

Board of Editors

Hua-Jing Han
Annika Wowra

Philip Hanspach

Philipp Klope

Scientific Advisory Council

Prof. Dr. D. Gärtner
Prof. Dr. J. v. Hagen
Prof. Dr. H. Hakenes
Prof. Dr. A. X. Hu

Prof. Dr. D. Krähmer
Prof. Dr. M. Kräkel
Prof. Dr. A. Kriwoluzky
JProf. Dr. M. Kuhn

Prof. Dr. K. Sandmann
Prof. Dr. M. Schularick
Prof. Dr. P. Sedlacek

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
ISSN

Bonn, Germany
2195-3449

Copyright: The Bonn Journal of Economics. All rights reserved. Every article published in The Bonn Journal of Economics is copyrighted material. No part of this publication may be reproduced, stored or transmitted in any form. Place of jurisdiction is Bonn, Germany.

Disclaimer: Every article and its contents are the responsibility of the author. It does not necessarily reflect the opinion of either the founders, the editorial board or the scientific advisory council.

THE BONN JOURNAL OF ECONOMICS

FOUNDED IN 2012



VOLUME 3

ISSUE 2

Founded by

Johannes Hermle Justus Inhoffen Tobias Ruof

DEPARTMENT OF ECONOMICS
UNIVERSITY OF BONN

THE BONN JOURNAL OF ECONOMICS

Christian Fischer

Collusion in Repeated Auctions and the Role of Communication **7**

Maja Marcus

Microfinance: A Cure for Poverty - A Macro Perspective **33**

Anna Papies

M&A Auctions and Toeholds **49**

Kaspar Zimmermann

Tax changes in the Federal Republic of Germany from 1950-1970 and their effects on economic growth **83**

Prof. Dr. Jürgen von Hagen

Sources of Fiscal Risk in the Euro Area **101**

JProf. Dr. Pia Pinger

Contexts, Compliance and Imperfect Recall - Using Two-Sample IV to Estimate Average Causal Effects of Child Human Capital Investments **119**

Collusion in Repeated Auctions and the Role of Communication

Christian Fischer*

Introduction

People of the same trade seldom meet together, even for merriment and diversion, but the conversation ends in a conspiracy against the public, or in some contrivance to raise prices. It is impossible indeed to prevent such meetings, by any law which either could be executed, or would be consistent with liberty and justice. But though the law cannot hinder people of the same trade from sometimes assembling together, it ought to do nothing to facilitate such assemblies; much less to render them necessary.

— Adam Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations* (1776)

Yet in the 18th century the Scottish philosopher Adam Smith recognized the businessman's enticement to communicate with his competitors about market share allocations and prices and to coordinate these towards outcomes that are beneficial to him and his small group of competitors but unfavourable to the large public of consumers. Until the middle of the 20th century, however, the

*Christian Fischer received his degree in Economics (M.Sc.) from the University of Bonn in 2013. The present article refers to his master thesis under supervision of Prof. Dr. Dennis Gärtner, which was submitted in September 2013. Today, Christian is affiliated with the Düsseldorf Institute for Competition Economics (DICE), Heinrich-Heine-Universität Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, email: fischer@dice.hhu.de. The author thanks Andrzej Skrzypacz for the provision of very helpful unpublished materials.

antitrust policies of our western economies were highly permissive allowing for industry associations in which firms exchanged information records, fixed prices, allocated market shares, and exchanged side-payments on a regular basis (Athey and Bagwell, 2001, p. 428).

The modern antitrust policy of many countries developed to be highly antagonistic to collusive practices. Recent examples are the U.S. Antitrust Division's Revised Amnesty Program from 1993 providing firms with incentives to self-report collusive conduct, and the United Kingdom's Competition Act from 1998 equipping the country's competition agency with powerful policy enforcement tools (Motta, 2004, pp. 1-12). Therefore, in order to collude successfully, firms are now forced to organize their coordinating activities more secretly and to reduce mutual communication to the bare minimum.

This article shall take the emerging threats that regulation imposes on collusive activities as a motivation and study for the case of auction markets how different communication structures among bidders can affect prices, and thus the extent to which collusion remains possible and effective. In recent years auctions have developed as an extremely popular way to allocate goods and resources. As Klemperer (2002) states, governments became keen to use them, e.g. to sell mobile-phone licenses, operate decentralized electricity markets, or privatize companies, and many business-to-business transactions which previously were negotiated bilaterally are now priced via auctions.

In this paper a particular focus is given to auctions that take place repeatedly over time, and thus to collusion that results from frequent interaction. As will become clear in the following, one highlight resulting from repetition is that side-payments between bidders can become unnecessary to sustain collusion. Under these circumstances, even with modern legislation at hand, competition agencies still face a difficult task to verify collusion and enforce punishments effectively.

Rather than providing a complete overview of the literature on collusion in repeated auctions, the following section contrasts the findings of models in the field that are seminal with respect to the respectively employed communication structure. In the ensuing main part of the paper I focus on the special case of tacit collusion from Skrzypacz and Hopenhayn (2004). The authors prove existence of an upper-bound on the collusive profits of bidders strictly below their first-best, given that side-payments among bidders are not possible. Their article also proposes a chips mechanism under which the payoffs of this collusive upper-bound can be approximated. I propose an algorithm that generally solves this mechanism and offer a very intuitive calibration in the context of the standard auction framework of Skrzypacz and Hopenhayn (2004). The calibration is tested with a parameter simulation and I find that it closely approximates the payoff upper-bound.

Seminal results on collusion with and without bidder communication

This section presents the results of seminal papers, that vary the assumptions on communication and information exchange among bidders. The case of tacit collusion is illustrated with the model of a repeated standard auction by Skrzypacz and Hopenhayn (2004). The outline of this model is fairly extensive because it provides the basis for my results on the chips mechanism. In a model of repeated Bertrand competition, Athey and Bagwell (2001) explore both cases – where firms are allowed to engage in communication and where this possibility is eliminated.¹ Hörner and Jamison (2007) apply the model setup of Athey and Bagwell (2001) to the case where behaviour is conditioned only on private histories and there-

¹The discussion of a Bertrand model in the auction context is justified because first price auction and Bertrand model are outcome-equivalent for the case of unit demand (Tirole, 1988, p. 364).

fore consider the extreme case, where there is no information exchange among firms at all. Additionally, the model of Aoyagi (2003) is discussed. This paper investigates the effect that institutionalizing communication through a coordination device called the ‘center’ has on market outcomes of repeated auctions. The analysis goes beyond the other papers by allowing values to be affiliated.

Skrzypacz and Hopenhayn (2004)

For a setting of repeated auctions, Skrzypacz and Hopenhayn (2004) analyse the effect that the dispense with any kind of communication between bidders has on the attainable collusive profits. The term ‘repeated auction’ expresses, that in every stage game of a repeated game an auction is played. According to Skrzypacz and Hopenhayn (2004), an analysis of collusion in auctions without communication between bidders is relevant due to several reasons: First, bidder communication before the auction is often illegal, and thus there exists a strong incentive to avoid it. And second, communication might be impractical, e.g. because the breaks between the stage games are too short. Beyond communication, the model restricts the available market information to a minimum. The public history of the repeated game only consists of the identities of the winners, which get revealed by the seller after the respective stage game. The authors focus their analysis on the effects on collusion resulting from the dynamic character of the game and the related inter-temporal incentive structure, and therefore exclude the possibility of side-payments among bidders from the model. Skrzypacz and Hopenhayn (2004) show, that for this framework, *efficient collusion* is not feasible, where efficiency means joint profit maximization. In particular, they show that for all discount factors $\delta < 1$, the average per-period payoffs of the cartel are uniformly bounded away from the payoffs of an efficient collusive scheme.²

²Skrzypacz and Hopenhayn (2004) show this result in Proposition 1 for the case of two bidders. The result has been extended to the case of N bidders by Blume and Heidhues (2001).

With the help of the chips mechanism that I implement in this paper it can be shown that there exist collusive schemes under which this upper-bound on payoffs can be approximated. In the following, the setup of Skrzypacz and Hopenhayn (2004) is summarized in further detail since my results on the chips mechanism rely on it.

In every stage game of an infinitely repeated game, one indivisible good is auctioned to a fixed set of N bidders via a standard auction that satisfies the Revenue Equivalence Theorem (RET). Assuming that the RET holds bears the implication, that the valuation v_i of every bidder $i \in N$ for the offered good is distributed independently from and identically to all other bidders in every stage game. Furthermore, all bidders have to be risk-neutral. The paper conforms with this and furthermore assumes, that all bidders not only have i.i.d. valuations within every stage game auction, but also that valuations are uncorrelated over time. Thus, the valuations of all players are drawn in every stage game from the same, stationary (and also commonly known) cumulative distribution function, $F(v)$, which is assumed to be continuously increasing and twice differentiable. The corresponding probability density function, $f(v)$, is strictly positive over the bounded interval $[0, v^h]$. The expected value of v is denoted by ev . Furthermore it is assumed that all players have a common and time-invariant discount factor δ . The reserve price of the seller is assumed to be zero. As equilibrium concept, the model uses Perfect Public Equilibrium (PPE), which goes back to Fudenberg, Levine, and Maskin (1994). Under this notion of equilibrium, players exclusively condition their strategies on the public history of the game, which in the present model only contains the identities of the winners of all preceding stage games.

All players $i \in N$ want to maximize their expected payoff from the repeated

auction, and therefore their strategy is to choose their bids in order to solve:

$$\max_{b_i \geq 0} \underbrace{v_i Q(b_i) - P(b_i)}_{\text{current period exp. payoff}} + \underbrace{\delta(Q(b_i)w_i^1 + (1 - Q(b_i))w_i^2)}_{\text{discounted exp. continuation payoff}} \quad (1)$$

$Q(b_i)$ denotes the probability of winning the current auction given bid b_i , $P(b_i)$ denotes the expected payment of i in the current auction when bidding b_i . The term w_i^1 represents the expected continuation payoffs of player i when winning the current auction, and w_i^2 the continuation payoffs when loosing it. The authors simplify the case where no auction participant bids by assuming that for this case there is a randomly chosen ‘sure bidder’, for which $b_i = 0$. The functions $Q(\cdot)$ and $P(\cdot)$ do not carry the player index i since every player i ’s valuation v_i , on which his bid is based, is drawn from the identical distribution function $f(\cdot)$.

Skrzypacz and Hopenhayn (2004) rewrite this problem to

$$\max_{b_i \geq 0} \underbrace{(v_i - \underbrace{\delta(w_i^2 - w_i^1)}_{=: c_i})Q(b_i) - P(b_i) + \delta w_i^2}_{=: \pi_i(b_i, c_i)}, \quad (2)$$

where $c_i = \delta(w_i^2 - w_i^1)$ expresses the discounted difference between the continuation payoff conditional on loosing and winning the current auction. Given this reformulation of the expected payoff, the bidding strategies can according to Skrzypacz and Hopenhayn (2004) be formulated as follows: For $v_i \in [0, c_i)$, player i will prefer not to bid (the sure bidder will choose $b_i = 0$). For $v_i \in (c_i, v^h]$, the bidding function is strictly positive and increasing in v_i . According to the authors, the reformulation to (2) implies that the repeated game can be reduced to a one-shot game with bidder valuations shifted from $[0, v^h]$ to $[-c_i, v^h - c_i]$, with δw_i^2 being a fixed, result-independent payment.

As indicated in (2), the expected payoff from the one-shot game can be summa-

alized through a payoff function π . By using the chips mechanism we will show that there exist strategy schemes that closely approximate the payoff upper-bound under tacit collusion. For this, a sensible choice of the “bid shifter”-variable c will turn out to be a crucial point of our calibration. Following Skrzypacz and Hopenhayn (2004) we will restrict the calibration on a uniform distribution of values of v_i on the $[0, 1]$ -interval. For this case, Table 1 in the Appendix provides a comparison of the payoffs achievable at the upper-bound with the case of first-best and further benchmark cases.³ Before coming to the mechanism, the results of models that vary the assumptions of Skrzypacz and Hopenhayn (2004) – in particular those on communication – will be presented.

Athey and Bagwell (2001)

Athey and Bagwell (2001) investigate collusion in an infinitely repeated Bertrand game with two firms. Side-payments that could effectuate collusive stability are not allowed. Firms have private cost realizations that are i.i.d. over time and players. This characteristic is comparable to the private and i.i.d. values in the auction model of Skrzypacz and Hopenhayn (2004). However, instead of being drawn from a continuous distribution, the model of Athey and Bagwell (2001) assumes costs to be binary – either high or low.

The basic version of their model allows for communication between firms to take place, and the stage game has the following structure. First, each of the two firms observes its realized cost type which is either high (H) or low (L). In the second step, each firm i announces its type to the other firm, $a^i \in A \equiv \{H, L, N\}$, where N stands for not announcing a type. Given the type announcements, each firm selects a price p^i and proposes its market share q^i for this round to the other firm. Finally, the market shares m^i are allocated. A stage game strategy is called

³The derivation of these values can be made available by the author upon request.

a *policy vector* consisting of a type announcement, a price choice, and a market share proposal.

In the repeated game, at the beginning of every stage, the players observe a private history containing former cost realizations, and policy vectors, and a public history that is much richer than that assumed in Skrzypacz and Hopenhayn (2004): it contains the cost announcements made by players in the former periods, the realized prices as well as the former market share proposals. Through choosing PPE as the equilibrium concept, equilibrium behaviour is only conditioned on public histories – the identical choice was made in Skrzypacz and Hopenhayn (2004).

In order to analyse the two cases, where firms communicate with each other and where they refrain from doing so, the authors concentrate their arguments on two subsets of the full set of PPE values: The first is the set of so-called *informative PPE*, where every firm always announces its cost type truthfully. The second set is that of *uninformative PPE*. In these equilibria, all firms never share any cost information with each other.

First, consider the case of informative PPE. The authors find that if the cost advantages of the low-type compared to the high-type are sufficiently large, there exists a patience level $\delta^* < 1$ such that for all $\delta \geq \delta^*$, first-best collusion can be established.⁴ A sufficient punishment for this result to hold is Nash reversion forever after a deviation occurs. This result is remarkable, as it constitutes a generalization of Fudenberg, Levine, and Maskin (1994), according to which first-best payoffs can only be achieved for $\delta \rightarrow 1$. Practically, this result implies, that given the truthful public announcement of cost levels and the observability of the realised market price, first-best outcomes can be established even for moderately impatient firms.

⁴See Athey and Bagwell (2001), Proposition 1.

Next, consider the case of uninformative PPE where no announcement of the cost-type is made in any round. Athey and Bagwell (2001) show, that with very simple strategy schemes, a lower bound on δ can be found on or above which first-best collusive results can be achieved by firms.⁵ Hence, also for the case of non-communication, first-best collusion can be established for firms with high patience levels. This result stands in stark contrast to Proposition 1 of Skrzypacz and Hopenhayn (2004), who find that collusive outcomes are bounded away from efficient collusion for any discount factor. Skrzypacz and Hopenhayn (2004) take up this point and argue that one of the reasons for this difference is the binomial character of the cost distribution in Athey and Bagwell (2001). They suggest, that as soon as there are more than two possible cost levels ('valuations' in the auction case), there exists an upper bound on collusive profits that lies below the level of efficient collusion.

Note however that the result on uninformative PPE is obtained under employment of a public history that contains the price choices of all firms in all previous periods. This is significantly more public information than available in Skrzypacz and Hopenhayn (2004), where only the identities of winners were contained in the public history. The paper of Hörner and Jamison (2007) takes up the Bertrand environment of Athey and Bagwell (2001) with two cost types and analyses uninformative collusive equilibria for the case of an empty public history, where firms condition their behaviour only on their private histories consisting of own price choices and if they made a sale in the respective period or not. Under employing the equilibrium notion of Perfect Bayesian Equilibrium, which is more general than PPE in the sense that it allows to condition behaviour on private histories, Hörner and Jamison (2007) construct a collusive scheme under which efficient collusion can be approximated arbitrarily, given that firms are

⁵In particular, see Athey and Bagwell (2001), Proposition 8.

sufficiently patient.⁶ This result, which also holds for $N > 2$, implies that the collusive upper bound resulting from non-communication and minimal public information found by Skrzypacz and Hopenhayn (2004) might have less relevance than one might expect. Through using PPE as the equilibrium standard, the possibility to condition economic behaviour on private histories gets excluded – but as Hörner and Jamison (2007) show, these might already be enough to attain fully efficient collusion. The relevance to consider private histories in the analysis of collusive equilibria is endorsed by the fact, that even for the case of more than two possible cost types, Hörner and Jamison (2007) show that first-best collusive profits can be reached by firms.

Aoyagi (2003)

Aoyagi (2003) investigates collusion in an infinitely repeated auction model with two risk-neutral bidders, where no side-payments are allowed. As in Skrzypacz and Hopenhayn (2004), values are drawn from a continuous distribution. Valuations are drawn independently over time. Within a stage auction, however, bidders' signals may be affiliated and thus might be correlated over players.

Communication in the game coordinates the bids of the two auction participants and is institutionalized in the stage game through an instruction device called the 'center'. The stage game looks as follows: At the beginning of the stage, each player i receives a signal s_i about her value. She then makes a report \hat{s}_i about her signal to the center. The center then chooses an instruction for each bidder on what bid to submit in the stage auction. All instructions made by the center can be observed by all players, as well as the bids made by the bidders. Obedience to the instruction rules can thus be observed by all players. By contrast, it cannot be observed if players' reports to the center are truthful.

⁶See Hörner and Jamison (2007), Theorem 1.

The collusive schemes considered by Aoyagi (2003) are implemented via PPE and work as follows: at the beginning of every period, the center's choice of an instruction rule is determined as a function of the public history consisting of former signal reports and former instructions made by the center, as well as the bid profiles of the preceding periods. At the beginning of each period, the players are informed publicly which instruction rule is used in that period. The center is defined as to implement 'grim-trigger' collusion, which means that the game starts in the collusive phase and instructs the players to bid according to the one-shot Nash equilibrium forever if one of the players deviates in some round in an observable manner. Deviation here signifies that a player does not bid in accordance with the center's instruction in some period.

As in Skrzypacz and Hopenhayn (2004), Aoyagi (2003) implements collusive stability through asymmetric continuation payoffs. But instead of conditioning this asymmetry on the fact of winning or losing the current stage auction, the asymmetry here is invoked through the height of the value reports made to the center. Aoyagi (2003) finds that there exist equilibria for the resulting dynamic schemes that can improve upon the payoffs from non-cooperative bidding. Excluding the possibility of value affiliation and reducing the consideration to FPA and SPA, he shows that collusion under the above communication structure can improve upon both the payoffs from non-cooperative bidding and static bid-rotation.⁷

Implementation of a chips mechanism

This section presents and elaborates on a chips mechanism under which payoffs close to the payoff upper-bound of the Skrzypacz and Hopenhayn (2004)-model can be obtained. Note, that the following implementation is for the case of a

⁷For further details, see Aoyagi (2003), Theorems 1 and 2.

second price auction (SPA) with two bidders that each have a uniform distribution of values on the $[0, 1]$ -interval. We make these simplifying assumptions because Skrzypacz and Hopenhayn (2004) use it for their calculations and we want to achieve comparability of our results. For the following elaborations it is useful to note that the bidding functions of the ‘bid-shifted’ game (2) can for a SPA be written as

$$b_i(v_i) = \begin{cases} v_i - c_i & \text{if } v_i \geq c_i \\ 0 & \text{if } v_i < c_i \end{cases}. \quad (3)$$

Since v_i is determined through the draw from an i.i.d. distribution, it is alone the choice of the bid shifters c_1 and c_2 that determines the height of the expected payoff π_i from a stage auction. We will see during the calibration of the chips mechanism that by setting the bid shifters suitably, the upper-bound payoffs can be closely approximated.

The mechanism

It follows a description of the chips mechanism as presented in Skrzypacz and Hopenhayn (2001). At the beginning of the game, each player gets the identical amount of T chips. After every auction, the winner of that auction gives one of his chips to the loser of the auction. If one player i runs out of chips, he is not allowed to participate in the following k stage auctions. After k rounds, where only the other player is allowed to bid, player i receives one chip back and he is readmitted for bidding.

Suppose there are T chips handed out to each player i at the beginning of the game. Then, for each player, there exist $(2T + 1)$ different states, characterized through the amount of chips that this player possesses in the respective state. $S = 2T + 1$ denotes the state where player i won T auctions more than his opponent (and thus has no chips left). $S = 1$ is the state where he lost T auctions

more than the other player (and thus possesses all the chips in the game). If both players won the same number of rounds, then $S = T + 1$. This state is also the initial state of the game.

Denote by $\pi(S)$ the expected payoff of the current auction in state S , by $q(S)$ the ex-ante probability of winning the stage auction in state S , and by $V(S)$ the expected sum of per-period profits in S . Skrzypacz and Hopenhayn (2001) write the chips mechanism as a set of equations, where for $S \neq \{1, 2T + 1\}$,

$$V(S) = (1 - \delta)\pi(S) + \delta(q(S)V(S + 1) + (1 - q(S))V(S - 1)) \quad (4)$$

and otherwise

$$V(1) = (1 - \delta^k)ev + \delta^k V(2) \quad (5)$$

$$V(2T + 1) = \delta^k V(2T) . \quad (6)$$

Skrzypacz and Hopenhayn (2004) solve this system of equations numerically for a uniform value distribution and optimize over the number of chips T emitted to each player, and the rounds of exclusion k . They state figures that suggest, that for discount factors at 0.9, 0.95, and 0.99, the mechanism can provide improvements over the per-period profits obtainable through simple collusive schemes. However, they do not show how they obtain their numerical results.

In the following, we develop an approach that solves the above system of equations for any number T of chips distributed to the bidders. Furthermore, a notion of state-dependent continuation payoffs is introduced, that implements the idea from the paper of making continuation payoffs conditional on winning and losing the current stage auction. Given this approach, it is shown by simulation, that the results of Skrzypacz and Hopenhayn (2004) for the uniform $[0, 1]$ - distribution of values are plausible.

