overall budget was in surplus and the primary

budget balance even posted a whopping surplus of three percent of GDP. In Ireland, the figures looked similar: 25 percent for the debt ratio and a little less than 1 percent for the primary surplus. In other words, both countries operated well within the Maastricht rules and were the poster children of sound fiscal management measured by the criteria of the Maastricht treaty. None of this prevented the collapse that followed. In the space of two years, the financial systems in Spain and Ireland imploded, the economies crumbled, unemployment soared, and both countries were, albeit to different degrees, forced to seek financial help from other European countries. The lesson of this episode seems to be that there was next to nothing in key indicators of public debt that indicated the imminent catastrophe. The build-up of financial fragility occurred on private sector balance sheets. From this perspective, the rise in public debt is merely an epiphenomenon and distracts from the real source of vulnerability: namely unsustainable leverage accumulation on private balance sheets.

What does history have to say about these issues? Looking at the experience of 17 advanced Western economies – Belgium, Canada, Australia, Denmark, Finland, France, Germany, Italy, Japan, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, U.K. and U.S. – it turns out that public debt to GDP ratios have indeed increased rather markedly in most, but not all, economies in the second half of the 20th century (Figure 1). Yet one must add that it is evident from the left hand panel in figure 1 that at least until the outbreak of the global financial crisis of 2008, public debt to GDP ratios had more or less stayed within their historical peace time range.

The right hand side of the graph shows the development of private debt – proxied by bank lending to the private sector. The visual impression is stark. The break with historical trends is much more evident in the private sector where debt and leverage levels have climbed to unprecedented heights in the post WWII period.



Source: Jorda, Schularick, and Taylor (2013)

Figure 1: Public and Private Debt in the Long Run

What the consequences are of such high levels of private debt in terms of financial stability or macroeconomic risks, is not yet well understood. But the long view from history clearly suggests that it is not only the rise in public debt that should be of concern for economists, but also the apparent changes in the borrowing behavior of private agents and much higher levels of private sector leverage.

If anything, a stronger focus on the interaction between public and private sector borrowing is also reinforced by recent comparative studies looking at the behavior of public and private sector credit before and after financial crises (Jorda, Schularick, and Taylor, 2013). Simply put, this study runs a horserace between private and public borrowing as a predictor of systemic financial crises. It specifies a standard forecasting framework relating the log-odds ratio of a financial crisis event occurring in a country in a given year to lagged changes and levels of the private and public debt-to-GDP ratio. Using data for 140 years of economic

history, a surprising result is that in Western economies financial crises appear to have very little to do with public debt. Financial crises are rarely preceded by a rapid build-up of public liabilities and are not more likely at high levels of public debt than at low levels. The idea that financial crises have their roots in fiscal misbehavior is, by and large, not supported by history. Clearly, some cases exist — the recent Greek experience comes to mind – but at least in peacetime such cases have been the exception and not the rule. By contrast, high rates of private credit growth are closely associated with higher crisis probabilities. Like in Ireland and Spain today, systemic financial crises can be typically traced back to developments in the private sector (Schularick and Taylor, 2012).

It is not hard to understand why economists tend to be outspoken about the risks of public debts, but much quieter about the risks of private debt accumulation. When governments borrow, the intuition is that incentive problems abound. The temptation to finance wasteful pet projects or serve special interests at the cost of future generations is too big to be contained. Private actors, by contrast, face no similar incentive problem, at least not in the absence of moral hazard. But this might not be the only reason why economists lost sleep over of the accumulation of public, not private debts. Excessive private sector debt accumulation – as it could be observed in recent years in countries like Spain, Ireland, the US and others – raises a number of theoretically much more demanding problems. It is considerably easier to explain the political economy logic of over-exploitation of common pools or problematic incentives for re-election-hungry politicians than it is to integrate an endogenous build-up of financial fragility into modern macroeconomic models. Yet the historical evidence shows that financial crises are typically 'credit booms gone bust' implying that they are endogenous to economic developments, not simply exogenous shocks.

The overall message is once more that taking the historical insights about the

causes of financial fragility seriously can help build better theories and design better policies. To some extent, this is already happening. Based on the historical evidence that connects financial crises to private sector credit indicators, many countries and international institutions are currently debating macro prudential policies that pay special attention to private credit indicators. This shows that historical insights, derived from serious quantitative research, can inform discussions about monetary and financial policies at a time when policymakers are searching for lessons from the recent crisis.

References

Brunnermeier, M. K., L. Garicano, P. R. Lange, M. Pagano, R. Reis, T. Santos, S. Van Nieuwerburgh, and D. Vayanos (2011): "European Safe Bonds: ESBies," Euro-nomics Group.

Della Paolera, G., and A. M. Taylor (2012): "Sovereign Debt in Latin America, 1820 – 1913," National Bureau of Economic Research Working Paper w18363.

Jorda, O., M. Schularick, and A. M. Taylor (2011): "When Credit Bites Back: Leverage, Business Cycles, and Crises," National Bureau of Economic Research Working Paper 17621, (Journal of Money, Credit and Banking, forthcoming).

——— (2013): "Sovereign Versus Banks: Credit, Crises, and Consequences," National Bureau of Economic Research Working Paper 19506.

Kroszner, R. S. (2010): "Implications of the Financial Crisis for the Grand Challenge Questions for the NSF/SBE," National Science Foundation.

Marichal, C. A. (1989): *Century of Debt Crises in Latin America: From Independence to the Great Depression, 1820-1930*. Princeton University Press.

O'Rourke, K. (2013): "Why Economics Needs Economic History," The Irish Economy, http://www.irisheconomy.ie/index.php/2013/07/05/why-economics-needs-economic-history/.

Reinhart, C. M., and V. R. Reinhart (2008): "Capital Flow Bonanzas: An Encompaasing View of the Past and Present," National Bureau of Economic Research Working Paper w14321.

Reinhart, C. M., and K. S. Rogoff (2009): *This Time is Different: Eight Centuries of Financial Folly*. Princeton University Press.

Rey, H. (2013): "Dilemma not Trilemma: The Global Financial Cycle and Monetary Policy Independence," Jackson Hole Conference Paper.

Schularick, M., and A. M. Taylor (2012): "Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870 – 2008," *American Economic Review*, 102, 1029 – 1061.

Schumpeter, J. (1954): *History of Economic Analysis*. Oxford University Press.

Sinclair, D. (2004): *The Land that Never Was: Sir Gregor MacGregor and the Most Audacious Fraud in History*. Da Capo Press.

The Bonn Journal of Economics