Solution algorithm and state-dependent continuation payoffs

In a first step, a simple solution algorithm for the chips mechanism is developed. Suppose that a fixed number T of chips is handed out to each player. Then, in order to obtain an expression for the ex-ante average expected per-period payoff, the mechanism can be solved for $V(T+1)$ as follows. Recall from before, that $S = T+1$ is the initial state of the mechanism and therefore can be used to calculate the ex-ante per-period payoff.

Take $V(1)$ and $V(2T+1)$ and plug the former into $V(2)$, the latter into $V(2T)$. After solving for $V(2)$ and $V(2T)$ respectively, insert $V(2)$ into $V(3)$ and $V(2T)$ into $V(2T-1)$. Keep repeating the described procedure until $V(T)$ and $V(T+2)$ are reached. Plug these last two equations into:

$$V(T+1) = (1-\delta)\pi(T+1) + \delta(q(T+1)V(T+2) + (1-q(T+1))V(T))$$

Simplifying yields an explicit, non-recursive expression for the desired average per-period payoff. Figure 1 provides an example for the plug-in procedure for the simple case where $T = 1$. Here, the system ought to be solved for $S = 2$.

Next, a notion for the state-dependent continuation payoff asymmetry is introduced. With its help, state-dependent expressions for the expected stage game payoffs and respective winning probabilities can be established.

Consider the expected payoff of the current auction in state S , $\pi(S)$. For the case of a SPA with uniform $[0,1]$ -value distribution the π -function for player i can in any state S be written as

$$\pi_i(c_i, c_{-i}) = \begin{cases} \frac{1}{6} - \frac{1}{6}c_{-i}^3 + \frac{1}{4}c_i^2c_{-i} - \frac{1}{2}c_i^2 + \frac{1}{2}c_{-i} & \text{if } c_i < c_{-i} \\ \frac{1}{6} + \frac{1}{3}c_i^3 - \frac{1}{4}c_i^2c_{-i} - \frac{1}{2}c_i^2 + \frac{1}{2}c_{-i} & \text{if } c_i \geq c_{-i} \end{cases}, \quad (7)$$

where $c_i = \delta(w^2 - w^1)$ is the discounted difference of the continuation payoff

of player i conditional on losing and winning the current auction in that state.⁸ Since the present mechanism does not only have one state where both players are active bidders, but $(2T - 1)$ states where this is the case, the π -function requires for every such state a reasonable choice of the c -variable. Skrzypacz and Hopenhayn (2004) do not mention what criterion they use in order to obtain their results. In the following, we propose a state-dependent choice criterion for the c -variable. Even though it is not the result of an optimization procedure, we argue that it is economically sensible. Additionally, by employing it in our simulation of the average per-period payoffs, we obtain results that are almost identical to those of Skrzypacz and Hopenhayn (2004).

Note that the above function for the current period expected payoff is valid for $c \in [0, 1]$. Furthermore note, that for given δ , c_i increases with the difference of the continuation payoff conditional on losing and winning the current auction, $w^2 - w^1$. The mechanism is defined in a way, that with winning more and more auctions, the current state S of player i increases. Simultaneously, the state of the player $-i$ decreases. When player i reaches the state $S = 2T + 1$, he gets excluded. Hence it seems sensible to assume, that $w^2 - w^1$ increases with the state S of a player: The closer the state of exclusion comes to a player, the more valuable losing a stage auction becomes to him, since this would increase the amount of chips he has and thus reduce the risk of exclusion in a future period. Choosing $w^2 - w^1 = \frac{S}{2T+1}$ and hence

$$c_i = \delta \frac{S_i}{2T+1} \tag{8}$$

implements this idea of a c -variable, that adjusts for a state-dependent risk of exclusion and lies inside the $[0, 1]$ -interval for all S . Note, that in all states but in

⁸A derivation of this function can be found in Skrzypacz (2000). This is also true for the below-mentioned winning probabilities.

the state $(T + 1)$ the players are in different states and possess different amounts of chips. The more rounds player i wins, the more chips he has to hand over to the losing player $-i$. While the state of player i rises by winning and his continuation payoff discrepancy c_i increases in the proposed setup, that of player $-i$ decreases accordingly. Since both players are treated symmetrically, the c -variable of player $-i$ can be written as a function of the state S_i of player i as follows:

$$c_{-i} = \delta \frac{(2T + 2) - S_i}{2T + 1} \quad (9)$$

Table 2 illustrates for the case where $T = 2$, that this functional setup treats the players symmetrically in their state-dependent continuation payoffs.

The bidding function (3) implies, that with having non-identical and state-dependent values for the c -variable, the players' probabilities to win an auction in some state S , $q(S)$, differ and depend on the particular state in which the auction is executed. The probability to win a stage auction given some (c_i, c_{-i}) -tuple can for the uniform distribution on the $[0, 1]$ -interval be written as

$$q_i(c_i, c_{-i}) = \begin{cases} \frac{1}{2} + (c_{-i} - c_i)(1 - \frac{1}{2}c_{-i}) & \text{if } c_i < c_{-i} \\ \frac{1}{2} - (c_i - c_{-i})(1 - \frac{1}{2}c_i) & \text{if } c_i \geq c_{-i} \end{cases} \quad (10)$$

Inserting (8) and (9) into (7) and (10) and simplifying gives expressions for the current period expected payoff and the winning probability, that for a given amount of chips T and discount factor δ , depend solely on the state S_i that player i is in. The resulting equations, which are used in the subsequent simulation, can be found in the Appendix.

Simulation of the chips mechanism

Given the solution algorithm and the formulas for the state-dependent expected stage game payoffs and winning probabilities, the ex-ante average expected per-period payoffs can be calculated and the performance of the mechanism measured. The resulting payoff values will be comparable to those from Table 1.

For the simulation, I solved the algorithm for different numbers of distributed chips – from $T = 1$ up to $T = 4$. For the cases where $T = 1$ and $T = 2$ the derivations have been attached to the Appendix. For increasing T , the resulting equations for $V(T+1)$ explode in computational complexity and I thus decided to limit the analysis to these four cases. However, as depicted in Figure 2, the results come strikingly close to those of Skrzypacz and Hopenhayn (2004) – already within this limited T -range.

With the resulting equations for different T at hand, the average per-period payoffs can be calculated numerically for any values of δ and k . Given a payoff-maximizing choice of k , the average per-periods payoffs have been calculated over the whole range of discount factors δ . Technically, this has been achieved by choosing 1000 equally spaced data points on the open $(0, 1)$ -interval for δ and for each of them, determining the payoff-maximizing k . Figure 2 shows the optimally attainable average per-period payoff given the respective discount factor δ . This relationship has been plotted for every amount of distributed chips T separately.

Inspection of the results shows that the chips mechanism, for $T \leq 4$, allows for substantial payoff improvements over both, non-cooperative bidding and bid rotation. The analysis of the results depicted in Figure 2 reveals, that within the limited range of examined T , the attainable average expected per-period payoffs lie at ~ 0.296 for $\delta \geq 0.9$. This result implies, that under the chips mechanism more than 97% of the collusive upper bound can be realized, and correspondingly, more than 88% of the payoffs under efficient collusion. This outcome shows, that

there exist relatively simple schemes, that allow for dramatically high collusive profits in standard auctions even if communication between bidders is eliminated. Comparison with the numerical results of Skrzypacz and Hopenhayn (2004) for arbitrary T shows that my results are almost identical to theirs, already for $T \leq 4$. Note, that for some T , Skrzypacz and Hopenhayn (2004) obtain a slightly larger $V(T+1)$ at $\delta = 0.99$. Their optimality result for $\delta = 0.95$ is practically identical to ours. At $\delta = 0.90$, our implementation of the model performs slightly better.⁹ By using methods from Abreu, Pearce, and Stacchetti (1986) and Abreu, Pearce, and Stacchetti (1990) the authors claim that for arbitrary values of T , the average expected per-period payoff converges to the collusive upper bound for $\delta \rightarrow 1$. Since they do not state their solution algorithm in the paper and our algorithm experiences severe computational complexity for increasing T , we leave this result unconfirmed.

Conclusion

This article provided a discussion of collusion in repeated auctions. Theoretical results have been presented and contrasted, where the argumentation was guided by the role that communication plays for the behaviour and collusive success of bidders. The comparative study of the related literature showed that with communication being feasible, first-best profits can easily be established. In the case where bidders do not communicate at all with each other the repetition of their interaction can provide sufficient incentives to realize collusive profits close to first-best. Even though Skrzypacz and Hopenhayn (2004) show that for the case of their two-bidder standard auction collusive profits are strictly bounded away from first-best, my implementation of the chips mechanism shows that there exist simple strategy schemes that allow for very profitable, yet tacit collusion.

⁹For this comparison we use Table 2 from Skrzypacz and Hopenhayn (2004).

The insight that even under non-communication, stable and highly profitable collusion can be established provides antitrust authorities with severe challenges. Where communicative collusion can regularly be verified by resilient evidence and therefore punished, there cannot exist hard proofs of collusion in the tacit case. It is therefore difficult to find adequate ex-post measures that could disincentivize firms from colluding tacitly. In order to advance towards auction environments that minimize the opportunities for collusive conduct – possibly even independently of a specific communication format – an integrated approach that applies the insights from economic theory through smart auction design to the needs of the respective market environment seems desirable and highly necessary.

References

- ABREU, D., D. PEARCE, AND E. STACCHETTI (1986): “Optimal cartel equilibria with imperfect monitoring,” *Journal of Economic Theory*, 39(1), 251–269.
- (1990): “Toward a theory of discounted repeated games with imperfect monitoring,” *Econometrica: Journal of the Econometric Society*, 58(5), 1041–1063.
- AOYAGI, M. (2003): “Bid rotation and collusion in repeated auctions,” *Journal of Economic Theory*, 112(1), 79–105.
- ATHEY, S., AND K. BAGWELL (2001): “Optimal collusion with private information,” *RAND Journal of Economics*, 32(3), 428–465.
- BLUME, A., AND P. HEIDHUES (2001): “Tacit collusion in repeated auctions,” *Discussion paper, Wissenschaftszentrum Berlin (WZB), Research Unit: Competition and Innovation (CIG)*.
- FUDENBERG, D., D. LEVINE, AND E. MASKIN (1994): “The folk theorem with imperfect public information,” *Econometrica: Journal of the Econometric Society*, 62(5), 997–1039.
- HÖRNER, J., AND J. JAMISON (2007): “Collusion with (almost) no information,” *RAND Journal of Economics*, 38(3), 804–822.

- KLEMPERER, P. (2002): “What really matters in auction design,” *The Journal of Economic Perspectives*, 16(1), 169–189.
- MOTTA, M. (2004): *Competition policy: theory and practice*. Cambridge University Press, Cambridge.
- SKRZYPACZ, A. (2000): “Optimal Collusion in Repeated Second Price Auctions when the Seller does not Reveal the Bids,” *Mimeo, University of Rochester*.
- SKRZYPACZ, A., AND H. HOPENHAYN (2001): “Tacit collusion in repeated auctions,” *Research Paper No. 1698, Graduate School of Business, Stanford University*.
- (2004): “Tacit collusion in repeated auctions,” *Journal of Economic Theory*, 114(1), 153–169.
- SMITH, A. (1863): *An Inquiry into the Nature and Causes of the Wealth of Nations*. A. and C. Black, Edinburgh (Original work published 1776).
- TIROLE, J. (1988): *The Theory of Industrial Organization*. MIT Press.

Appendix

Tables

Type of collusive scheme	Efficient collusion	Collusive upper-bound	Non-cooperative bidding	BRS ¹⁰	Chips mechanism
Average per-period payoff for each player	0.333	0.305	0.167	0.25	0.296

Table 1: Collusive profits for schemes under a uniform distribution of values

S_i	1	2	3	4	5
$\frac{S_i}{2T+1}$	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{3}{5}$	$\frac{4}{5}$	1
$\frac{(2T+2)-S_i}{2T+1}$	1	$\frac{4}{5}$	$\frac{3}{5}$	$\frac{2}{5}$	$\frac{1}{5}$

Table 2: Illustration of the symmetric treatment in the chips mechanism for $T = 2$

Figures

$$\begin{aligned}
 V(1) &= (1 - \delta^k)ev + \delta^k V(2) \\
 V(2) &= (1 - \delta)\pi(2) + \delta(q(2)V(3) + (1 - q(2))V(1)) \\
 V(3) &= \delta^k V(2)
 \end{aligned}$$

Figure 1: Illustration of the chips solution algorithm for $T = 1$

Figure 1: Illustration of the chips solution algorithm for $T = 1$

Stage game payoffs and winning probabilities for the chips mechanism

In the following, the state-dependent stage game expected payoffs and winning probabilities for the chips mechanism are stated. The subsequent expression for the stage game expected payoff is obtained by inserting (8) and (9) into (7)

¹⁰Under this bid rotation scheme (BRS), the two bidders bid iteratively in every second round, and thus decrease the effective number of players per stage auction to one.

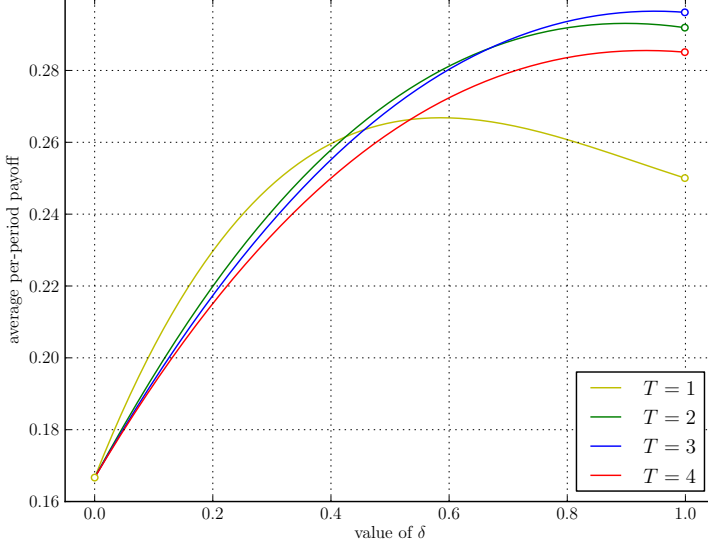


Figure 2: Avg. expected per-period payoffs under the chips mechanism for optimal k

and simplification. The result is a function of S - the parameters T and δ are exogenously given.

$$\pi_i(S_i) = \begin{cases} \frac{2(2T+1)^3 + 6\delta(2T+1)^2(2T+2-S_i) - 6\delta^2 S_i^2(2T+1) - \delta^3(S_i - 2(T+1))(S_i^2 + 8S_i(T+1) - 8(T+1)^2)}{12(2T+1)^3} & \text{if } S_i < T+1 \\ \frac{2(2T+1)^3 + 6\delta(2T+1)^2(2T+2-S_i) - 6\delta^2 S_i^2(2T+1) + \delta^3 S_i^2(7S_i - 6(T+1))}{12(2T+1)^3} & \text{if } S_i \geq T+1 \end{cases}$$

Inserting (8) and (9) into (10) and simplifying gives the probability of winning the current auction as a function of state S .

$$q_i(S_i) = \begin{cases} \frac{(2T+1)^2 - 4\delta(S_i - T - 1)(2T+1) - 2\delta^2(S_i - T - 1)(S_i - 2(T+1))}{2(2T+1)^2} & \text{if } S_i < T+1 \\ \frac{1}{2} + \frac{\delta(\delta S_i - 4T - 2)(S_i - T - 1)}{(2T+1)^2} & \text{if } S_i \geq T+1 \end{cases}$$

Applying the chips mechanism solution algorithm

The solution algorithm suggests, that for every T , the resulting equation system is solved for $V(T+1)$, which gives an expression of the ex-ante average expected per-period payoff.

For $T = 1$, the following equation system applies:

$$\begin{aligned} V(1) &= (1 - \delta^k)ev + \delta^k V(2) \\ V(2) &= (1 - \delta)\pi(2) + \delta(q(2)V(3) + (1 - q(2))V(1)) \\ V(3) &= \delta^k V(2) \end{aligned}$$

As explained above, the algorithm suggests to plug $V(1)$ and $V(3)$ into $V(2)$ and solve for $V(2)$. This yields:

$$V(2) = \frac{1}{1 - \delta^{k+1}} \left[(1 - \delta)\pi(2) + \frac{1}{2}\delta(1 - q(2))(1 - \delta^k) \right]$$

Since for $T = 1$, $q(2) = \frac{1}{2}$, this simplifies to:

$$V(2) = \frac{1}{1 - \delta^{k+1}} \left[(1 - \delta)\pi(2) + \frac{1}{4}\delta(1 - \delta^k) \right]$$

For $T = 2$, there results a system of five equations:

$$\begin{aligned} V(1) &= (1 - \delta^k)ev + \delta^k V(2) \\ V(2) &= (1 - \delta)\pi(2) + \delta(q(2)V(3) + (1 - q(2))V(1)) \\ V(3) &= (1 - \delta)\pi(3) + \delta(q(3)V(4) + (1 - q(3))V(2)) \\ V(4) &= (1 - \delta)\pi(4) + \delta(q(4)V(5) + (1 - q(4))V(3)) \\ V(5) &= \delta^k V(4) \end{aligned}$$

Plugging $V(1)$ into $V(2)$ and $V(5)$ into $V(4)$ and simplifying yields

$$\begin{aligned} V(2) &= a \left[(1 - \delta)\pi(2) + \delta(q(2)V(3) + \frac{1}{2}(1 - q(2))(1 - \delta^k)) \right] \\ V(4) &= b \left[(1 - \delta)\pi(4) + \delta(1 - q(4))V(3) \right], \end{aligned}$$

where

$$\begin{aligned} a &= \frac{1}{1 - \delta^{k+1}(1 - q(2))} \\ b &= \frac{1}{1 - \delta^{k+1}q(4)} . \end{aligned}$$

Finally, plugging these two equations into $V(3)$ and solving for $V(3)$ gives the solution

$$\begin{aligned} V(3) = c &\left\{ (1 - \delta)\pi(3) + \delta \left[q(3)b(1 - \delta)\pi(4) + (1 - q(3))a \left((1 - \delta)\pi(2) \right. \right. \right. \\ &\left. \left. \left. + \frac{1}{2}\delta(1 - q(2))(1 - \delta^k) \right) \right] \right\} , \end{aligned}$$

where

$$c = \frac{1}{1 - b\delta^2(1 - q(4))q(3) - a\delta^2q(2)(1 - q(3))} .$$

Microfinance: A Cure for Poverty - A Macro Perspective

Maja Marcus *

Introduction

In 2006 the Norwegian Nobel Committee awarded the Nobel Peace Prize to Muhammad Yunus, the “father of microfinance” with the words: “Yunus’s long-term vision is to eliminate poverty in the world. That vision can not be realized by means of micro-credit alone. But Muhammad Yunus and Grameen Bank have shown that in the continuing efforts to achieve it, microcredit must play a major part.” (The Economist, 2009; The Nobel Peace Prize, 2006). For more than thirty years microfinance has played a major part in poverty reducing policies. According to many studies, microfinance has had a tremendous, macroeconomic impact on poverty over the years. Recently, however, there has been a shift in findings and the excitement surrounding the magic of microfinance has simmered down. Critics have pointed out that many of the earlier studies have methodological and statistical errors and are therefore strongly biased (Bateman, 2011, p. 1ff.). As a significant amount of money from development aid, donations and governmental funds is being aimed towards microfinance, it is important to find out whether microfinance is in fact a cost-effective and beneficial policy for poverty reduction.

*Maja Marcus received her degree (B.Sc.) from the University of Bonn in September 2014. The present article refers to her bachelor thesis under supervision of JProf. Dr. Moritz Kuhn, which was submitted in December 2013.

With this in mind, the objective of this thesis is to identify the qualitative and quantitative impact of microfinance on poverty from a macro perspective as well as to determine the policy implications of the results of this analysis. In doing so, it will pose the overall question of whether microfinance is truly the cure for poverty as it has been deemed by earlier research.

Analyzing the current discourse on the impact of microfinance, this thesis finds that microfinance does not seem to affect poverty over the short- and medium-term. While, there is a strong correlation between microfinance and poverty, studies fail to convincingly identify causality. This thesis determines that the main reason why microfinance fails to have a significant impact is its surprising inability to promote microenterprises. The only significant effect of microfinance observed is a change in the inter-temporal consumption choices of borrowers (Duflo, Benerjee, Glennerster, and Kinnan, 2013, p. 3ff.). Yet, this thesis argues for a long-term effect of microfinance due to its likelihood of decreasing vulnerabilities to external shocks, positively impacting health as well as causing intergenerational spillovers. With this in mind, this thesis stresses the importance for policies that aim at increasing any observable effects of microfinance. One initiative proposed here is for microcredit schemes to be designed in a more sophisticated manner in order to target the specific needs of the borrower subgroups.

The thesis will be structured as follows: after introducing the prerequisites for the analysis in this thesis, two papers elaborating on the impact of microfinance with opposing views will be presented and analyzed. As a last step, policy implication will be identified. The conclusion will outline the results of this thesis and recommend possible future research proposals.

Prerequisites for an Analysis

In the last decades, microfinance has rapidly increased all over the world. At the same time, poverty has seen immense reductions in terms of global aggregates. The negative correlation between poverty and microfinance becomes quickly apparent, when scrutinizing the empirical data (compare figures 1 and 2). As microfinance has been created in an effort to reduce poverty, this thesis investigates whether their relationship extends beyond a correlation to the point of a causal linkage. In order to precisely analyze this it is vital to first define microfinance and poverty correctly.

While microfinance can describe a very broad range of basic financial services, this thesis primarily deals with microcredit in its most common form: group-lending. In a group-lending scheme, a microfinance institution (MFI) will lend money to a group of people that are lacking access to alternative means of borrowing as a result of their poverty. As the borrowers are treated as a single entity, a joint liability is induced that creates social pressure between the members to not default, making up for a lack in collateral and reducing the costs of borrowing (Todaro and Smith, 2011, p. 741f.).

The concept of poverty has many diverse dimensions and therefore definitions vary with different approaches. This thesis focuses on the definition most advantageous for quantitative comparison, which describes an individual living in poverty when his basic material needs are not being met (Hulme and Mosley, 1996, p. 105). The material well-being of an individual is often quantified by income and consumption (ibid, p.105f.). The classification of poverty occurs when income and consumption are below a standardized cut-off point such as a relative or absolute poverty line (The World Bank, 2013).

A Macroeconomic Approach by Imai et al.

The 2012 paper “Microfinance and Poverty - A Macro Perspective” by Imai, Gaiha, Thapa and Annim seeks to answer the question raised by recent research about whether microfinance truly has an impact on poverty reduction. The authors hypothesize that microfinance does reduce poverty at the macro level and verify this in the results of their study (Imai, Gaiha, Thapa, and Annim, 2012, p. 1675). Implying a very positive outlook for the significance of microfinance, the authors state that with a 10% increase in MFI loan per capita comes a 0.325% reduction in poverty (Imai, Gaiha, Thapa, and Annim, 2012, p. 1680). They conclude that microfinance is a viable poverty reduction policy and assert that recent research doubting the effect of microfinance on the macro level is highly inaccurate (Imai, Gaiha, Thapa, and Annim, 2012, p. 1684).

A Critique

With their research, Imai et al. provide promising results for the argument that microfinance leads to a decline in poverty. However, there are factors worth considering when evaluating their results.

To begin with, the authors recognized the need for an instrumental variable approach in order to deal with reverse causality. In their effort to measure the effect microfinance has on poverty, they utilize the gross loan product (GLP) of a country as an estimate for microfinance. Yet, as GLP and poverty are likely to influence each other, Imai et al. choose two instrumental variables for estimating GLP: the “cost of enforcing contract and a lag of 5-year average of gross loan portfolio weighted by the number of MFIs for each country” (Imai, Gaiha, Thapa, and Annim, 2012, p. 1677). However, it is questionable whether their choice in instruments eliminates the problem of reverse causality. For instance, the cost of enforcing contracts is a strong indicator for the economic institutions of a country.

Countries with higher income and less poverty can afford better institutions. Moreover, studies have shown that better institutions facilitate development and poverty reduction in a country (Todaro and Smith, 2011, p. 84ff.). Therefore, it cannot necessarily be assumed that the cost of enforcing contracts is not, to some extent, directly or through omitted variables cause or effect of poverty. This casts doubt on the causality implied in the results of this study.

Furthermore, it must be noted that the time frame used for the evaluation of this study, the years 2003-2007 (Imai, Gaiha, Thapa, and Annim, 2012, p. 1684), is set within a time of remarkably high growth in developing countries. Since 2000, GDP growth, consumption growth, and similar macroeconomic indicators have taken off (The Economist, 2013). This increases the possibility that there are variables omitted in the study's regressions that are the driving force behind economic growth and poverty reduction. Due to this omitted variable bias, a decline in poverty may then be falsely attributed to a rise in microfinance in their model. Taking this into account, the study of Imai et al. does not provide convincing evidence on a causal relationship between microfinance and poverty, but rather confirms their correlation.

Finally, even if one is to accept the results of this study, it shows very little insight into the economic mechanisms through which loaning to an individual or a micro-enterprise will decrease poverty on a macro scale. Microfinance could, for instance, lead to a rise in income if invested in microenterprises, the diversification of income sources and the like. It could also affect inter-temporal decision-making, smoothing consumption and decreasing vulnerability due to external shocks such as illness or natural disasters. Additionally it is possible, that microfinance has an impact on education, health or housing and will lead to positive spillover effects (Hermes and Lensink, 2011, p. 875). As a macroeconomic approach fails to achieve an understanding of the exact mechanisms that lead to poverty

reduction, the following study takes a more micro-founded approach in analyzing the effects microfinance has on poverty.

A Micro-Foundation by Duflo et al.

The paper “The Miracle of Microfinance? Evidence from a Randomized Evaluation” by Duflo, Banerjee, Glennerster and Kinnan reports on the execution and results of a randomized control trial (RCT) in which microfinance is introduced to parts of India. The objective of their study, carried out from 2005 until 2010, is to analyze the effect that microfinance has on consumption, business income, and creation, and thereby its alleged impact on poverty (Duflo, Banerjee, Glennerster, and Kinnan, 2013, p. 3ff.).

Results

As mentioned above, consumption levels are a good indicator of poverty. This implies that a rise in consumption may possibly be one channel through which microfinance could lead to poverty reduction. Consequently, Duflo et al. examine this factor closely in the evaluation of their RCT. Disappointingly, Duflo et al. could not find a macroeconomic poverty reducing effect of microfinance through increased consumption levels. However, they did detect a change in inter-temporal consumption decisions. While the level of consumption seems to be unaltered, they found a difference in composition. Spending on durable goods was significantly increased in the treatment group. At the same time, the expenses on festivals and temptation goods, such as alcohol and tobacco, were cut back. As Duflo et al. argue, this suggests that the households decreased unnecessary costs to finance the micro-loan as well as the subsequent investment on a durable good (Duflo, Banerjee, Glennerster, and Kinnan, 2013, p. 19ff.).

One other important channel through which microfinance can contribute to

poverty reduction is income. A major argument for the positive impact of microfinance is that it supposedly enables individuals to create micro- and small enterprises and to raise profitability of already existing businesses, giving them more lucrative sources of income. Hence, Duflo et al. pay close attention to this mechanism in their study, but do not find it verified in their results. The authors discover that borrowers' investments generally do not seem to translate into bigger business profits, a greater probability of becoming an entrepreneur or increased income. Duflo et al. explain this by arguing that the marginal business that begins through microfinance in an area already offering unprofitable opportunities will be even less lucrative. They found that the only considerable and significant boost in profits happened in the upper tail of previously existing businesses which were already comparably profitable and large (Duflo, Benerjee, Glennerster, and Kinnan, 2013, p. 22ff.). Taking this into consideration, the researchers' results provide important counter-evidence to much of the earlier literature claiming microfinance significantly reduces poverty through business growth.

Discussion

The results presented in the paper by Duflo et al. give much insight into the channels through which microfinance has an impact on poverty, and more importantly where it fails to have an effect. Nevertheless, this thesis argues for a more time-distinguished interpretation of their results. While the 5 year-long study of Duflo et al. gives insight into the short- and medium-term effects of microfinance on poverty, it might have neglected long-term effects that evolve over the life span of an individual or longer.

One argument for this addresses health issues of borrowers. As Duflo et al. have discovered, microfinance leads to a decline in expenditure for temptation

goods. The decreased consumption of alcohol or tobacco could have favorable long-lasting health effects for an individual. This, in turn, can decrease his lifetime health expenditures and increase his lifetime income, as studies have shown that healthier people are more productive and earn higher wages (Todaro and Smith, 2011, p. 399f.).

Another mechanism through which microfinance could have long-term effects is the possibility of decreasing the vulnerability of the poor arising from external shocks. Duflo et al. found in their study that microfinance changes inter-temporal consumption decisions. This can lead to consumption smoothing, crucial in times of an unexpected external shock or periods of cyclical downturns in order to cope with the crisis (Chowdhury, 2009, p. 8). Furthermore, Duflo et al. observe a boost in the probability that a household owns more than one business (Duflo, Benerjee, Glennerster, and Kinnan, 2013, p. 22ff.). This fact points to a diversification in income sources, which spreads the risk of being affected by a crisis. One study supporting the hypothesis that microfinance decreases vulnerability to external shocks finds microfinance acting as a recovery tool after a natural disaster (Hermes and Lensink, 2011, p. 877).

These are just some examples of the various ways through which microfinance may have an effect on the lifetime income of individuals and when aggregated on poverty on a macro scale. If microfinance does have a long-term impact, it could also potentially lead to intergenerational spillover effects. For instance, less poverty and less vulnerability in a household may positively affect the education of a child, which, in turn, will influence his future income. However, a sizable challenge to a more profound discussion on long-term effects is that an exact quantification of these on the macro scale may prove to be difficult to obtain.

Policy Implications

Taking the newest research into account, it becomes apparent that microfinance does not seem to be the cure for poverty. However, one should not neglect the aforementioned impact that microfinance does have. Regarding this, it is of value to address how any observed effects can be enhanced. While there are many suggested improvements and complimentary efforts for increasing the effectiveness of microfinance, such as various supply and demand side policies (Chowdhury, 2009, p. 2ff.), this thesis further points out the following consideration.

Targeting

Considering how fast the establishment of MFIs spread across the world, it is surprising to discover that many MFIs struggle with fairly low take-up and high dropout rates (Duflo, Benerjee, Glennerster, and Kinnan, 2013, p. 33). This points to a misunderstanding of the actual needs of borrowers and might add to an explanation of why microfinance is relatively ineffective.

One possible solution for this is a better targeting of borrowers. In the previous analysis it becomes apparent that microfinance has different effects on different levels of poverty. This stresses the importance of not treating the poor as one homogenous group, but to distinguish between the poorest of the poor, the so-called “core poor” (Hermes and Lensink, 2011, p. 876), and the relatively better off poor. For instance, the core poor, on the one hand, are much more risk averse and therefore rather invest in working capital and consumption over productive activities (Hulme and Mosley, 1996, p. 787). The relatively better off poor, on the other hand, might struggle more with transiting a microenterprise to a small and medium enterprise (SME) due to a lack of credit for SME’s (Bateman, 2011, p. 3). Addressing the specific needs of the relevant households should be a first priority for MFIs if their goal is to actively reduce poverty.

Not only the distinction of different poverty levels, but also the distinction of different employment sectors is important when considering how to increase the effects of microfinance. In her paper “Microfinance and Investment: A Comparison with Bank and Informal Lending” Lucia Dalla Pellegrina shows that MFIs have so far not been able to effectively engage the agricultural sector in borrowings. One major obstacle is that microcredit is usually given out over short periods of time, whereas the average production cycle in agriculture is much longer. The solely short-term investment possibilities are therefore not profitable for a borrower working in agriculture (Pellegrina, 2011, p. 882ff.). However, especially for the rural population, which is particularly prone to poverty (The World Bank, 2013), the agricultural sector is a key element in regard to poverty reduction. Therefore, designing microcredit schemes that are better at reaching this sector could lead to a much higher impact of microfinance.

Conclusion

The once prevalent assumption that microfinance has a significant impact on poverty from a macro perspective has recently lost some of its credibility. While discussions are still being shaped by controversial views, more evidence challenging the positive impact of microfinance is accumulating. This thesis has come to the conclusion that microfinance most likely does not have a significant effect on poverty over the short and medium run, but arguments can be made for a positive long-term influence.

Research shows that microfinance does not have a significant impact on microenterprises and thereby on disposable income. Nonetheless, it was determined that microfinance has an impact on the inter-temporal consumption choices of households. This however does not seem to translate into significant poverty reduction in the short and medium term. Yet, it is possible that microfinance

does have a macroeconomic impact on poverty in the long run, through intergenerational spillovers, decreased vulnerabilities, and health benefits for borrowers. Targeting the needs of the borrowers' subgroups might substantially increase any observed effects.

While recent research has provided more insight into the effects of microfinance, there is still plenty to be investigated. The robustness of the results of Duflo et al. should be further tested and extended to other micro-financial services. Additionally, the hypothesis of long-term effects as well as possible enhancements for any impact observed need to be further examined.

Microfinance does not seem to be the cure for poverty as it had once been lauded. However, there are some effects that are observable today and there may be others in the long run that remain undiscovered. Further identifying these effects and enhancing them with the right policy mix will not eradicate poverty on a macro scale, but it may have a positive impact on it in the future.

References

- BATEMAN, M. (2011): "Microfinance as a development and poverty reduction policy: is it every-thing its cracked up to be?," Overseas Development Institute Background Note.
- CHOWDHURY, A. (2009): "Microfinance as a Poverty Reduction Tool A Critical Assessment," *DESA Working Paper*, (89).
- DUFLO, E., A. BENERJEE, R. GLENNERSTER, AND C. KINNAN (2013): "The miracle of microfinance? Evidence from a randomized evaluation," *NBER Working Paper*, (18950).
- HERMES, N., AND R. LENSINK (2011): "Microfinance: Its Impact, Outreach, and Sustainability, in World Development," *World Development*, 39(6), 875–881.
- HULME, D., AND P. MOSLEY (1996): *Finance against Poverty*, vol. 1.
- IMAI, K., R. GAIHA, G. THAPA, AND S. ANNIM (2012): "Microfinance and Poverty: A Macro Perspective," *World Development*, 40(8), 1675–1689.

PELLEGRINA, L. (2011): "Microfinance and Investment: A Comparison with Bank and Informal Lending," *World Development*, 39(6), 882-897.

THE ECONOMIST (2009): "A partial marvel. Microcredit may not work wonders but it does help the entrepreneurial poor," Accessed: 21.11.2013.

——— (2013): "Not always with us. The world has an astonishing chance to take a billion people out of extreme poverty by 2030," <http://www.economist.com/node/21578643/print>, Accessed: 22.12.2013.

THE NOBEL PEACE PRIZE (2006): "The Nobel Peace Prize for 2006, Oslo 13.10.2006," <http://www.nobelprize.org/nobelprizes/peace/laureates/2006/press.html>, Accessed: 05.12.2013.

THE WORLD BANK (2013): "Choosing and Estimating a Poverty Line," <http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTPOVERTY>, 12.12.2013.

TODARO, M., AND S. SMITH (2011): *Economic Development*. 11 edn.

Appendix

Progress in Reporting, 1997-2010

Date	Number of Programs Reporting	Total Number of Clients Reached	Number of Poorest Clients Reported
12/31/97	618 institutions	13,478,797	7,600,000
12/31/98	925 institutions	20,938,899	12,221,918
12/31/99	1,065 institutions	23,555,689	13,779,872
12/31/00	1,567 institutions	30,681,107	19,327,451
12/31/01	2,186 institutions	54,932,235	26,878,332
12/31/02	2,572 institutions	67,606,080	41,594,778
12/31/03	2,931 institutions	80,868,343	54,785,433
12/31/04	3,164 institutions	92,270,289	66,614,871
12/31/05	3,133 institutions	113,261,390	81,949,036
12/31/06	3,316 institutions	133,030,913	92,922,574
12/31/07	3,552 institutions	154,825,825	106,584,679
12/31/09	3,589* institutions	190,135,080	128,220,051
12/31/10	3,652 institutions	205,314,502	137,547,441

* The small increase in the number of institutions reporting December 31, 2009 data is due, in part, to the fact that we subtracted from the list of "practitioners" more than 88 Networks who assist with the collection of practitioner action plans but have no clients themselves.

Figure 1: Global Increase of Microfinance. Source: Maes, J.P., L.R. Reed: State of the Microcredit Summit Campaign Report 2012

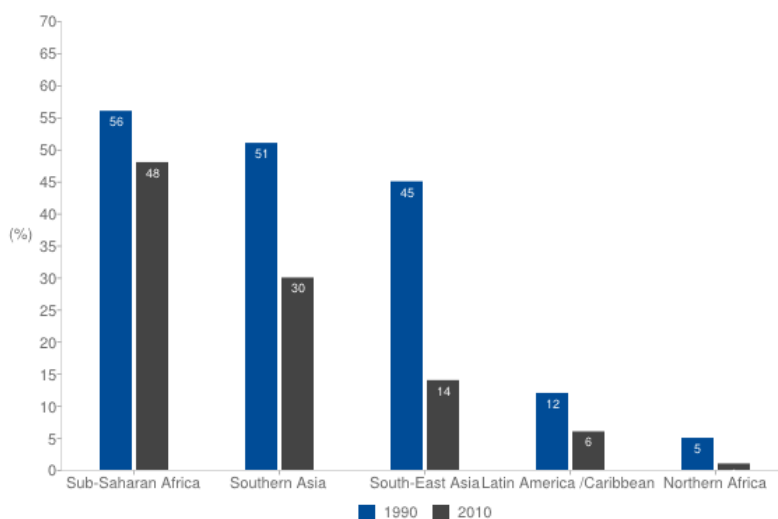


Figure 2: Regional Decrease in Poverty. Source: UNDP: Eradicate extreme poverty and hunger, in: http://www.undp.org/content/undp/en/home/mdgoverview/mdg_goals/mdg1/ (01.12.2013)

M&A Auctions and Toeholds

Anna Papies *

Introduction

Corporate takeovers, or more generally M&A activities, are a pervasive business practice with recently more than forty thousand transactions all over the world¹. With large sums of money involved, there is a lot at stake for a company planning to acquire (parts of) another firm, which is why the prevalent practice involves the consultation of experts, such as investment banks or financial advisers. The opportunities which such a deal provides range from matters of synergies and efficiency-enhancing corporate restructuring to strategic aspects, such as the potentially advantaged competitive position a transaction might result in.

The usual sales process for corporate stakes is an auction mechanism, which differs in one crucial aspect from a standard auction procedure: the process is not initiated by the auctioneer at prespecified terms and conditions, but evolves when a potential buyer indicates his interest in the target. Thus, the decision for a certain auction design and the adaptation to the specific situation must be made, more or less, ad hoc and therefore require a sound preparation in advance. As well-prepared as a target company has to be, as well-considered must be a potential bidder's approach to it.

* Anna Papies received her degree (M.Sc.) from the University of Bonn in September 2013. The present article refers to her master thesis under supervision of Prof. Dr. Audrey Xianhua Hu, which was submitted in September 2013.

¹See Appendix A.1

In this thesis², the focus is on a particular form of strategic pre-auction behavior, the acquisition of a so-called toehold, in two-bidder contests³. A toehold is a small stake in the target company, acquired before the start of the actual auction process, and even before the target is brought into play. As long as this stake does not exceed a five percent threshold, its acquisition need not be announced publicly, i.e. to the target's board and the Securities and Exchange Commission or, outside the USA, a comparable authority.

The strategic special feature of a toehold is the two sidedness of the incentives it provides. On the one hand, a toeholder has the incentives of a buyer who wants to buy the outstanding shares at a preferably low price. On the other hand, he has some of the incentives of a seller wanting to sell his toehold at a price as high as possible. Therefore, depending on whether he expects to win or lose the auction with a higher probability, the toeholder either prefers a high or a low selling price. The advantages a toehold entails are both of a strategic and a practical nature. Practically, a toehold is a way of purchasing shares at a low price before the market knows that the target is in play, which might, if known, have an effect on the price development. Strategically, as I will show in the course of this thesis, a toehold puts its owner in a highly favorable position. Both in case of winning *or* losing the auction, a toeholder makes some profit, either from the value received from the acquisition of the outstanding shares, or from the sale of his toehold at a potentially higher price than he had to pay. Consequently, he can submit a higher bid, or rather stay in the bidding for a little longer in an open-bid auction, than he could without a toehold. Due to his double-sided incentives and profits, in expectation, this is profitable. The more aggressive bidding behavior

²This thesis contains some technical derivations from Dr. Audrey Xianhua Hu. I gratefully acknowledge her valuable contribution.

³The assumption of (only) two bidders is empirically supported by Bradley, Desai, and Kim (1988), who find in their sample of 73 multi-bidder contests only 8 cases with more than two bidders.

not only increases the bidder's probability of winning, but also initiates a cycle of an aggravated and self-enforcing 'winner's curse' for the opponent.

The winner's curse usually occurs in common value settings, i.e. when all bidders attach the same value to the object for sale and each bidder has a private estimation of the uncertain value. Winning the auction is bad news then, implying that the own estimation is the highest, and thus one has, most probably, overestimated the value and paid more for the object than it is worth. In standard settings, rational agents react to the winner's curse by shading their bids, taking the bad news winning implies into account.

For a bidder without a toehold, competing against a toehold, the situation is strategically similar. Winning against an opponent who bids more aggressively is bad news for the above-mentioned reasons. Hence, a rational bidder should shade his bid, i.e. bid more conservatively.⁴ Knowing this, the toehold can bid even more aggressively and the cycle proceeds. Hence, a toehold's benefits are magnified by its interaction with the winner's curse, whose effects are, in turn, multiplied up by the toehold. We will see that with toeholds, even in a private value setting, where each bidder knows his own value but not those of the others, a kind of ex-post regret similar to the winner's curse is likely to occur.

Which of these settings, common or private values, is the more adequate one, cannot be answered in a general way, but rather depends on the takeover's specific circumstances. Generally, common values can better be aligned with hostile takeovers, where potential raiders have similar intentions and plans for the target company, once having replaced the (potentially badly performing) management. These can be, for instance, value-increasing measures that the incumbent management could have taken, but for whichever reasons has not done so. Private values, however, are more in line with friendly takeovers. Therein, different bid-

⁴Of course, this requires that the bidder knows about his opponent's toehold. In the forthcoming formal analysis this will, by assumption, be the case.

ders come with both different prerequisites and firm-specific conditions, but also differing skills in managing the target's assets. Thus, they accomplish different (levels of) synergistic gains. In practice, though, most takeovers are characterized by a combination of both private and common value elements.

A further issue in modeling a takeover is that of the most suitable auction format. As mentioned above, a takeover auction is usually initiated by the bid or the announced interest of a bidder, followed by counter bids that arrive sequentially. On these grounds, it seems natural to consider an open ascending-bid auction, like the English auction (EA). In fact, this is the prevalent choice in practice. An alternative would be a sealed-bid first- or second-price auction. However, it is hard to justify, or rather commit to, a sealed-bid auction when, after the result is revealed, a topping bid comes up. Under Delaware law, which is the predominant U.S. corporate law, the target management is obliged to accept the highest bid. So even if a management is not primarily concerned with the shareholders' interest, it is induced to find and accept the highest among the submitted bids, thus maximizing shareholder value.

The present thesis is organized as follows. First, I have a closer look at two seminal papers and present them in a somewhat more transparent way. The first section deals with a model (Burkart (1995)) analyzing the effects of toeholds in a private value setting with either one single or two symmetric toeholds, and shows the basic impact on optimal bidding behavior. The paper considered in the following (Bulow, Huang, and Klemperer (1999), henceforth BHK) is based on a common value setting and extends the analysis to a more general and complex environment with two toeholders that need not have identical initial stakes. Subsequently, I give two empirical examples and then proceed outlining some related literature and making suggestions for further research.

The Private Value Setting

Burkart (1995) analyzes an English auction for a takeover target with two bidders and shows how the introduction of toeholds alters the bidding strategy of a toeholder. In a simple setting with one toeholder competing against a non-toeholder, he demonstrates the strategic considerations and mechanics behind the bidding behavior and shows that overbidding, i.e. bidding more than one's value, is optimal. As a consequence, the toeholder might make a loss in case of winning and the auction outcome is no longer guaranteed to be efficient, like in the standard setting without toeholds.

The Model

There are two risk-neutral bidders with exogenously acquired toeholds $\theta_k \in [0, \frac{1}{2})$, $k = i, j$. Their private valuations $v_k \stackrel{i.i.d.}{\sim} [0, 1]$ are independently⁵ distributed according to $G_k(\cdot)$, where $G_k(0) = 0$ and $\frac{\partial G_k}{\partial v_k} > 0$ for all $v_k \in [0, 1]$. Furthermore, $G_k(\cdot)$ is twice continuously differentiable and satisfies the monotone hazard rate condition, i.e.

$$\frac{d}{dv_k} \frac{g(v_k)}{(1 - G(v_k))} \geq 0.$$

The valuations can be interpreted as the value per share of the target company⁶ once it has been acquired by the winning bidder. The value under the incumbent management, v_{im} , is normalized to be zero without loss of generality (w.l.o.g.). While the private values are known only to the respective bidder, both its distribution and v_{im} , as well as the size of the toeholds, are common knowledge. A bidder's strategy, i.e. his bid, is denoted by $b_k \sim F_k(\cdot)$ and $b_k \geq 0$.

⁵The independence assumption is equivalent to an interpretation of private valuations in terms of bidder-specific synergistic gains.

⁶The total number of shares is normalized to 1.

To focus on the mere auction process and rule out undesired strategic effects, it is assumed that negotiation, collusion and any kind of retrading after the actual auction are not possible. Moreover, bids cannot be withdrawn and, conditional on winning the auction, must hold for all outstanding shares (including those of the rival bidder). Throughout this thesis, I further abstract away from any stock price effects or reactions of the firm's market value.

Although takeover auctions are best modeled using an open ascending auction, here, a second-price auction (SPA) is analyzed. In a standard setting, these two formats are strategically equivalent when there are only two bidders⁷. Burkart (1995) shows that this result is not invalidated by the introduction of toeholds.

The One-Toehold Case

Assume that bidder i has a positive toehold, while bidder j does not own any shares prior to the auction, i.e. $\theta_i > 0$ and $\theta_j = 0$. For bidder j , the situation is equivalent to the standard SPA without toeholds, where a dominant strategy is to bid one's value. The following Lemma establishes this point (see Appendix A.2 for a proof).

Lemma 1. *For a bidder without a toehold who competes against a rival with a positive toehold, it is optimal to bid his value, i.e. $b_j^* = v_j$.*

For a toeholder, however, there is no dominant strategy. To see this, assume (w.l.o.g.) that in case of a tie, bidder j wins the auction. Given b_j , for bidder i it is optimal to bid b_j if $v_i < b_j$ so as to maximize the price he receives for selling his toehold. If, on the other hand, $v_i > b_j$, he wants to win the auction by submitting a bid $b_i^* > b_j$, which is, if bids are not restricted to increase by fixed amounts, not determined.

When bidder j bids his value, i.e. $b_j = v_j$, the cumulative distribution function

⁷Cf. Krishna (2009) for instance.

of his bid and his value are the same ($G_j(\cdot) = F_j(\cdot)$). Bidder i 's optimal bid is determined by solving the following maximization problem:

$$\max_{b_i} \theta_i b_i [1 - F_j(b_i)] + v_i F_j(b_i) - (1 - \theta_i) \int_0^{b_i} b_j f_j(b_j) db_j,$$

where the first term is the profit from the sale of the toehold when losing the auction, the second term is the value received and the third term the expected payment⁸ in case of winning.

The first-order condition implies

$$\theta_i(1 - G_j(b_i)) + v_i g_j(b_i) = b_i g_j(b_i) \Leftrightarrow b_i^* = v_i + \theta_i \frac{1 - G_j(b_i)}{g_j(b_i)} > v_i, \quad (1)$$

where $G_j(b_i)$ is substituted for $F_j(b_i)$, and accordingly, $g_j(b_i)$ for $f_j(b_i)$. Due to the monotone hazard rate condition, the problem is concave, so the second-order condition holds and b_i^* maximizes bidder i 's expected profit. From the above expression (1), one can see that overbidding occurs for all values $v_i \in [0, 1]$ ⁹. If v_i takes on its maximum value of one, however, bidder i does *not* overbid, but bids his value v_i . This follows from the fact that bidder j bids his value in equilibrium, so $b_j = v_j$, and, on the other hand, from the insight that bidder i never bids *less* than his value. Thus, $G_j(b_i) = 1$ and $b_i^* = v_i$. This is, however, just a special case and occurs with zero probability, so it is not critical in any way.

The (Symmetric) Two-Toehold Case

To derive an explicit solution for the above equilibrium, I assume that the bidders' values are independently and uniformly distributed on $[0, 1]$. Thus, we have

⁸Strictly speaking, this is the expectation of the payment conditional on winning, multiplied by the fraction of shares that the bidder does not yet own.

⁹Burkart shows that overbidding occurs as a general feature of bidding with toeholds, i.e. that it is individually rational, irrespective of the number of opponents and their potential toeholds and bidding behavior.

$G_k(b_k) = b_k$, and $g_k(\cdot) = 1$. From (1) we then have

$$b_i = v_i + \theta_i(1 - b_i) \Leftrightarrow b_i = \frac{v_i + \theta_i}{1 + \theta_i}.$$

With the simplification to uniformly distributed values, we can extend the analysis to the (symmetric) two-toehold case, where $\theta_i = \theta_j = \theta > 0$. From now on, the respective expressions for bidder j can be computed in an analogous way to those of bidder i (if not stated otherwise) and therefore will not be restated. For the symmetric case, the first-order condition for bidder i implies

$$\begin{aligned} & \theta(1 - F_j(b_i) - f_j(b_i)b_i) + f_j(b_i)v_i - (1 - \theta)b_i f_j(b_i) \stackrel{!}{=} 0 \\ \Leftrightarrow & \quad \theta(1 - F_j(b_i)) + f_j(b_i)v_i = f_j(b_i)b_i \\ \Leftrightarrow & \quad b_i = v_i + \theta \frac{1 - F_j(b_i)}{f_j(b_i)}. \end{aligned}$$

We are looking for a symmetric equilibrium with strategies that monotonically increase in the valuations, i.e. $b'_i(v_i) > 0$. Thus, b_i is invertible and we can rewrite $b_i = h_i(v_i)$ as $h_i^{-1}(b_i) = v_i$. This yields

$$F_j[b_i] = F_j[h_i(v_i)] = G_j[h_i^{-1}(b_i)] = v_i = G_i(v_i) = F_i[b_i],$$

where the second equality follows from monotonicity. Note that from

$$\frac{\partial F_j(b_i)}{\partial v_i} = f_j(b_i) \cdot b'_i(v_i) = 1,$$

we have $f_j(b_j) = \frac{1}{b'_i(v_i)}$. Plugging in these terms, we can rewrite the first-order condition as

$$b'_i(v_i)\theta(1 - v_i) - b_i = -v_i.$$

Multiplying both sides by $(1 - v_i)^{\frac{1-\theta}{\theta}}$ yields

$$\begin{aligned} b'_i(v_i)\theta(1 - v_i)^{\frac{1}{\theta}} - b_i(1 - v_i)^{\frac{1-\theta}{\theta}} &= -v_i(1 - v_i)^{\frac{1-\theta}{\theta}} \\ \Leftrightarrow \frac{\partial}{\partial v_i} \left[b_i(v_i)\theta(1 - v_i)^{\frac{1}{\theta}} \right] &= -v_i(1 - v_i)^{\frac{1-\theta}{\theta}}. \end{aligned}$$

Integrating over v_i on $[v_i, 1]$, we have

$$\begin{aligned} -b_i(v_i)\theta(1 - v_i)^{\frac{1}{\theta}} &= -\int_{v_i}^1 v_i(1 - v_i)^{\frac{1-\theta}{\theta}} dv_i \\ \Leftrightarrow -b_i(v_i)\theta(1 - v_i)^{\frac{1}{\theta}} &= -\left[-v_i\theta(1 - v_i)^{\frac{1}{\theta}} \right]_{v_i}^1 + \theta \int_{v_i}^1 (1 - v_i)^{\frac{1}{\theta}} dv_i \\ \Leftrightarrow b_i(v_i)\theta(1 - v_i)^{\frac{1}{\theta}} &= \theta v_i(1 - v_i)^{\frac{1}{\theta}} - \frac{\theta^2}{1 + \theta} \left[(1 - v_i)^{\frac{1+\theta}{\theta}} \right]_{v_i}^1 \\ \Leftrightarrow b_i(v_i) &= v_i + \frac{\theta}{1 + \theta}(1 - v_i) \Leftrightarrow b_i(v_i) = \frac{v_i + \theta}{1 + \theta}. \end{aligned} \quad (2)$$

Note that in this case, unlike in the one-toehold case, the outcome is efficient, as the bidder with the higher valuation always wins. Interestingly, the above expression is the same as the toehold holder's optimal bid in the one-toehold case with uniformly distributed values.¹⁰ This result is due to the following two effects:

- (i) Given b_j , bidder i 's overbidding increases his probability of winning from

$$\begin{aligned} pr\{b_j < v_i\} &= 1 - G_i(b_j) = 1 - b_j \quad \text{to} \\ pr\{b_j < \frac{v_i + \theta}{1 + \theta}\} &= pr\{(1 + \theta)b_j - \theta < v_i\} \\ &= 1 - G_i((1 + \theta)b_j - \theta) = (1 + \theta)(1 - b_j). \end{aligned}$$

- (ii) Given that bidder i overbids by Δ , the risk that he pays more than his

¹⁰Thus, the comparative statics results from the preceding section carry over to this special case.

value increases from

$$\begin{aligned}
 pr\{v_i < v_j < v_i + \Delta\} &= pr\{v_j < v_i + \Delta\} - pr\{v_j < v_i\} = \Delta \quad \text{to} \\
 pr\{v_i < \frac{v_j + \theta}{1 + \theta} < v_i + \Delta\} &= pr\{v_j < (1 + \theta)(v_i + \Delta) - \theta\} - pr\{v_j < (1 + \theta)v_i - \theta\} \\
 &= (1 + \theta)(v_i + \Delta) - \theta - (1 + \theta)v_i + \theta = (1 + \theta)\Delta
 \end{aligned}$$

when bidder j overbids according to (2).

We see that bidder i 's overbidding increases his probability of winning, but at the same time his opponent's overbidding increases his (bidder i 's) risk of paying more for a share than it is worth to him in case of winning. Since these probabilities are increased by the same factor, $1 + \theta$, the two effects cancel each other out and there is no overall effect on the optimal bidding strategy induced by the introduction of a second (symmetric) toehold. Note that this result critically depends on the assumption of uniformly distributed values. Thereby, a linear relation between the valuation and the optimal bid is established, which is the reason why the two opposing effects exactly countervail each other.

A general result for two positive toeholds in a private value setting has, to the best of my knowledge, not yet been derived so far. In the following section, we will see that within the framework of a common value setting, BHK (1999) succeed in establishing a general solution for two positive, but not necessarily symmetric toeholds.

The Common Value Setting

The Model

In the model analyzed by BHK, just like in Burkart (1995), there are two risk-neutral bidders $i, j \in k$. The difference is, though, that both acquire a toehold,

$\theta_k \in (0, \frac{1}{2})$, prior to bidding for the target company that is sold via a standard ascending-bid auction. Bidders attach a purely common value $v(t_i, t_j)$ to it which is symmetric and strictly increasing in the private signals that they receive before the auction begins. These signals are denoted by $t_k \sim U[0, 1]$. By $b_k(t_k)$, which is assumed to strictly increase in t_k , we denote the price at which a bidder quits if the other one has not yet done so. Bidder i 's "marginal revenue" is defined as $MR_i(t_i, t_j) \equiv v(t_i, t_j) - (1 - t_i) \frac{\partial v}{\partial t_i} - (t_i, t_j)$ and we assume that a higher signal implies a higher marginal revenue: $t_i > t_j \Rightarrow MR_i(t_i, t_j) > MR_j(t_i, t_j)$. Bidder k 's equilibrium expected profit (conditional on his signal) is $\pi_k(t_k)$ and Π_k denotes his unconditional expected profit (averaged across all possible signals). Finally, Π_0 is the expected profit accruing to all shareholders except the two bidders.

Unlike Burkart's model, both bidders face the same problem here. Since neither bidder's strategy can be expressed by simple standard auction considerations, the analysis is by far more complex. Nevertheless, BHK show that there exists a unique symmetric Nash equilibrium.

To find the optimal bidding strategies, we think of a bidder choosing a drop-out price at which he finds it optimal to quit the auction and sell his toehold to his opponent at this price. If for any given signal vector (t_i, t_j) a bidder, say bidder j , wins the auction at price $b_i(t_i)$ at which bidder i quits, bidder j 's profit is

$$v(t_i, t_j) - (1 - \theta_j)b_i(t_i),$$

i.e. the (realized) common value minus the selling price, multiplied by the fraction of shares that he does not yet own. If, on the other hand, bidder j drops out at a price p , he sells his stake in the company to bidder i , which yields a profit of

$\theta_j p$. Let the drop-out price p be a function of bidder i 's signal s , i.e. $p \equiv b_i(s)$. Since b_i is strictly increasing in its argument, one can say that bidder j drops out at a signal s of bidder i given his bid function $b_i(t_i)$. Assume that the bid functions exist and are differentiable. In a Nash equilibrium, they are such that they maximize a bidder's expected payoff, given his opponent's strategy. Thus, bidder j chooses a drop-out signal s (of bidder i) to maximize his expected payoff at the beginning of the auction, taking $b_i(t_i)$ as given:

$$\pi_j(s, t_j) = \int_0^s [v(t_i, t_j) - (1 - \theta_j)b_i(t_i)]dt_i + (1 - s)\theta_j b_i(s).$$

The first term relates to the event of winning, i.e. t_i is below the drop-out level s , whereas the last term represents the case in which $t_i > s$, such that bidder j loses the auction. The first-order condition implies

$$\begin{aligned} \frac{\partial}{\partial s} \pi_j(s, t_j) &= v(s, t_j) - (1 - \theta_j)b_i(s) - \theta_j b_i(s) + (1 - s)\theta_j b'_i(s) \\ &= v(s, t_j) - b_i(s) + (1 - s)\theta_j b'_i(s) \stackrel{!}{=} 0, \forall t_j. \end{aligned}$$

This defines the optimal s as a function of t_j , though, it is more useful to view it as an implicit definition of t_j as a function of s . For this purpose, I define

$$\phi_i(\cdot) \equiv b_i^{-1}(b_j(\cdot)).$$

This is well defined with $\phi'_j > 0$, because we assume that the bid function be increasing.

To make this expression more intelligible, one can think of $b_i^{-1}(b_j)$ as a signal received by bidder i , who bids as if he had received t_j and mimicked bidder j 's strategy (b_j): $b_i(b_i^{-1}(b_j(t_j))) = b_j(t_j)$. Alternatively, $b_i(\phi_i(t_j)) = b_j(t_j)$ specifies

two types¹¹ of bidders, $\phi_i(t_j)$ and t_j , who quit at the same price. In equilibrium, we have $t_j = \phi_j(t_i)$ and, equivalently, $t_i = \phi_i(t_j)$ which implies $b_i(t_i) = b_j(t_j)$ and $v(\phi_i(t_j), t_j) = v(t_i, \phi_j(t_i))$. Rearranging terms, the necessary condition for an equilibrium is thus:

$$-b_i(s) + (1-s)\theta_j b'_i(s) = -v(s, \phi_j(s)). \quad (3)$$

To derive from this the explicit expression for the optimal bid function, some technical re-arrangements are helpful (see Appendix A.3). Substituting by the so-derived result, we get the following equilibrium bidding strategy for bidder i :

$$b_i(t_i) = \frac{\int_{t_i}^1 v(s, 1 - (1-s)^{\frac{\theta_i}{\theta_j}})(1-s)^{\frac{1}{\theta_j}-1} ds}{\theta_j(1-t_i)^{\frac{1}{\theta_j}}}.$$

Since $\phi_j(t_i)$ is explicitly determined by (A.3) (cf. A.3), this equilibrium is unique. In the (standard) model without toeholds, this is not possible. Therein, every different weakly increasing function $\phi_j(t_i)$ yields a different equilibrium $b_i(t_i) = v(t_i, \phi_j(t_i)) = b_j(\phi_j(t_i))$. In such equilibria, each bidder stays in the bidding until the value has reached the critical level that is determined by assuming that the opponent receives the same signal as oneself (this is the standard case of an English auction with two bidders).

Profit Analysis

One of the most interesting matters from the perspective of an auction designer is that of the expected revenue, i.e. the expected sale price in the context of

¹¹I use the term “type” here in an analogous way to “signal”, as it makes the explanation more intuitive.

corporate takeovers. In our case, this can be written as:

$$\frac{\Pi_0}{(1 - \theta_i - \theta_j)}.$$

To see why, note that the expected sale price, multiplied by the shares held by the non-bidding shareholders, must equal their expected surplus:

$$E[P] \cdot (1 - \theta_i - \theta_j) = \Pi_0,$$

where $E[P]$ denotes the expected sale price. Notice that Π_0 is the residual surplus that remains after the bidding shareholders received their (expected) stakes of the total surplus, so it can be expressed as follows:

$$\Pi_0 = \int_{t_i=0}^1 \int_{t_j=0}^1 v(t_i, t_j) dt_j dt_i - (\Pi_i + \Pi_j),$$

where $\int_{t_i=0}^1 \int_{t_j=0}^1 v(t_i, t_j) dt_j dt_i$ is the expected (common) value of the firm, i.e. the total surplus that is at stake.

Comparative Statics

The main message of the BHK paper is that a toehold helps a bidder win the auction. To see this, I analyze a bidder's probability of winning and consider its reaction to changes in the size of the toeholds. Bidder i wins against bidder j if $t_j \leq \phi_j(t_i) (\hat{=} b_j(t_j) \leq b_i(t_i))$,¹² so his probability of winning *conditional on his signal* is $\phi_j(t_i)$. His *unconditional* probability of winning can be calculated by

¹²Note that by this formulation, we assume (w.l.o.g.) that bidder i is the bidder with the higher signal, since in (a symmetric) equilibrium we have $b_i(\cdot) = b_j(\cdot)$.

averaging across all possible values of t_i :

$$\begin{aligned}\int_0^1 \phi_j(t_i) dt_i &= \int_0^1 1 - (1 - t_i)^{\frac{\theta_j}{\theta_i}} dt_i = \left[t_i + \frac{\theta_j}{\theta_i + \theta_j} (1 - t_i)^{\frac{\theta_i + \theta_j}{\theta_j}} \right]_0^1 \\ &= 1 - \frac{\theta_j}{\theta_i + \theta_j} = \frac{\theta_i}{\theta_i + \theta_j} \equiv p_i(EA).\end{aligned}$$

Bidder i 's probability of winning increases with his own and decreases with bidder j 's toehold:

$$\begin{aligned}\frac{\partial[\theta_i/(\theta_i + \theta_j)]}{\partial\theta_i} &= \frac{\theta_i + \theta_j - \theta_i}{(\theta_i + \theta_j)^2} = \frac{\theta_j}{(\theta_i + \theta_j)^2} > 0, \\ \frac{\partial[\theta_i/(\theta_i + \theta_j)]}{\partial\theta_j} &= \frac{-\theta_i}{(\theta_i + \theta_j)^2} < 0.\end{aligned}$$

In the limit, when θ_i goes to zero, bidder i 's probability of winning approaches zero, given that θ_j is positive. This is because bidder j always has an incentive to bid above his value, while bidder i does not. Just as in Burkart's model, bidder i without a toehold optimally bids his value when competing against a toeholder. So he has no reason, and thus no incentive, to overbid.

The question how bidder j 's bid reacts to an increase in bidder i 's toehold cannot be answered in a general way. The result critically depends on bidder j 's signal, i.e. his expectations about the common value.

For computational ease, for the rest of this section I consider the linear example from BHK (1999) with $v(t_i, t_j) = t_i + t_j$ instead of the general model. With this modification, the optimal bid function is

$$b_j(t_j) = \frac{\int_{t_j}^1 [s + 1 - (1 - s)^{\frac{\theta_j}{\theta_i}}] (1 - s)^{\frac{(1 - \theta_i)}{\theta_i}} ds}{\theta_i (1 - t_j)^{\frac{1}{\theta_i}}}.$$

Recollecting terms, we have¹³

$$b_j(t_j) = t_j + \frac{\theta_i}{1 + \theta_i}(1 - t_j) + 1 - \frac{1}{1 + \theta_j}(1 - t_j)^{\frac{\theta_j}{\theta_i}}. \quad (4)$$

Now we can take the derivative with respect to bidder i 's toehold:

$$\begin{aligned} \frac{\partial b_j(t_j)}{\partial \theta_i} &= (1 - t_j) \left[\frac{1 - \theta_i - \theta_i}{(1 + \theta_i)^2} \right] - \frac{1}{1 + \theta_j} (1 - t_j)^{\frac{\theta_j}{\theta_i}} \ln(1 - t_j) \left(-\frac{\theta_j}{\theta_i^2} \right) \\ &= \frac{1 - t_j}{(1 - \theta_i)^2} \left[1 + \frac{1}{1 + \theta_j} \cdot \frac{\theta_j}{\theta_i^2} (1 - t_j)^{\frac{\theta_j - \theta_i}{\theta_i}} (1 + \theta_i)^2 \ln(1 - t_j) \right]. \end{aligned}$$

This expression is positive if bidder j 's signal is below some critical level and negative otherwise. The intuition is that a type of bidder j with a low signal, thus having pessimistic expectations about the target firm's value, regards bidder i 's increasing toehold as an "opportunity to lose" the auction with a higher probability. Consequently, he bids more aggressively to bid up the price he receives for selling his toehold. If bidder j 's signal is high, however, a larger θ_i aggravates the winner's curse that he might suffer from in case of winning. Thus, he bids more conservatively.¹⁴

From the perspective of an auction designer, or the target management in case of a takeover, the most interesting question still has to be answered: How do toeholds affect the expected selling price? On an intuitive basis, this can be answered by standard considerations: the more fierce the competition among the buyers (bidders), the higher the price they will have to pay. In our context, this translates to: the more symmetric bidders are ex ante, i.e. the more equal their toeholds are, the more fair or homogenous, and thus harsh, is the competition

¹³See Appendix A.4 for a detailed derivation.

¹⁴Note that this result occurs in the linear case studied here, but not in general. Which of the two conflicting effects, the increased "opportunity to lose" or the aggravated winner's curse, prevails, depends on the form of the value function.

among them. If toeholds are more asymmetric, however, the bidder with the larger stake has an advantage that is known to both of them before the start of the auction. Consequently, his probability of winning, and thus his bid, is higher, and his opponent, knowing that he will suffer from an aggravated winner's curse in case of winning, will bid more conservatively.

In some sense, this result reminds of a well-known finding in the theory of optimal auctions, stating that the "Expected revenue from an absolute English auction with $N+1$ bidders exceeds expected revenue from an English auction with N bidders followed by a take-it-or-leave-it offer to the last remaining bidder if either (i) bidders' values are private; or (ii) bidders' signals are affiliated¹⁵" (Bulow and Klemperer, 1996, p.187)). Since with sufficiently asymmetric toeholds there is no 'serious' competitor from the perspective of the advantaged bidder, making toeholds more symmetric has a somewhat similar effect to adding an additional ('serious') bidder: it levels the playing field and so enhances both competition and the expected selling price. The following numerical example demonstrates the results.

Example 1. Assume $v(t_i, t_j) = t_i + t_j$, like in BHK's linear example. Consider two scenarios where the sum of toeholds is the same, but their distribution is either (a) symmetric or (b) asymmetric, respectively. Further assume that a bidder receives a signal $x = E[t_k] = 0.5$.

¹⁵Since in our setting signals are independent, they are, by definition of affiliation, also affiliated, so the result applies.

(a) Symmetric toeholds: $\theta_i = \theta_j = 0.02$

$$\begin{aligned}\Pi_i &= \Pi_j \approx 0.17, \\ b_i(x) &= b_j(x) = 2 - \frac{100}{51}(1-x) \approx 1.02, \\ \Pi_0 &\approx 0.67.\end{aligned}$$

(b) Asymmetric toeholds: $\theta_i = 0.03, \theta_j = 0.01$

$$\begin{aligned}\Pi_i &\approx 0.3, \Pi_j \approx 0.07, \\ b_i(x) &= 2 - \frac{100}{101}(1-x) - \frac{100}{103}(1-x)^3 \approx 1.38, \\ b_j(x) &= 2 - \frac{100}{103}(1-x) - \frac{100}{101}(1-x)^{\frac{1}{3}} \approx 0.73, \\ \Pi_0 &\approx 0.63.\end{aligned}$$

As predicted, Π_0 and thus the expected sale price $\frac{\Pi_0}{1-\theta_i-\theta_j}$, is higher with symmetric toeholds. Furthermore, the high-toehold bidder bids more aggressively in the asymmetric setting and receives a higher expected profit, whereas the low-toehold bidder bids more conservatively and has a lower expected profit, which is in line with the results derived above.

Empirical Examples

Due to the high frequency of M&A activities, there is a multitude of empirical examples. However, toeholds are not that easily observed and rather scarcely documented, especially if they fall below the five percent threshold. Nonetheless, I found two examples that are particularly illustrative: the attempted takeover of the Manchester United Football Club (MUFC) by the British Sky Broadcasting Group PLC (BSkyB) and, on the other hand, the Glaxo-Wellcome merger.

British Sky Broadcasting Group PLC/Manchester United

In 1999, an attempted takeover of MUFC by BSkyB was blocked by the Department of Trade & Industry, because an investigation by the Monopolies and Mergers Commission (MMC) had pointed out its anticompetitive effects in terms of the sale of Premier League broadcasting rights. The point in these sales is that football matches between Premier League clubs are the most lucrative branch for pay television companies. Since the (exclusive) rights to broadcast these matches are sold as a package including *all* matches that are telecasted, they are particularly valuable and constitute a competitive advantage. The rights are sold, at regular intervals, via an auction, and since 1992, no competitor ever managed to win against BSkyB until then. Harbord and Binmore (2000) report that the 1992 contract included a ‘meet the competition’ clause, allowing BSkyB to match any other bid in the subsequent auction, which put them in a highly advantaged position. In exactly the same way as in the model of toehold bidding, BSkyB’s competitors suffer from an aggravated winner’s curse: Their chances of winning are rather low, provided that any reasonable bid would be matched by BSkyB, but if they do actually win, they must attach a very high value to the object that BSkyB is not willing to pay, so most probably they overpay.

This clause, however, would no longer be in place in the 2001 auction. A way to abide its competitive advantage for BSkyB is the acquisition of MUFC, England’s most successful team. MUFC, like all other Premier League clubs, receives a share of the revenues created by the sale of the broadcasting rights. When owning MUFC, this revenue would accrue to BSkyB, thus it is similar to a toehold in the auction for broadcasting rights: In case of losing the auction, BSkyB would still participate in the gains, which allows it to bid more aggressively than any of its competitors. In case of winning, on the other hand, it would benefit twofold. According to Harbord and Binmore (2000), after the MMC’s decision, broadcast-

ers recognized the potential of toeholds. Consequently, many of them now own stakes in several football clubs, which has, all in all, a similar effect to owning a whole club. This is unlikely to attract the attention of the MMC or similar authorities and allowed (by Football Association rules), as long as a broadcaster does not own more than ten percent of more than one club.

Glaxo-Wellcome

Another empirical example of how a seemingly small advantage can drastically change the auction outcome is that of the Glaxo-Wellcome merger in 1995. Although there were no toeholds in terms of initial stakes in the target, a strategically similar effect occurred as a consequence of Glaxo's advantaged position. According to Klemperer (1998), many drugs companies attached a similar common value to the target (Wellcome), but "[...]there were also particular synergies that made Wellcome worth a little more to Glaxo than to any other potential bidder." ((Klemperer, 1998, p.9)). Provided with this advantage, Glaxo was prepared and willing¹⁶ to top any other bid. Consequently, Glaxo won with its bid of £ 9 billion and no other bidder even entered the bidding. This outcome was highly unsatisfactory, as there were two potential bidders offering £ 10 billion and £ 11 billion, respectively. The reason for this is that the Wellcome Trust, holding 39.5 percent of Wellcome, did not agree to accept a bid and guarantee its success, i.e. commit to not accepting Glaxo's topping bid. Note that the entry-detering effect of Glaxo's behavior was based on a threat that, hypothetically, could have been pretended. This shows that sometimes, it might be sufficient to signal readiness or to build up a reputation for aggressive bidding behavior to effectively improve one's position in a takeover contest.

¹⁶Cf. (Financial Times, 1995, p.32)

Related Literature and Further Research

The literature in the field of strategic bidding with toeholds and their specific roles is not that numerous. To give a short overview and embed the papers discussed above, I present the essential contributions in the following subsection. After this, I make some suggestions for further research that might be interesting and add to the existing literature.

Related Literature

Singh (1998) analyzes a single-toeholder model that is quite similar to Burkart (1995)'s, but differs in one crucial aspect, thus providing a further result: The winning bidder can renege on his bid. Without this possibility, Singh derives the same results: A toeholder bids more aggressively and possibly overpays in equilibrium, whereas a non-toeholder's strategy stays unaffected by the introduction of a toehold. This changes when reneging is no longer prohibited. Singh shows that the toeholder is induced to bid even more aggressively, while his opponent shades his bid. Observe that now, there is a dominant strategy for the toeholder, which is to match his opponent's bid. Hence, assuming (w.l.o.g.) that in case of a tie he is the winner, he always wins the auction and can, should his value be below the final bid, withdraw his offer. With the possibility of withdrawal, although he overbids, the toeholder does *not* overpay, which is inseparable without this possibility. Knowing that he can never win, the non-toeholder finds it optimal to shade his bid, since bidding his value now is associated with making zero profit. Bidding less than the value, however, gives him a chance of making some profit when his opponent's value is sufficiently low, so that he withdraws his bid ex post. In private value settings like the present one, bids are no strategic substitutes, i.e. the non-toeholder's strategy is not affected by his rival's more

aggressive bidding¹⁷. Thus, a toehold benefits the target's shareholders since the more aggressive bidding, in expectation, results in a higher selling price. With retractable bids, this is no longer the case, as the non-toeholder's shaded bid presents an upper bound of the price. Singh shows that the above-mentioned benefits increase in the size of the toehold, thus there are two measures that the target management can take to attain a high selling price. On the one hand, it can increase the toehold by selling or giving away further shares, thus fostering competition. On the other hand, it should ensure credible bidding by the prohibition of the ex-post withdrawal of winning bids, if necessary by means of break-up fees or other withdrawal penalties.

The empirical evidence regarding toehold bidding and the predictions of models in this context is highly heterogeneous and therefore, oftentimes, not perfectly convincing. Although this issue is, deliberately, mostly excluded from the present thesis, I want to mention the following paper, as it analyzes a model that helps explain and consolidate some puzzling empirical observations.

Betton, Eckbo, and Thorburn (2009) develop a threshold model in order to explain the "toehold puzzle", which comprises the observation that bidders scarcely acquire toeholds, but, at the same time, pay large takeover premiums. In addition, toeholds seem to follow a bimodal distribution and, on average, take on either a value of zero or a rather high one¹⁸. The authors assume that toeholds displease the target management, reducing their private benefits of control, so it might react by rejecting merger negotiations. In this case, bidders incur rejection costs. As a result, there is a threshold where the acquisition of a toehold, which entails transaction costs, becomes profitable in the first place. This threshold is the critical percentage of target shares that a bidder has to own so that the

¹⁷This holds for the standard setting but *not* for the special case of retractable bids.

¹⁸Although a purchase of five percent of the target shares has to be disclosed publically, this does not have to take place immediately. Consequently, a toeholder can acquire more shares unnoticedly and thereby build up a larger toehold.

benefits of this toehold equal the rejection costs. It is estimated to be about nine percent in the sample of more than 10,000 initial bids in U.S. mergers and tender offers. Empirically, however, with twenty percent, the threshold is slightly larger in the examined sample, and positive toeholds can be observed in only thirteen percent of the cases. For hostile takeovers, separately analyzed, this percentage is somewhat larger (fifty percent). This evidence can help explain the observed decline in toehold acquisition since the 1980s, which is, thus, attributable to the decreased occurrence of hostile takeovers.

Further Research

There are two directions for further research that I found particularly interesting while working on this thesis. The first one is that of the value setting. Most models assume either private or purely common values, presumably for simplicity. At the same time, there is a broad agreement that in practice, most takeovers entail elements of both settings. Klemperer (1998) introduces the notion of “almost common values” which allows for both common value and a small private value component in a bidder’s value function. The following expression is proposed as an example:

$$v_1 = (1 + \alpha_1)t_1 + t_2,$$

$$v_2 = t_1 + (1 + \alpha_2)t_2,$$

where α_i , $i \in \{1, 2\}$ denotes a bidder-specific (small) private component.

Klemperer (1998) shows that with almost common values, the outcomes of standard auction settings are strongly influenced and, for example, a standard ascending auction might be a suboptimal choice of auction format. However, the focus is more on (small) asymmetries among bidders and not so much on the

phenomenon of a mix of common and private values as such. Klemperer (1998) uses the change in the bidders' value function as a way to model slight asymmetries that can have various forms and toeholds, for instance, can be one of them. Thus, there is no prediction about how the bidding in an auction with toeholds changes in reaction to a change in the value setting. Goeree and Offerman (2003) address exactly this aspect and present a further way to combine both common and private values in a single model. They introduce multi-dimensional signals that can be reduced to a one-dimensional surplus signal in the following way. The common value component, V , is the arithmetic mean of all bidders' unbiased estimates, v_i , $i \in \{1, \dots, n\}$, of the true value:

$$V = \frac{1}{n} \sum_{i=1}^n v_i.$$

The private value component, c_i , is interpreted as a cost signal, but can be considered as a synergistic advantage in the context of takeover auction just as well. A bidder's total surplus, when he wins the auction and pays b , is $V - c_i - b$. To obtain a one-dimensional signal, the following reduction is performed:

$$s_i = \frac{v_i}{n} - c_i,$$

where s_i denotes some kind of type or a bidder-specific surplus.

The authors show that the more emphasis is put on the common value component, the lower are both the total expected surplus and the seller's revenue. However, the main focus of this paper, apart from the strategic analysis of the effect of different value settings, is on efficiency issues. The central result is that more uncertainty concerning the common value has a negative effect on efficiency. An interesting issue for further research would be a rigorous analysis of how the strategic effects of toeholds are affected when values are a combination of private

and common components, especially when these components' weights are varied. A second possible direction for further research is that of asymmetric value functions. Most of the literature uses the assumption of symmetric value functions in common value settings, i.e. the bidders' private information is assumed to influence the value in the same way. Of course, this is a simplification, as in reality, it is not clear why all agents should have equal skills in assessing a target's worth. BHK dedicate a short paragraph to this issue and state that if the low-information bidder, i.e. the bidder to whose signal the value function is less sensitive, has a larger toehold than his opponent, two of the formerly derived results are no longer valid.

A model that analyzes in a formal way the impact of asymmetries in value functions on strategic bidding behavior in the context of toeholds would be an interesting task for further research.

By now, the results of different publications in the field of toeholds lack comparability. Though, a generalization is not an easy task, as BHK (1995, p.3) indicate: "Obtaining explicit solutions of the equilibrium of an ascending-bid auction under private information and mixed valuation, however, turns out to be difficult." Nonetheless, an advancement in this field would be of both academic and practical use and, at best, facilitate valuable new findings.

Conclusion

Toeholds are an effective and influential strategic instrument. They increase a bidder's probability of winning, his optimal bid and expected profit, basically because they induce him to bid more aggressively. The two-sided incentives provided by a toehold confront the bidder with a constant trade-off between a high and a low selling price, and force him to soundly balance the advantages of winning and losing the auction.

Furthermore, the auction outcome may be inefficient as the valuation for the target is not the crucial element in determining the winner. In the one-toeholder case studied above, the expected revenue is the higher, the larger the size of the toehold. As the non-toeholder's strategy is, as long as the possibility of withdrawing the winning bid is ruled out, not affected by the toeholder's overbidding, bidding more aggressively automatically implies a higher expected selling price and revenue.

With a common value, however, an aggravated and self-enforcing form of the winner's curse may occur and make non-toeholders or weak bidders bid more conservatively. Efficiency is not an issue in this case as with common values any allocation is, trivially, efficient. The expected revenue is higher with symmetric than with asymmetric toeholds, so a larger toehold decreases the selling price, unless the other bidder's toehold increases by the same amount. Furthermore, a larger toehold decreases the opponent's probability of winning and, at least in the linear example, increases his optimal bid and expected profit if he receives a low signal. With a high signal, the opposite is true.

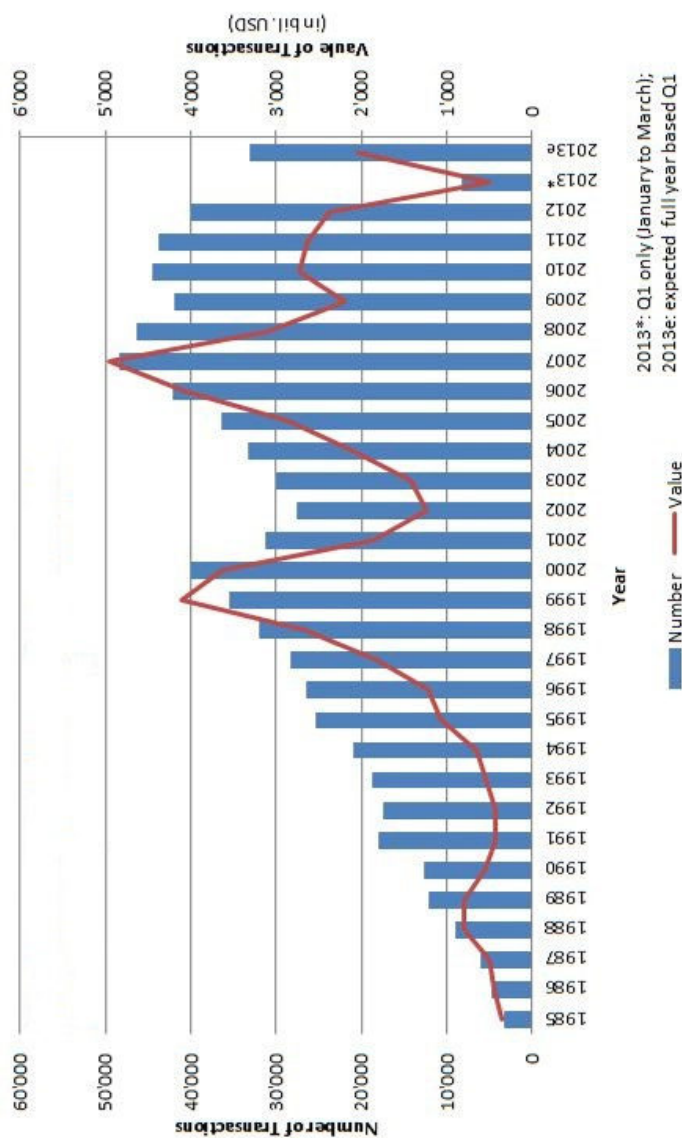
The literature in the field of toeholds is, for the most part, technically rather advanced and not very numerous. Hopefully, this thesis can help in the former matter and make the seminal papers discussed above somewhat more transparent and accessible. A useful advancement in the second issue could be in terms of analyzing the impact of the degrees of common/private value elements and asymmetries in the bidders' impact on value functions. Finally, it would be interesting to find out whether the above-mentioned results also hold in more general settings, for instance, in a private value setting with two or more toeholders.

References

- BETTON, S., B. E. ECKBO, AND K. S. THORBURN (2009): "Merger negotiations and the toehold puzzle," *Journal of Financial Economics*, 91(2), 158–178.
- BRADLEY, M., A. DESAI, AND E. H. KIM (1988): "Synergistic gains from corporate acquisitions and their division between the stockholders of target and acquiring firms," *Journal of Financial Economics*, 21(1), 3–40.
- BULOW, J., M. HUANG, AND P. KLEMPERER (1999): "Toeholds and takeovers," *Journal of Political Economy*, 107(3), 427–454.
- BULOW, J., AND P. KLEMPERER (1996): "Auctions Versus Negotiations," *The American Economic Review*, 86(1), 180–194.
- BULOW, J., P. KLEMPERER, AND M. HUANG (1995): "Toeholds and Takeovers: General Characterization, Existence, and Uniqueness of Equilibrium,," Manuscript; Stanford University, Nuffield College, Oxford University.
- BURKART, M. (1995): "Initial shareholdings and overbidding in takeover contests," *The Journal of Finance*, 50(5), 1491–1515.
- FINANCIAL TIMES (1995): "Wellcome still smarting over handling of Trust's stake," .
- GOEREE, J. K., AND T. OFFERMAN (2003): "Competitive Bidding in Auctions with Private and Common Values*," *The Economic Journal*, 113(489), 598–613.
- HARBORD, D., AND K. BINMORE (2000): "Toeholds, takeovers and football," *European competition law review*, 21(2), 1–4.
- KLEMPERER, P. (1998): "Auctions with almost common values: The Wallet Game and its applications," *European Economic Review*, 42(3), 757–769.
- KRISHNA, V. (2009): *Auction theory*, vol. 94. Academic press.
- SINGH, R. (1998): "Takeover bidding with toeholds: The case of the owner's curse," *Review of Financial Studies*, 11(4), 679–704.

Appendix

2.1 Number and Value of Announced Mergers & Acquisitions, World-wide, 1985-2013*



Source: <http://www.imaa-institute.org/statistics-mergers-acquisitions.html>

2.2 Proof of Lemma 1

I show that neither bidding more, nor bidding less than one's value can be optimal:

- (i) Overbidding¹⁹ only has an impact if $v_j < b_i < b_j'$. Bidder j then wins an auction he would otherwise have lost and makes a profit of $v_j - b_i < 0$ instead of 0, which leads to a loss of $b_i - v_j$ and thus does not make him better off.
- (ii) Underbidding is consequential only if $b_j' < b_i < v_j$. In this case, bidder j loses the auction and foregoes a profit of $v_j - b_i$ which he would have received if he had bid his private value and therefore would have won.

q.e.d.

2.3 The optimal bid function in the common value setting

First, notice that

$$\begin{aligned} \frac{\partial}{\partial s}[(1-s)^{\frac{1}{\theta}} b(s)] &= -\frac{1}{\theta}(1-s)^{\frac{1}{\theta}-1} b(s) + (1-s)^{\frac{1}{\theta}} b'(s) \\ &= \frac{1}{\theta}(1-s)^{\frac{1}{\theta}-1} [-b(s) + (1-s)\theta b'(s)]. \end{aligned}$$

Multiplying both sides of (3) by $\frac{1}{\theta_j}(1-s)^{\frac{1}{\theta_j}-1}$, we obtain

$$\frac{\partial}{\partial s}[(1-s)^{\frac{1}{\theta_j}} b_i(s)] = -v(s, \phi_j(s)) \frac{1}{\theta_j} (1-s)^{\frac{1}{\theta_j}-1}.$$

Integrating over s on $[t_i, 1]$ yields

$$-(1-t_i)^{\frac{1}{\theta_j}} b_i(t_i) = -\frac{1}{\theta_j} \int_{t_i}^1 v(s, \phi_j(s)) (1-s)^{\frac{1}{\theta_j}-1} ds,$$

¹⁹In this context, overbidding (underbidding) means submitting a bid that exceeds (falls short of) the valuation for a share of the target company.

so that the bid function reduces to

$$b_i(t_i) = \frac{\int_{t_i}^1 v(s, \phi_j(s))(1-s)^{\frac{1}{\theta_j}-1} ds}{\theta_j(1-t_i)^{\frac{1}{\theta_j}}}.$$

Now we can derive ϕ_i and ϕ_j to further simplify and substantiate the expression.

When t_i and t_j are realized signals, in a Nash equilibrium (3) implies

$$b'_i(t_i) = \frac{b_i(t_i) - v(t_i, \phi_j(t_i))}{(1-t_i)\theta_j}, \quad (5)$$

$$b'_j(t_j) = \frac{b_j(t_j) - v(t_j, \phi_i(t_j))}{(1-t_j)\theta_i}. \quad (6)$$

Dividing (A.1) by (A.2) yields

$$\frac{b'_i(t_i)}{b'_j(t_j)} = \frac{(1-t_j)\theta_i}{(1-t_i)\theta_j}.$$

This relationship will be the essential ingredient that makes the resulting equilibrium unique, in contrast to the (standard) model without toeholds, where we have a continuum of equilibria that depend on the concrete form of $\phi_k(\cdot)$.

Further observe that $b_i(t_i) \equiv b_j(\phi_j(t_i))$ implies $b'_i(t_i) = b'_j(\phi_j(t_i))\phi'_j(t_i)$.

Therefore, we derive

$$\begin{aligned} \phi'_j(t_i) &= \frac{(1-\phi_j(t_i))\theta_i}{(1-t_i)\theta_j}, \text{ or equivalently} \\ \frac{d}{dt_i} \ln(1-\phi_j(t_i)) &= -\frac{\theta_i}{(1-t_i)\theta_j}. \end{aligned}$$

Integrating this expression using $\phi_j(0) = 0$ gives

$$\begin{aligned} \ln(1 - \phi_j(t_i)) &= -\frac{\theta_i}{\theta_j} \int_0^{t_i} \frac{1}{(1-s)} ds \\ &= \frac{\theta_i}{\theta_j} \int_0^{t_i} d\ln(1-s) = \frac{\theta_i}{\theta_j} \ln(1-t_i), \end{aligned}$$

which implies

$$\begin{aligned} (1 - \phi_j(t_i))^{\theta_j} &= (1 - t_i)^{\theta_i} \\ \Leftrightarrow \phi_j(t_i) &= 1 - (1 - t_i)^{\frac{\theta_i}{\theta_j}}. \end{aligned} \quad (7)$$

2.4 The optimal bid function in the linear example

Applying the linear example, the optimal bid function is

$$\begin{aligned} b_j(t_j) &= \frac{\int_{t_j}^1 [s + 1 - (1-s)^{\frac{\theta_j}{\theta_i}}] (1-s)^{\frac{(1-\theta_i)}{\theta_i}} ds}{\theta_i (1-t_j)^{\frac{1}{\theta_i}}} \\ &= \frac{1}{\theta_i (1-t_j)^{\frac{1}{\theta_i}}} \left[\int_{t_j}^1 s (1-s)^{\frac{1-\theta_i}{\theta_i}} ds + \int_{t_j}^1 (1-s)^{\frac{1-\theta_i}{\theta_i}} ds - \int_{t_j}^1 (1-s)^{\frac{1+\theta_j-\theta_i}{\theta_i}} ds \right]. \end{aligned}$$

Solving the integrals one by one yields:

(i)

$$\begin{aligned} \int_{t_j}^1 s (1-s)^{\frac{1-\theta_i}{\theta_i}} ds &= \left[-\theta_i (1-s)^{\frac{1}{\theta_i}} s \right]_{t_j}^1 + \int_{t_j}^1 \theta_i (1-s)^{\frac{1}{\theta_i}} ds \\ &= \theta_i (1-t_j)^{\frac{1}{\theta_i}} t_j - \theta_i \left[\frac{\theta_i}{1+\theta_i} (1-s)^{\frac{1+\theta_i}{\theta_i}} \right]_{t_j}^1 \\ &= \theta_i (1-t_j)^{\frac{1}{\theta_i}} t_j + \frac{\theta_i^2}{1+\theta_i} (1-t_j)^{\frac{1+\theta_i}{\theta_i}}, \end{aligned}$$

(ii)

$$\int_{t_j}^1 (1-s)^{\frac{1-\theta_i}{\theta_i}} ds = - \left[\theta_i (1-s)^{\frac{1}{\theta_i}} \right]_{t_j}^1 = \theta_i (1-t_j)^{\frac{1}{\theta_i}},$$

(iii)

$$\int_{t_j}^1 (1-s)^{\frac{1+\theta_j-\theta_i}{\theta_i}} ds = - \left[\frac{\theta_i}{1+\theta_j} (1-s)^{\frac{1+\theta_j}{\theta_i}} \right]_{t_j}^1 = \frac{\theta_i}{1-\theta_j} (1-t_j)^{\frac{1+\theta_j}{\theta_i}}.$$

Recollecting terms, we have:

$$b_j(t_j) = t_j + \frac{\theta_i}{1+\theta_i} (1-t_j) + 1 - \frac{1}{1+\theta_j} (1-t_j)^{\frac{\theta_j}{\theta_i}}.$$

Tax changes in the Federal Republic of Germany from 1950-1970 and their effects on economic growth

Kaspar Zimmermann *

Introduction

Tax changes play a central role in economic policy making. For a good economic policy it is therefore crucial to have solid scientific knowledge about the effects of a tax change on the economy. Despite this importance, the debate in research about the size and impact of tax changes on the economy is still ongoing. One of the key problems for researchers in this area is simultaneity: a tax change is likely to influence both the current and future economic development. Additionally, the variables used to measure the tax change (e.g. the tax revenue of a country) are also influenced by the economic situation (Cloyne, 2013, p. 1509).

There are two predominant approaches to solve the simultaneity problem and identify the impact of tax shocks. Blanchard and Perotti (2002) try to solve this problem of simultaneity by using a structural vector autoregression (SVAR) to identify tax shocks, that are uncorrelated with the main economic variables. They assume that a policy reaction to a shock is not happening in the same quarter

*Kaspar Zimmermann received his degree (B.Sc.) from the University of Bonn in August 2013. The present article refers to his bachelor thesis under supervision of Prof. Dr. Moritz Schularick, which was submitted in August 2013.

as the shock itself. Equipped with this assumption they are able to construct a time series of cyclically adjusted tax revenues. Using US data, they estimate that cutting taxes by one percent of GDP increases GDP by one percent.

Romer and Romer (2010) present a new approach to overcome the simultaneity problem. They construct a new time series of legislated exogenous tax changes by using a narrative approach. An exogenous tax change is a tax change that is not correlated with economic indicators influencing the economic situation in the short run. To construct the time series and to separate exogenous and endogenous tax changes, historic documents such as congressional reports or presidential speeches are used. This narrative approach was used for the first time by Friedman and Schwartz (1963) in their seminal book *A Monetary History of the United States, 1867-1960* studying monetary policy. Romer and Romer (1989, 2004) further develop this approach in their studies on the monetary policy of the Federal Reserve.

Romer and Romer (2010) find a big and permanent effect of tax changes on GDP. A tax cut by one percent of GDP increases GDP by approximately 3 percent after three years. By now other researchers have applied their new approach to other countries. Cloyne (2012, 2013) constructs a dataset for Great Britain from 1949 to 2009 and finds an effect of around 2.5 percent of GDP after three years. Hayo and Uhl (2013) and Uhl (2013) find a maximal effect on 2.4 percent using a dataset for Germany from 1974 to 2009.

This thesis constructs a new dataset for Germany from 1950 to 1970 applying the narrative approach to post World War 2 Germany and uses this time series to estimate the impact of tax changes on economic growth. The results for the new dataset are similar to those of Romer and Romer (2010) and Cloyne (2013), but the confidence intervals are too wide to make broad claims.

Identification strategy

The following section describes the identification strategy of this paper using an easy model of how tax changes affect output growth. As already pointed out the main focus lies on the solution of the simultaneity problem:

$$\Delta Y_t = \alpha + \beta \Delta T_t + \epsilon_t \quad (1)$$

where ΔY_t is a measure for the real changes of economic activity in period t ¹, ΔT_t a measure of the tax changes in period t and ϵ_t captures other influences on output. Output is also influenced by previous tax changes but since the identification strategy stays the same with or without them they are ignored here for simplicity.

Economic activity is influenced by many factors besides tax changes such as the monetary policy of the central bank or changes in government expenditure. All these possible individual factors ϵ_t^i are included in ϵ_t ($\epsilon_t = \sum_{i=1}^K \epsilon_t^i$). We can think of some ϵ_t^i 's to be correlated and others to be uncorrelated with each other.

ΔT_t can be decomposed into:

$$\Delta T_t = \sum_{i=1}^K b_t^i \epsilon_t^i + \sum_{j=1}^L \omega_t^j, \quad (2)$$

where ω_t^j are further influences on tax policy. Therefore some ϵ_t^i have a direct effect on output via (1) and on the tax changes via (2) at the same time making a direct identification impossible. The key idea from Romer and Romer is to

¹Romer and Romer (2010) use the difference of the logarithms of real GDP while in this paper the real growth rate of GNP is used. In the thesis a section is dedicated to the construction of this seasonally adjusted quarterly series of real GNP for Germany from 1949 to 1970. This series is based on the "Vierteljahrshefte zur Wirtschaftsforschung" of the German Institute of Economic Research (DIW). While this section is omitted here, the appendix contains Figure 1 summarizing the results.

separate these tax changes from those which are uncorrelated with the economic indicators and can therefore be seen as exogenous. We assume therefore by construction that every individual exogenous ω_t^j is uncorrelated with the ϵ_t^i 's and the b_t^i 's.

Romer and Romer (2010) propose a new approach to identify ω_t by looking at legislated tax changes. These are separated events such that an individual motivation can be attributed to each change. Some of these motivations are uncorrelated with the economic activity in the short run. In the following, these changes in the legal code will be called exogenous tax changes and are the key measure of tax changes used in this paper. Tax changes which are correlated with economic indicators will be called endogenous. With this separation it is possible to construct a dataset of $\sum \omega_t^j$ and $\sum b_t^i \epsilon_t^i$. To estimate the effects of a tax change on output equation (3) with $v_t = \sum_{i=1}^K (1 + \beta b_t^i) \epsilon_t^i$ is used. Based on the arguments from above it is possible to conclude that a simple least squares regression of $\sum \omega_t^j$ on output should lead to an undistorted estimate of these tax changes. Therefore equation (3) represents the theoretical foundation of this approach.

$$\Delta Y_t = \alpha + \beta \sum_{j=1}^L \omega_t^j d + v_t \quad (3)$$

Based on these considerations, several validity checks can be used to test if an undistorted regression is possible and if the ω_t are truly exogenous. When estimating the effects of a tax change, controlling for variables influencing current and short run economic performance should not change the results of the regression. Further, it should not be possible to predict exogenous tax changes using economic indicators.

Construction of the new dataset

The individual analysis of each significant legislated tax change is given in the appendix of my thesis. Here I only provide a brief and very condensed overview of the construction method.

For 1965-1970, I have good detailed data based on publications of the Ministry of Finance. For 1950-1964 I had to use a wide range of legislative as well as secondary sources to construct a complete dataset of all major tax changes. I only include tax changes increasing or decreasing expected revenue by more than 50 million Mark per year from 1950-1959 and 100 million Mark from 1960-1970. The biggest entry barrier for a tax change to be included in the analysis is therefore around 0.1 percent of GDP for 1950 and 0.065 percent of GDP for 1960. Several assumptions were needed regarding size and timing of each individual tax change. In the constructed dataset each tax change is dated when liabilities actually changed, however the promulgation date of each law is reported in the narrative analysis such that future research can also account for anticipation effects. As an estimate of the size of a tax change the expected change in revenue of the government in the following year is used.

The crucial part in the construction process is the categorization of tax changes into four different types of motivation (Spending-driven, Countercyclical, Deficit-driven and Long-run). Romer and Romer argue that all tax changes in their sample could be assigned to one of their four categories while Cloyne (2012) works with 8 subcategories². I use the four subcategories by Romer and Romer.

Spending-driven and countercyclical motivations are contingent on the economic conditions and therefore endogenous. Often taxes are increased to finance

²The four endogenous motivations used by Cloyne (2012) are demand management, supply stimulus, deficit reduction/balance of payment crisis and spending driven. He classifies the exogenous tax changes into the following categories: long-run performance, ideological, external and deficit consolidation.

new expenditure directly. For example the *Verkehrsfinanzgesetz 1955* increased the tax on oil to finance new investment on infrastructure. This type of tax change is classified as spending-driven. Tax changes with a countercyclical motivation were particularly common during the last years of the time series, when the federal government tried to control the economy using Keynesian policies. Also reactions to exogenous shocks to the economy, such as the construction of the Berlin wall on 13th August 1961 or the appreciation of the Mark on 27th October 1969 are classified as countercyclical in the time series.

Tax changes which are implemented to reduce the deficit or have a long-term objective can be seen as exogenous. An inherited budget deficit or budget surplus can be explained by spending decisions and economic developments of the past. Therefore a tax change to balance the budget does not depend on the current economic situation and should be classified as exogenous. In some cases it can be difficult to separate deficit-driven changes from those legislated to finance increased spending. In case of doubt they are classified as endogenous. Often politicians want to increase the long run growth perspectives or change the legal code due to ideological or distributional concerns. These tax changes are classified as “Long-run changes”. The removal of economic obstacles or a more efficient tax code - both justifications of the *Große Steuerreform 1955*- are typical goals of a tax change classified into the long-run category. Tax changes motivated by equity or redistributive concerns are also included in this category.

Of the 45 changes, 31 are categorized as exogenous, 12 as endogenous and 2 consist of exogenous and endogenous parts. Figure 2 shows all exogenous tax changes relative to the nominal GDP. A negative value represents a tax cut and a positive a tax increase. In many quarters no changes occur. This can be explained by the fact that only legislated tax changes are used and in these quarters no significant change in the tax code was implemented. Exogenous tax

changes occurred over the whole period but the figure shows that most big changes occurred right at the start of our sample. The largest tax change occurred in the first quarter 1955 with a tax cut of 1.8 percent of GDP legislated in the *Große Steuerreform*. The arithmetic mean of the new sample is -0.08 percent of GDP and the standard deviation is 0.29 percentage points.

The series in Figure 3 are less volatile since only 14 legislated tax changes contain endogenous parts. At the end of the time series more tax changes take place. This reflects the shift to a more demand-side-oriented policy perspective of the federal government. The largest endogenous tax change took place in 1951, with a tax increase of 1.5 percent of GDP³. The arithmetic mean of this time series is 0.04 percentage points and the standard deviation is 0.21.

Effects of tax changes on economic growth

The effects of a tax change are estimated using the following regression:

$$\Delta Y_t = \alpha + \sum_{i=0}^{12} b_i \Delta T_{t-i} + \epsilon_t \quad (4)$$

where ΔY_t is the growth rate of real GNP in Quarter t and ΔT_t is the exogenous legislated tax change in quarter t . It is likely that output is also influenced by previous tax changes, so I include 12 lags of the tax series. If the constructed time series is really exogenous we should get undistorted estimates of the effects of tax changes on output using this regression.

In a second step lagged economic growth is included in the regression:

$$\Delta Y_t = \alpha + \sum_{i=0}^{12} b_i \Delta T_{t-i} + \sum_{j=1}^{12} c_j \Delta Y_{t-j} + \epsilon_t \quad (5)$$

This should control for the basic dynamics of GNP and several other factors

³ *Gesetz zur Änderung des Umsatzsteuergesetzes und des Beförderungsteuergesetzes*

correlated with it. Further this setup can be seen as a test of hidden motivation. Politicians could have had endogenous motives when changing the tax code, but justified it with arguments which are classified as exogenous. As a consequence the effects estimated in regression (4) could be distorted by the dynamics of the economy e.g. a return to normal growth rates after a crisis. As the theory discussed above suggests we should not observe big differences between both regressions.

Figure 4(a) illustrates the results of regression (4) showing the estimated effects of a tax increase of one percent of GDP on real GNP together with the 68 percent confidence interval. This corresponds to the common practice in similar papers (Hayo and Uhl, 2013, p. 13). A one percent tax increase has a negative effect of around -0.5 percent on GNP in the first quarter. This effect fluctuates over the course and reaches a maximal negative effect of -1.54 ($t = -0.86$) after 12 quarters. In all quarters the confidence interval reaches from a positive upper bound to a negative lower bound. Therefore, we cannot be certain about the sign or size of the multiplier.

Figure 4(b) shows the results of estimating regression (5), which controls for lagged GNP. Compared with figure 4(a), the impulse response function is shifted down. The effect in the first quarter is around -0.4 percent. The maximal effect of -2.87 ($t = -1.05$) is reached after 12 quarters. However, the estimates have again very wide confidence intervals, even though only 68 percent confidence intervals are included. This could be explained by the relatively short time horizon, including only 21 years. Romer and Romer (2010) include data for 57 years in their regression and Cloyne (2013) constructs a dataset for 54 years. A further discussion of the differences of the two regressions is done at the end of this paper.

The thesis presents a variety of different regressions and specifications. As a first step the two exogenous motivations, deficit-driven and long-run tax changes,

are separated and individual time series are constructed. For each of these series both Regression (4) and (5) are estimated. Figure 5 illustrates the results for all four different specifications. For long-run tax changes the effects are persistent and negative, reaching a maximal amount of -2.70 percent ($t = -0.87$) or -5.04 percent ($t = -1.06$) respectively. Deficit-driven tax changes do not seem to share this feature but instead fluctuate around zero in both specifications.

Apart from the main regressions, several robustness checks are included and presented in the thesis. Using different time series as a measure of output does not seem to change the results dramatically, but changing the investigated time period changes the results remarkably. The effects of a regression only including data from 1953Q1 to 1965Q4 are strong and persistent. The effects for the 1958 to 1970 interval are entirely different. After initial negative effects in the first quarters a positive effect is observed. Big difference in the two estimations could be explained by the dramatic changes in trend growth during the examined period, with high growth episodes at the beginning and normal growth episodes at the end of the series. Further the very high effects in the first years could be partly explained by the fact that several large cuts of the marginal tax rate were legislated in the years after the war. Further, the *Große Steuerreform 1955* seems to have a big impact on the results.

Conclusion

The thesis presents the first narrative analysis of all significant legislated tax changes from 1950 to 1970 following the Romer and Romer (2010) approach. This approach enables us to categorize tax changes in those whose legislation was connected to the economic conditions (exogenous) and those which were implemented for other reasons (endogenous). Given this narrative analysis, a new dataset of tax changes as a percentage of GDP in Germany is constructed.

This allows us to estimate the economic effects of the tax changes in a second step.

In all main regressions presented in the thesis, the hypothesis that the effect of a tax change on the economic condition is zero or close to it, cannot be rejected. The estimations presented in this summary can therefore only provide very limited information about the size of the tax multiplier. Still, the estimated coefficients point in the direction of Romer and Romer (2010) estimating higher coefficients than under the Blanchard and Perotti (2002) method. A further caveat is created by the differences between the two central regressions which the theory predicts to. This means that the short run dynamics of economic growth influence the estimates of the tax multiplier. Since the dataset was constructed with the aim of being independent from the short run changes in output, this finding suggests the identified tax changes may not be truly exogenous. Possible explanations could be hidden motivation of the politicians when justifying their policy, or sources and factors which were not taken into account. Another possible explanation is presented in the robustness checks of the paper. When the *Große Steuerreform 1955* is excluded from the regression, the impulse response functions of both specifications have a common course.

The obvious extension of this thesis is the construction of a dataset from the Second World War to present day, to get better estimates and confidence intervals. This larger dataset could build on this thesis, as well as the work by Uhl (2013), but should not neglect the years from the end of the Second World War to 1950 since several interesting changes in the tax code occurred during this time period. This new dataset could be rich enough to compare the effects of changes of the different tax types as proposed by Mertens and Ravn (2013). Another extension would be to investigate the effects of tax changes on the different components of GDP, as done by Romer and Romer (2010) and Cloyne (2013). Quarterly data

could be constructed using the *Vierteljahrshefte zur Wirtschaftsforschung* which were also used for this paper.

The Federal Republic of Germany was in an individual unique situation after the Second World War. Therefore it is unlikely that tax changes today would have the same effects as in the investigated time period. This paper should be seen as an historic assessment of tax changes and their effects in the period of the *Wirtschaftswunder* and provides only limited information to the questions of tax policies today.

References

- BLANCHARD, O., AND R. PEROTTI (2002): “An empirical characterization of the dynamic effects of changes in government spending and taxes on output,” *Quarterly Journal of Economics*, 117(4), 1329–1368.
- BUNDESARCHIV (several years): *Kabinettsprotokolle der Bundesregierung, online 1949-1967*. <http://www.bundesarchiv.de/cocoon/barch/0000/k/index.html>, accessed at 17.08.2013, several meetings.
- BUNDESMINISTERIUM DER FINANZEN (1954 - 1960): *Allgemeine Vorbemerkungen zum Bundeshaushaltsplan für das Rechnungsjahr*. Bundesdruckerei, Bonn, several volumes.
- (1961-1971): *Finanzbericht*. Stollfuß Verlag, Bonn, several volumes.
- (2012): *Zusammenstellung der Steuerrechtsänderungen seit 1964*. http://www.bundesfinanzministerium.de/Content/DE/Standardartikel/Themen/Steuern/Weitere_Informationen/zusammenstellung-der-steuerrechtsaenderungen-seit-1964.html, accessed at 17.08.2013.
- BUNDESTAG (1949-1970a): *Plenarprotokolle, Stenographische Berichte*. Heger, Bonn, several volumes.
- (1949-1970b): *Drucksachen*. Heger, Bonn, several volumes.
- CLOYNE, J. (2012): “Discretionary tax shocks in the United Kingdom 1945 – 2009: a narrative account and dataset,” <http://ideas.repec.org/e/pc185.html>, zugegriffen am 17.08.2013.

- (2013): “Discretionary tax changes and the macroeconomy: new narrative evidence from the United Kingdom,” *American Economic Review*, 104(4), 1507–1528.
- DEUTSCHES INSTITUT FÜR WIRTSCHAFTSFORSCHUNG (1951-1971): *Vierteljahrshefte zur Wirtschaftsforschung*. Duncker & Humblot, Berlin, mehrere Bände.
- DZIADKOWSKI, D. (2003): “50 Jahre Reformversuche bei der Einkommensteuer – Von der “Kleinen Steuerreform” bis zur vorgezogenen Steuerentlastung 2003,” in *50 Jahre Steuerreformen in Deutschland*, ed. by Akademie für Steuer- und Wirtschaftsrecht. Stollfuß Verlag, Bonn.
- FRANZEN, K. (1994): *Die Steuergesetzgebung der Nachkriegszeit in Westdeutschland (1945-1961)*. Kamloth, Bremen.
- FRIEDMAN, M., AND A. J. SCHWARTZ (1963): *A monetary history of the United States, 1867-1960*. Princeton University Press, Princeton.
- HAYO, B., AND M. UHL (2013): “The macroeconomic effects of legislated tax changes in Germany,” *Oxford Economic Papers*, Forthcoming.
- KRÜER-BUCHHOLZ, W. (1982): *Steuerpolitik und Steuerreformen in der Bundesrepublik Deutschland unter besonderer Berücksichtigung der “Großen Steuerreform” 1975: eine Untersuchung der Steuerreformmaßnahmen in ihrem historischen Zusammenhang, ihren wirtschaftlichen und politischen Grenzen sowie ihren sozialen Auswirkungen*. Minerva Publikation, München, (zugleich Dissertation Universität Bremen 1978).
- MERTENS, K., AND M. O. RAVN (2013): “The dynamic effects of personal and corporate income tax changes in the United States,” *American Economic Review*, 103(4), 1212–1247.
- MUSCHEID, J. (1986): *Die Steuerpolitik in der Bundesrepublik Deutschland 1949 - 1982*. Duncker & Humblot, Berlin, Volkswirtschaftliche Schriften, Heft 365.
- PRESSE- UND INFORMATIONSAMT DER BUNDESREGIERUNG (1949-1970): *Bulletin*. Deutscher Bundes-Verlag, Bonn, several volumes.
- ROMER, C., AND D. H. ROMER (1989): “Does monetary policy matter? A new test in the spirit of Friedman and Schwartz,” *NBER Macroeconomics Annual* 1989, 4, 121–184.
- (2004): “A new measure of monetary shocks: derivation and implications,” *American Economic Review*, 94(4), 1055–1084.
- (2009): “A narrative analysis of postwar tax changes,” <http://emlab.berkeley.edu/users/dromer/papers/nadraft609.pdf>, zugegriffen am 17.08.2013.

——— (2010): “The macroeconomic effects of tax changes: estimates based on a new measure of fiscal shocks,” *American Economic Review*, 100(3), 763–801.

UHL, M. (2013): “A history of tax legislation in the Federal Republic of Germany,” http://www.uni-marburg.de/fb02/makro/forschung/magkspapers/11-2013_uhl.pdf, zugegriffen am 17.08.2013.

Appendix

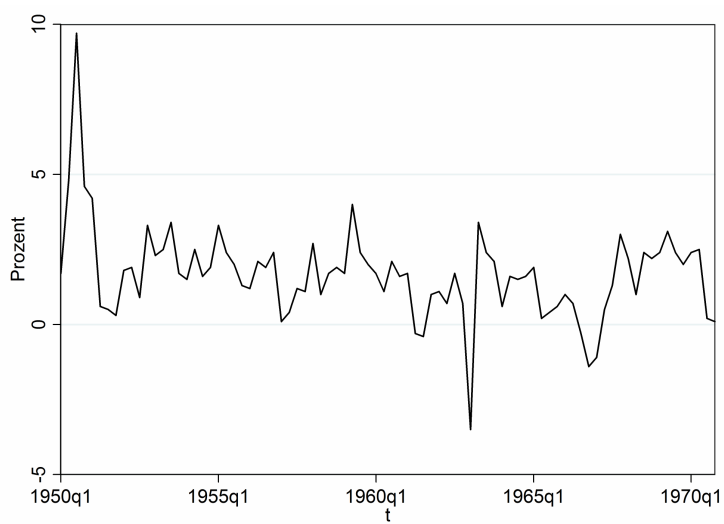


Figure 1: Change of Gross National Product



Figure 2: Exogenous tax changes relative to nominal annual GDP



Figure 3: Endogenous tax changes relative to nominal annual GDP

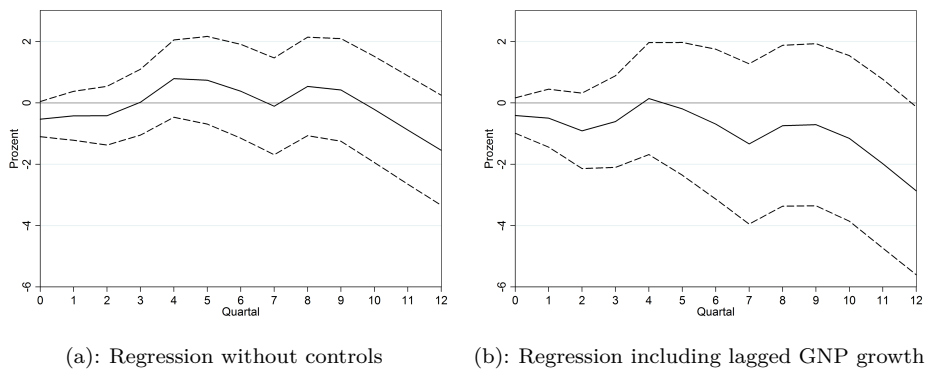


Figure 4: Estimated effects of a tax increase of one percent of GDP on GNP

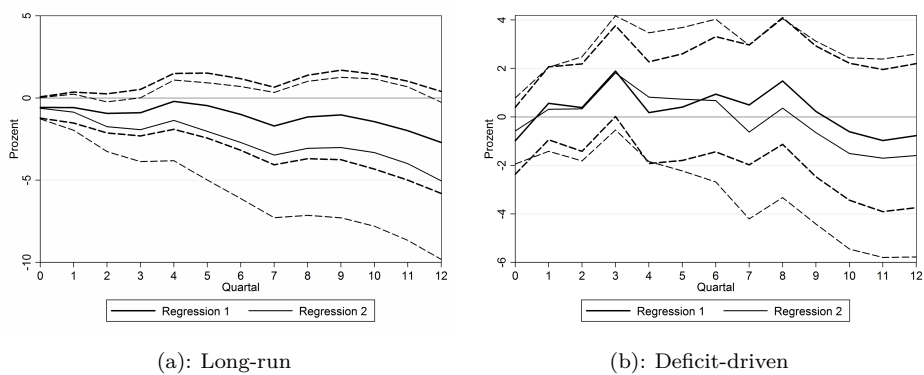


Figure 5: Estimated effects of a tax increase of one percent of GDP on GNP

Sources of Fiscal Risk in the euro area

Prof. Dr. Jürgen von Hagen *

Introduction

The sovereign debt crisis that started in the euro area in 2010 was not supposed to happen. Building on the “Excessive Deficit Procedure” (EDP) of the Maastricht Treaty and the “Stability and Growth Pact” (SGP) of the late 1990s, the member states of the euro area had vested themselves with an elaborate system of fiscal rules and processes based on a host of statistical indicators, all aiming at a high degree of fiscal discipline and the sustainability of public finances in all member states. Euro-Area governments must comply with conditional and unconditional fiscal targets and report annually on their fiscal strategies, intentions, policies, and outcomes. This machinery is watched over by the European Commission and Eurostat. The commitment to common numerical rules, been dubbed “government by statistics,”¹ was believed to compensate for the lack of a strong fiscal authority coordinating the fiscal policies of the member states, an institutional deficiency that had been criticized especially by economists in the US.² The EU has recently expanded the scope and the depth of “government by

*Department of Economics, Institute for International Economic Policy, University of Bonn, Lennéstrasse 37, 53113 Bonn, Germany, phone: +49 (0)228 73-9199, email: vonhagen@uni-bonn.de

¹(Pisani-Ferry, 2010, p.2)

²See Jonung and Drea (2010) for a summary of the debate in the US

statistics” through the “Six Pack” and the “Two Pack.” The “Six Pack,” which entered into force on 13 December 2011, strengthens the enforcement of the EDP and the SGP and adds new rules for government spending and a new “Macroeconomic Imbalance Procedure” also based on a range of statistical indicators.³ The “Two Pack”, which entered into force on 30 May 2013, further develops the processes for monitoring the member states’ public finances.⁴

The debt crisis has revealed how much “government by statistics” has failed. A first reason for its failure is that the fiscal indicators on which the European framework relies are backwards-looking; they measure the outcomes of past policies but they provide only a very limited look into the future.⁵ The second reason is that the indicators focus on nominal budgetary flows and stocks of explicit financial liabilities of the government without taking into account contingent liabilities and their riskiness.⁶ Even 20 years after the Maastricht Treaty, it remains largely unclear how the rules would assure the sustainability of public finances and how deviations from them would endanger it.⁷ The economics underlying European Commission evaluations of the public finances of the member states are often unclear and the conclusions often seem to be driven by political considerations. If anything, “government by statistics” has created a culture of problem denial, allowing policymakers to argue that everything is fine as long as the numbers comply with the rules.⁸ In this paper, we review the origins of the

³European Commission (2011)

⁴European Commission (2013)

⁵This shortcoming of traditional fiscal indicators has long been recognized in the context of evaluating the fiscal consequences of aging societies. See e.g. Velculescu (2010).

⁶They also largely ignore the implications of budgetary operations on the net wealth of government. Moriyama and Milesi-Ferretti (2004) for example, point out that, during the run-up to EMU, most European governments seem to have limited the growth of gross debt by reducing gross assets, reducing net wealth and the government’s ability to react to unforeseen events as a result.

⁷Schubert (2013) reminds us that the core criteria for general government debt and deficits, 60% and 3% of GDP, were the outcome of a historical coincidence.

⁸An example is the budget debate in Portugal during October 2013, which features the argument that one of the budgetary risks for 2014 is the reclassification of state-owned enterprises as part of the general government sector (See Portuguese Ministry of Finance, 2013, p.78).

euro-area sovereign debt crisis in order to identify the sources of fiscal risk. On this basis, we draw some conclusions for the future design of the euro area's fiscal framework.

What makes a sovereign debt crisis?

The sovereign debt crisis hit six countries the hardest: Greece, Portugal, Ireland, Spain, Italy, and Cyprus. Of these, Greece, Ireland, Portugal, and Cyprus received official bail-outs from the EU and the IMF, and Spain received some financial assistance from the EU. What made these countries different from the rest of the euro area?

Asymmetric Shocks

The first and most popular answer to our question is that these countries had been hit by negative "asymmetric shocks." It is convenient, because it makes the crisis look like an exogenous event not related to domestic policies. This is the main argument for the creation of risk-sharing through intergovernmental transfers in the euro area and the main justification of the call for a fiscal union in Europe.

Table 1 shows the average real GDP growth rates for 2002-2006 and the growth rates for each year from 2007 to 2012 for the euro area as a whole and the six crisis countries. In the five years before the crisis, the euro area's average growth rate was 1.76 percent. Greece, Ireland, Spain and Cyprus had average growth rates above the euro area, while Italy and Portugal grew considerably less than the group as a whole. During the years of the financial crisis, 2007-2009, euro-area real GDP growth fell from 3.00 percent to -4.39 percent. Ireland was the only country among the six that suffered a real growth rate of one cross-section standard deviation less than the euro area average, and this only in 2008, when

Irish real GDP fell by 2.1 percent. Cyprus and Spain continued to grow faster than the euro area during the years of the financial crisis. Portugal and Greece did so in 2009. In 2008, these two countries had slightly lower growth rates than the euro area. Ireland and Greece grew faster than the euro area in 2007. Only Italy consistently had growth rates below the euro area in 2007-2009.

To sharpen the notion of an asymmetric shock, we calculate the difference in real GDP growth rates between 2007 and 2009. Table 1 shows that euro-area real GDP growth fell by 7.49 percent during that period. We then subtract this difference from the each country's difference growth rates between 2007 and 2009. This difference in difference is negative only for Ireland during 2007-2009, and even there is within one cross-section standard deviation from the euro-area average. For Greece, Spain, Italy, Cyprus, and Portugal, the result is positive, indicating that these economies experienced small *positive* asymmetric shocks during the period of the financial crisis. The crisis countries were not hit by larger negative shocks than the euro area on average.⁹

A possible counter-argument to this is that the six countries considered here were indeed hit by larger negative shocks than the euro area on average, and that these shocks are not reflected in their real GDP growth rates because of their larger fiscal stabilization efforts. If so, one would expect that these countries had much larger cyclical deficits than the euro area had on average during 2007-2009. Table 2 shows that this is not the case. Only Ireland and Spain had cyclical deficits larger than the euro area, but the differences are marginal. Greece, Ireland and Spain had much larger cyclical deficits than the euro area in 2010 and 2011, but this can be attributed to the fiscal adjustments following the emergence of the public debt crisis. Between 2009 and 2011, the euro area

⁹I do not argue here that business cycles in the euro-area were symmetric. The countries that were hit by significant asymmetric shocks in 2007-2009 were Estonia, Slovenia, Slovakia, and Finland.

returned to a real GDP growth rate of 1.44 percent. It is in this period, that several crisis countries experienced negative asymmetric growth. As shown in Table 1, Greece had a very strong negative deviation from euro-area growth. In Portugal and Cyprus, the deviation was sizable, though less than one standard deviation away from the euro-area growth rate. Yet, in view of the strong fiscal contractions in these countries during that period, their negative growth performance cannot be attributed to exogenous developments. Ireland and Italy had stronger improvements in real growth than the euro area during those years.

One version of the asymmetric-shocks argument attributes the debt crisis to the volatility of international capital flows. Greece, Ireland, Portugal, and Spain experienced very sizable capital inflows in the years preceding the financial crisis that suddenly reversed during the crisis (see Table 5), leading to a sharp decline in aggregate demand.¹⁰ Portfolio capital inflows already dried up Spain in 2007–2008 and turned into moderate net outflows in 2011 and 2012. Ireland’s financial balance turned from (–19.1) percent of GDP in 2007 to 0.7 percent of GDP in 2009. Greece and Portugal followed with a strong turnaround of portfolio inflows from (–12.4) to 9.6 percent of GDP in Greece and from (–9.0) percent of GDP to 5.6 percent in Portugal between 2009 and 2010. Their sharp turnarounds were offset by the building up of large negative balances within the European TARGET2 system, so that the overall balance on the financial account remained largely unaffected by market developments.¹¹ In Cyprus, capital inflows fell by one half as a percentage of GDP between 2010 and 2011, but the country was still able to attract net inflows. Italy, finally, experienced a more moderate reversal of portfolio inflows by 4.7 percent of GDP between 2010 and 2011, which again was offset partially, at least, by “other investment” inflows. These so-called sudden stops indicate that private investors became increasingly weary of

¹⁰See Merler and Pisani-Ferry (2012) for a detailed account of the European sudden stops.

¹¹See Sinn and Wollmershäuser (2011)

financing persistent current account deficits.

Still, the crisis countries were not the only countries facing such turnarounds, and their sudden stops of capital inflows were not the most severe ones in the euro area, either. Between 2008 and 2009, Estonia and Latvia suffered sudden increases in their financial account balances by 17.5 and 16.7 percent of GDP, respectively. Considering private portfolio investment, the Netherlands and Austria experienced turnarounds of 13.1 and 11.8 percent of GDP between 2008 and 2009, respectively, Belgium and Malta had turnarounds of 14.1 and 11.8 percent of GDP between 2009 and 2010. Thus, with regard to sudden stops the crisis countries are no exceptions in the euro area. We conclude that the asymmetric-shock hypothesis does not do much to explain the debt crises in Europe.

Asymmetric Policies

Table 2 reports the average structural and cyclical deficits for the euro area and the six countries considered here on average for 2003-2006 and each year after 2006. Ireland and Spain had very strong fiscal positions with considerable structural surpluses in the years before the financial crisis. In contrast, Greece, Italy, Portugal, and Cyprus all had sizable structural deficits exceeding three percent of GDP during that period. They approached the global financial crisis with relatively weak fiscal balances. In 2007 and 2008, the euro area on average and all crisis countries except Italy still showed comfortable cyclical budget surpluses.

Table 3 considers the patterns of adjustment in fiscal aggregates during the crisis. We compute the changes in the ratios of general government revenues, general government expenditures, various expenditure categories and primary deficits to GDP. Again, we compare the euro area average with the six crisis countries. Thus, Table 3 gives a difference-in-difference analysis of fiscal adjustments. Boldface numbers highlight country-specific differences in excess of one

cross-section standard deviation among the non-crisis euro area countries. We summarize the information of Table 3 in Table 4. Here, a dark cell indicates that the country's fiscal indicator under consideration significantly deviates from the average of the non-crisis euro-area during 2007-2009.

The table reveals several interesting points. First, there are some clear differences between Italy and Portugal on the one hand and Greece, Ireland, Spain, and Cyprus on the other. The first two countries do not show fiscal adjustments significantly different from the average of the euro area other than their increase in the gross debt ratio. The four other countries do show significantly different fiscal adjustments. The decline in general government revenues was much stronger than on average in the euro area and so was the increase in the primary deficits. This indicates that the tax systems of these countries had a much higher degree of cyclical elasticity than the average system in the euro area. In contrast, differences on the spending side of the budget are less clear-cut. For Greece and Spain, the increase in total spending was not significantly different from the euro-area average in 2007-2009, for Ireland it was and this is due to the government's efforts to bail out the main banks.

However, Greece, Ireland, Spain, and Cyprus all show significantly stronger increases in social benefits, government final consumption, and compensation of employees than the euro-area average. Spending in these areas is generally difficult to reverse for political reasons and, therefore, the adjustments translate into longer-lasting budgetary effects than what one would wish to fend-off a temporary recession. This suggestion is supported by the evidence given in the last row of the table. There, we show the share of the change in a country's structural deficit between 2007 and 2009 in the country's change in the overall deficit. A large share would indicate that most of the fiscal adjustment to counteract the recession following the financial crisis was undertaken by structural rather than

cyclical measures. While the average share of the structural deficit adjustment in the euro area was 42.5 percent, all crisis countries except Italy had shares above two-thirds, Greece, Portugal, Spain and Cyprus even had shares of 75 percent and above. Finally, Greece, Ireland, Spain and Cyprus also experienced a significantly stronger increase in interest expenditures compared to the euro-area average.

Thus, the crisis countries stand out for their strongly asymmetric fiscal policies compared to the euro-area average. They seem to have used relatively more sticky and structural fiscal policy tools than the rest of the group. The exception to this is Italy which, apart from a stronger increase in its debt ratio behaved in a not significantly different way compared to the average of the euro-area.

Banking Crises

After a long string of years with budget surpluses that had brought the debt ratio down to well under 30 percent, Ireland had realized budget deficits of 7.3 percent and 14.3 percent of GDP in 2008 and 2009 respectively. As a result of the financial crisis that started in 2007 and the collapse of a huge real estate bubble, the country faced a severe banking crisis.¹² At the end of September 2008, the government issued a blanket guarantee on all bank deposits, thus turning bank deposits into the equivalent of government debt. The government's fiscal operations providing funds to the main Irish banks caused its deficit to rise to 30.9 percent of GDP in 2010, of which 20.2 percent was due to the expenditures for bank support. The fiscal cost of recapitalizing the Irish banks amounted to 46.3 billion euros or 30 percent of Irish GDP in 2009-10. Lane (2013) estimates the total cost of bank recapitalizations to the Irish government during 2009-2011 at 41 percent of 2011 GDP.

¹²See Fernandez-Villaverde, Garicano, and Santos (2013), Honohan (2010) and Commission of Investigation (2011) for accounts of the Irish real estate bubble and financial crisis.

Spain also faced a banking crisis due to the collapse of house prices in the course of the global financial crisis of 2007-2008.¹³ Large-scale fiscal operations to support troubled banks would have caused further increases in public debt, and, anticipating this, markets seemed increasingly reluctant to lend to the Spanish government. In June 2012, shortly after the collapse of Bankia, an institution with assets amounting to one third of Spanish GDP, Spain requested and obtained financial assistance from the EU.

Similar to Spain and Ireland, Cyprus had enjoyed sizeable capital inflows for a number of years before the crisis, but these inflows started to dry up in 2011, putting the Cypriot banking system under intense pressure. The government's request for financial assistance in July 2012 thus came in the context of its efforts to prevent its banking system from collapsing. Cyprus was granted a financial program of € 9 billion from the ESM and € 1 billion from the IMF in March 2013, which included conditionalities for recapitalizing, restructuring and downsizing the banking sector.

Ireland, Spain, and Cyprus thus stand out for the link between a troubled banking sector and the government which was willing to rescue failing banks at the expense of the tax payer. From an economics point of view, the banking sector was a large conditional liability of the government that was not accounted for in the fiscal data. Moreover, the risk associated with these liabilities was more of a political than of an economic nature. Accounts of the Irish and the Spanish crises show that the political decisions to bail out the banks were not driven by economic rationale. Tight personal relationships between bankers and politicians seem to have been at the core of the problem.¹⁴

¹³See Fernandez-Villaverde, Garicano, and Santos (2013) for an account of the Spanish real estate boom and banking crisis.

¹⁴See e.g. Commission of Investigation (2011), Honohan (2010), Lane (2013), Kanda (2010), Fernandez-Villaverde, Garicano, and Santos (2013).

Conclusions

Three different patterns of sovereign debt crises arise from this review. The first, exemplified by Italy and Portugal, is a generally weak fiscal situation, one that might have been sustainable in normal times but turned out to be unsustainable after the economy had been hit by a negative shock. The second, exemplified by Greece, Spain, and Cyprus, is that of a significant decline in government revenues due to an adverse macroeconomic shock combined with a lack of sufficiently flexible tools on the expenditure side of the budget to react. The third, exemplified by Ireland, Spain, and Cyprus, is the unaccounted exposure to large contingent liabilities. These patterns point to three sources of sovereign risk in the euro area: (1), vulnerability to negative economic shocks due to persistent budgetary weaknesses in normal times, (2), excessive degrees of elasticity of tax revenues to negative economic shocks combined with the lack of sufficiently flexible fiscal instruments to offset them, and, (3), the exposure to unaccounted contingent liabilities intertwined with domestic politics. These sources are not addressed in any way by the EU's current "government by statistics," which, therefore, remains ineffective in protecting the euro area against future crises. Furthermore, all three sources of risk are the result of how national government revenue and expenditure systems are designed. Design choices reflect national fiscal preferences. If euro area member states reserve the right for themselves to make these decisions as they see fit, they should bear the consequences, since there is no accountability otherwise. Bailing out crisis-ridden governments only weakens the incentive to design and implement fiscal systems which are robust under the conditions of a monetary union. In a European fiscal union, governments would have to give up this right as a price for receiving bailouts.

References

- COMMISSION OF INVESTIGATION (2011): "Misjudging risk: Causes of the systemic banking crisis in Ireland," *Report of the Commission of Investigation into the Banking Sector in Ireland*, Government Publications, <http://www.bankinginquiry.gov.ie>.
- EUROPEAN COMMISSION (2011): "EU Economic Governance "Six Pack" Enters Into Force," *Memo/11/898*, 11 December.
- (2012): "Fiscal Sustainability Report," *European Economy*, 8.
- (2013): "Two Pack Enters Into Force, Completing Budgetary Surveillance Cycle and Further Improving Economic Governance in the Euro Area," *Memo/13/457*, 27 May.
- FERNANDEZ-VILLAYERDE, J., L. GARICANO, AND T. SANTOS (2013): "Political credit cycles: The case of the euro zone," Discussion paper, London:CEPR, CEPR Discussion Paper 9404.
- HONOHAN, P. (2010): "The Irish Banking Crisis Regulatory and Financial Stability Policy 2003-2008," *Report to the Irish Minister for Finance by the Governor of the Central Bank (Dublin)*, pp. 2003–2008, <http://www.bankinginquiry.gov.ie>.
- JONUNG, L., AND E. DREA (2010): "It Can't Happen, It's a Bad Idea, It Won't Last: U.S. Economists on the EMU and the Euro, 1989-2002," *Econ Journal Watch*, 7, 4–52.
- KANDA, D. (2010): "Asset Booms and Structural Fiscal Positions: The Case of Ireland," Working Paper 10/57, Washington DC: IMF.
- LANE, P. R. (2013): *Ireland: Macro-Financial Review*. Mimeo, Trinity College Dublin.
- MERLER, S., AND J. PISANI-FERRY (2012): "Sudden stops in the euro area," Discussion paper, Brussels: Bruegel Policy Contributions 2012/06.
- MORIYAMA, K., AND G. M. MILESI-FERRETTI (2004): *Fiscal adjustment in EU countries: a balance sheet approach*. Washington DC:IMF, Working Paper 04/143.
- PISANI-FERRY, J. (2010): "Euro-area governance: what went wrong? How to repair it?," Discussion paper, Bruegel policy contribution 2010/05.
- PORTUGUESE MINISTRY OF FINANCE (2013): "Orçamento do Estado para 2014," Discussion paper, Lisbon, October, Relatorio.

SCHUBERT (2013): “Wie das Maastricht Kriterium im Louvre entstand,” *Frankfurter Allgemeine Zeitung*.

SINN, H.-W., AND T. WOLLMERSHÄUSER (2011): “Target-Kredite, Leistungsbilanzsalden und Kapitalverkehr: Der Rettungsschirm der EZB,” Discussion paper, Ifo Working Paper.

VELCULESCU, D. (2010): *Some Uncomfortable Arithmetic Regarding Europe's Public Finances*. Working Paper 10/177, International Monetary Fund.

Appendix

	Real GDP Growth						Change in real GDP growth 2007-2009
	2002-06 Average	2007	2008	2009	2010	2011	2012
Cyprus	1.49	2.09	3.21	2.53	-0.68	-0.90	-1.82
Greece	2.55	0.54	-0.59	1.25	-6.93	-8.54	-5.77
Ireland	3.28	2.45	-2.48	-1.07	-2.76	0.00	1.55
Italy	-0.71	-1.32	-1.53	-1.11	-0.27	-1.06	-1.76
Portugal	-1.03	-0.63	-0.38	1.48	-0.05	-2.99	2.11
Spain	1.58	0.48	0.52	0.64	-2.31	-1.02	0.16
Euro Area	1.76	3.00	0.38	-4.39	1.99	1.44	-0.61
Std.Dev.	2.00	2.40	2.50	3.30	1.00	2.00	4.9

Note: Individual country data are growth rate differentials relative to the euro area. Std.Dev. is the cross-section standard deviation for the 11 non-crisis euro-area countries. Boldface entries indicate growth rates which are at least one standard deviation below the euro-area average. Source: European Commission AMECO

Table 1: Asymmetric Shocks

	Euro	GRE	IR	IT	SPA	POR	CYP	Euro	GRE	IR	IT	SPA	POR	CYP
	Structural Deficit (% of GDP)							Cyclical component of deficit (% of GDP)						
2003-06	2.80	6.31	-1.48	5.15	-0.46	5.33	4.34	-0.13	-0.07	0.17	-0.60	-0.53	0.43	0.17
2007	2.13	7.62	1.81	3.51	-0.93	3.73	-2.44	-1.41	-1.37	-1.89	1.72	-1.00	-0.47	-1.06
2008	2.95	9.62	7.57	3.84	4.46	4.49	0.79	-0.88	-0.70	-0.21	0.93	-0.21	-0.06	-1.72
2009	4.51	14.77	9.82	4.19	8.55	8.69	6.45	1.82	0.63	2.05	-1.97	2.05	1.21	-0.33
2010	4.46	8.83	9.11	3.70	7.41	8.81	5.70	1.06	2.31	2.25	-0.97	2.25	0.35	-0.41
2011	3.55	5.36	7.65	3.61	7.25	6.56	6.59	0.63	4.43	1.49	-0.87	1.49	0.85	-0.52
2012	2.14	0.96	7.41	1.43	5.46	4.17	6.72	1.12	5.80	0.65	-1.71	0.65	1.60	-0.01
2013	1.39	2.04	6.92	0.49	4.36	3.65	5.39	1.53	6.26	0.14	-2.21	0.14	2.09	2.70

Source: European Commission AMECO

Table 2: Structural and Cyclical Deficits

	Euro	GRE	IR	IT	SPA	POR	CYP	Euro	GRE	IR	IT	SPA	POR	CYP
	2007-2009													
Total Revenues	-0.4	-2.4	-2.2	0.5	-6.0	-1.5	-4.7	0.4	4.1	0.2	-0.3	0.6	5.4	-0.4
Total Expenditures	5.2	6.8	11.4	4.3	7.1	5.4	4.9	-1.8	-2.0	-0.5	-2.0	-1.1	-0.4	-0.3
Social Benefits not in kind	2.0	2.9	4.9	2.2	3.1	2.4	1.8	-0.3	1.3	0.5	0.1	0.7	0.3	1.3
Gen. Govt. Final Consum.	2.3	3.1	3.3	1.9	3.0	2.2	2.7	-0.7	-3.1	-1.9	-1.0	-0.4	-2.1	-0.1
Compensation of Employees	1.0	2.0	2.3	0.7	1.8	0.6	1.6	-0.5	-1.1	-0.9	-0.6	-0.4	-1.4	-0.2
Interest	-0.1	0.4	1.0	-0.3	0.2	-0.3	-0.4	0.2	1.2	1.2	0.2	0.7	1.2	-0.2
Primary deficit	5.8	8.5	12.9	4.2	12.9	7.1	10.1	-2.4	-8.1	-1.8	-2.0	2.4	-7.0	0.4
Share of structural def.	42.5	78.1	67.0	15.6	76.1	74.6	92.5	45.6	167.9	79.4	35.7	97.9	85.5	-264.5
Gross debt	13.6	22.5	39.8	13.1	17.6	15.3	-0.3	8.0	40.6	41.6	15.3	4.4	24.6	12.5

Note: Bold figures denote deviations from euro-area average in excess of one cross-section standard deviation among the non-crisis countries.
Source: European Commission, AMECO

Table 3: Changes in Selected Budgetary Aggregates (percent of GDP), 2007-2009 and 2009-2011

	Greece	Ireland	Italy	Portugal	Spain	Cyprus
Decline total revenues						
Increase total spending						
Social benefits						
Final consumption						
Personnel						
Interest						
Primary deficit						
Gross deficit						
Share structural deficit						

Source: Table 3. Lighter colors for Cyprus indicate significant differences in 2010-2011.

Table 4: Patterns of Fiscal Adjustments, 2007-2009

	2006	2007	2008	2009	2010	2011	2012
Balance on Financial Account (net borrowing -)							
Cyprus	-6.8	-11.5	-16.5	-10.9	-9.5	-4.3	-4.3
Greece	-9.8	-12.3	-12.9	-10.6	-9.4	-8.6	-2.1
Ireland	-2.7	-19.1	-9.1	0.7	-4.6	-3.8	1.3
Italy	-1.7	-1.7	-2.0	-2.5	-5.6	-4.6	-0.5
Portugal	-9.5	-8.5	-11.2	-10.4	-9.0	-5.7	1.1
Spain	-8.6	-9.6	-9.2	-5.0	-4.1	-2.7	0.1
Net Portfolio Inflows							
Cyprus	-8.0	-4.9	-23.4	6.0	-8.3	-4.8	-0.3
Greece	-3.5	-8.0	-7.2	-12.4	9.6	9.1	51.1
Ireland	-11.9	-11.3	21.3	-1.4	-2.3	3.7	3.2
Italy	-3.0	-1.2	-4.8	-1.8	-2.5	2.2	-1.9
Portugal	-2.4	-5.9	-9.0	-9.0	5.6	2.7	12.7
Spain	-18.9	-8.5	0.1	-4.8	-3.4	3.0	4.0

Note: All entries are percentages of GDP.

Source: IMF, Balance of Payments Statistics

Table 5: Sudden Stops

Contexts, Compliance and Imperfect Recall - Using Two-Sample IV to Estimate Average Causal Effects of Child Human Capital Investments

JProf Dr. Pia R. Pinger^{* *}

Abstract

Recently, much research has focused on determining critical and sensitive periods in child development. In this context, the use of non-experimental data can lead to biased estimates if parental investments and human capital outcomes jointly depend on unobserved confounders. To deal with this problem, many studies relate so-called contextual variation, in the form of events that are exogenous from the point of view of the individual, to later human capital outcomes. Yet, being a child when a contextual shock materializes does not necessarily imply individual suffering, such that the average causal effect of early-life investments on the individual cannot be determined from contextual information alone. This paper explains how instrumental variable estimation can be used to obtain causal estimates from contextual variation and information about individual human capital in-

^{*}Department of Economics, University of Bonn, University of Bonn, Adenauerallee 24-42, D-53113 Bonn, Germany, e-mail: pinger@uni-bonn.de, phone: +49 228 73-9206 or +49 228 73-9238 (secretariat).

^{*}This article draws on earlier work with Gerard van den Berg and Johannes Schoch (see van den Berg, Pinger, and Schoch, 2014). I thank Jonas Dovern for valuable comments.

vestments. It then discusses how combining information from two different samples can be used to compute causal effects in the presence of recall bias.

Introduction

Human capital is vital to individual success in life and constitutes one of the most valuable economic resources. But how and when does it develop, and what are the most important human capital investments? Finding answers to these questions is not an easy task. First, experiments are ideal from a scientific point of view, but not always ethical or possible. Second, recall bias is omnipresent, since most individuals do not fully recollect investments that took place in early childhood. In this context, the use of non-experimental data can lead to biased estimates, e.g. if childhood investments are imperfectly recalled or if outcomes and investments are jointly determined by unobserved confounders.

To deal with these problems, numerous studies have examined the effect of so-called contextual variation, i.e. events that are exogenous from the point of view of the individual child, on health and socio-economic outcomes later in life. These studies mostly compare outcomes of individuals who experienced a certain exogenous shock during childhood to outcomes of other individuals. Contextual variations that have been analyzed in this literature range from exposure to radioactive fallout, terrorist attacks, flue pandemics, famines and changes in the supply of fast food restaurants to weather shocks (Almond, Edlund, and Palme, 2009; Camacho, 2008; Almond, 2006; Currie, DellaVigna, Moretti, and Pathania, 2010; Maccini and Yang, 2009). The rationale for using such contextual variation is to determine age periods during which educational and health investments are particularly important.

However, having experienced a contextual change early in life does not necessarily imply that the human capital investment process is affected and vice versa.

Therefore, estimates from contextual variation only inform us about *population-wide* effects. Most times however, the estimate of interest is the *average causal* effect of certain childhood investments (the average causal effect on the individual). The difference between the population-wide, or Intention-To-Treat (ITT) effect, and the average causal effect on the individual is well-known and in general it is possible to use the contextual variation as an Instrument Variable (IV) for the treatment of interest, in order to obtain the average causal effect. Data on outcomes, linked to important contextual variables, usually contain detailed location information, allowing researchers to make statements about changes in a child's living environment. Such data, however, often lack information about individual investments. In other cases, information about human capital investments is available, but subject to recall bias if experienced at very young ages.

This article, summarizes how instrumental variable estimation can be used in a Local Average Treatment Effects (LATE) framework to obtain causal estimates from contextual variation and information about individual suffering for the case where both variables are binary.¹ It then discusses how average causal effects can be computed in the presence of imperfect recall or if information about compliance is entirely missing and has to be retrieved from another data source.

Contextual Effects

Exposure to contextual variation around birth or during childhood allows researchers to estimate so-called reduced form or ITT effects.² Take the example of fast food restaurants from Currie, DellaVigna, Moretti, and Pathania (2010). If we are interested in the effect of unhealthy food consumption on child obesity, it might not be possible to conduct an experiment, where French fries and ham-

¹See also the seminal paper by Angrist and Pischke (2008) and the exposition in Imbens and Angrist (1994).

²In the language of clinical trials, the ITT effect compares individuals who were *offered* a treatment to those who were not.

burgers are randomly allocated to some families with small children, but not to others. It might however be possible to make use of some random event at the macro level, such as the opening of a fast food restaurant nearby, which lowers the monetary and non-monetary costs of accessing unhealthy food. In this case the nearby opening of a fast food restaurant is the contextual variation ($Z \in \{0, 1\}$), which randomly affects the actual human capital investment variable (eating more than a threshold amount of unhealthy food), denoted by $D \in \{0, 1\}$, after conditioning on some covariates X . This can be helpful, if eating unhealthy food (D) and child obesity (Y) jointly depend on unobserved confounders, such as parenting practices (see figure 1 for a causal graph in the spirit of Pearl, 2000). If Z is randomly assigned and if it moves all individuals ‘in the same direction’, that is, easier access to unhealthy food does not induce anyone to reduce its consumption (in IV-talk this means that there are no defiers), the ITT effect can be obtained by looking at the difference in average outcomes between individuals who were exposed and those who were unexposed to the contextual variation Z :

$$E[Y|X = x, Z = 1] - E[Y|X = x, Z = 0] \quad (1)$$

This estimate has a causal interpretation: it will tell us the causal effect of the *offer of* treatment, i.e. the effect of living close to a place where a fast food restaurant opens its doors.

While the above effect is easily estimated and causal, two problems remain. First, if we are interested in the average causal effect of unhealthy food consumption, information about low-cost access to unhealthy food might not be very informative. In the example above, not all families who live near a fast food restaurant will actually visit the restaurant and some families who live very far away from a fast food restaurant might eat lots of unhealthy food. Similarly, living close to a fast food restaurant might simply lead to substitution away from

unhealthy food prepared at home, but might not change the overall amount of unhealthy food consumed (Currie, DellaVigna, Moretti, and Pathania, 2010). In other words, while proximity to a fast food restaurant reduces the costs of accessing unhealthy food, it will only induce some families to change their eating habits. The observed difference in child obesity rates between families who live close to a fast food restaurant and others will then only provide a qualitative assessment of the average causal effect of eating unhealthy food for the individual child.³ Second, selection into the treatment of interest, i.e. eating unhealthy food, is likely to be correlated with unobserved characteristics and therefore endogenous. If information about actual treatment take-up, that is actual fast food consumption, is available, the contextual variation can, however, serve as an instrument for the treatment of interest.

Instrumental Variable Estimation

Because not everybody changes their behavior in response to the contextual variation, the ITT effect is too small relative to the average causal effect. The effect on the individual can, however, be obtained by dividing the ITT effect by the difference in compliance (or take-up) rates between treatment and control groups. Under the classical IV assumptions, this gives the LATE, i.e. the average difference in potential outcomes for those individuals whose human capital investments were affected by the instrument (the compliers):

$$E[Y^1 - Y^0 | X = x, D_{Z=1} > D_{Z=0}] = \frac{E[Y | X = x, Z = 1] - E[Y | X = x, Z = 0]}{E[D | X = x, Z = 1] - E[D | X = x, Z = 0]} \quad (2)$$

³Note that the effect of easy access to fast food (in the form of a fast food restaurant in the neighborhood) might itself be a policy relevant effect.

The assumptions under which this holds are that Z does not have a direct effect on Y , other than through its (nonzero) effect on D (a fast food restaurant in the area influences child obesity only because it increases the consumption of unhealthy food); that there are no defiers in the sample; that the probability of compliance is not affected by the realization of Z (individuals in neighborhoods where a fast food restaurant opens are equally likely to increase their consumption of fast food as individuals in other neighborhoods would be were there a fast food restaurant opening in their neighborhood); and that the support of the covariates in the subpopulations of individuals affected and unaffected by the contextual variation is the same.

It is important to note the limitations of estimating a local effect. Unless the (hypothetical) difference in potential outcomes $Y^1 - Y^0$ is the same in the entire population, this implies that the estimated effect only holds for those individuals whose human capital investment process is affected by the contextual change. Hence agents induced to treatment by a given Z (opening of a fast food restaurant) need not be the same agents induced to treatment by an unrelated policy change (such as allowing schools to serve fast food for lunch). Therefore, researchers need to ask about the external validity and policy relevance of the LATE they estimate, unless the instrument-induced effect of treatment is the policy-relevant effect in question.⁴

Equation 2 gives the LATE for every value of X . Frölich (2007) shows that the average effect for the subpopulation of compliers can be obtained from:

$$E[Y^1 - Y^0 | D_{Z=1} > D_{Z=0}] = \frac{\int E[Y|X = x, Z = 1] - E[Y|X = x, Z = 0] f(x) dx}{\int E[D|X = x, Z = 1] - E[D|X = x, Z = 0] f(x) dx}. \quad (3)$$

Looking at equations 2 and 3, it becomes obvious that the LATE is simply the ITT effect (equation 1) scaled by the rate of compliance. The magnitude by which the average causal effect exceeds the reduced-form population effect hinges on the

⁴Recent research shows that a linear marginal treatment effects model (from which information about treatment effects for different parts of the population can be inferred) can be identified even with a single binary instrument (Mogstad, Brinch, and Wiswall, 2014).

percentage of individuals whose human capital investment process is affected by the contextual variation. Only at the extreme, if compliance to the contextual variation is perfect, the reduced-form effect conforms to the individual effect.

Missing Compliance Information or Imperfect Recall

To obtain the average causal effect from the ITT (numerator of equation 2), researchers thus need information about take-up of the treatment of interest (denominator of equation 2). In the above example, this would imply information about actual fast food consumption per child. The most straightforward way of obtaining such information is by conducting a survey which asks children about their fast food consumption. However, obtaining such compliance information is not always easy. Often, quite some time can pass between the contextual change and the compliance survey. This can, for instance, be the case when the contextual variation is a famine or a flue pandemic during childhood and the outcome of interest is adult health. In such cases, the parents of the children might long be dead and recall of the surviving children might depend on the age period they are asked about. In general, recall of a any investment may be difficult if this period took place before age 4.

Take a second example where the contextual variation (Z) is a famine and where undernutrition is the treatment (D). For this case, van den Berg, Pinger, and Schoch (2014) find that children aged 6-16 during a famine are more than twice as likely to report hunger than younger children (see figure 2).

To deal with such a situation, data combination can be useful. Equation 3 corresponds to the ratio of two estimators, which implies that the numerator and the denominator can be estimated from two different samples. Thus, if children before age 4 do not remember the investment, a compliance estimate from older children can be obtained, as long as $E[D|X = x, Z = 1] - E[D|X = x, Z = 0]$ is

the same in both samples.⁵ This assumption is not innocuous. Yet, in many cases it is possible to induce whether the denominator is larger or smaller in the substitute sample when compared to the original sample. If it is larger, then the two-sample estimate provides a lower bound for the average causal effect.

Conclusion

Research on the formation of human capital is difficult, because non-experimental measures of childhood investments are almost never free of confounding. Therefore, many recent papers have relied on contextual changes as quasi-experimental settings, which allow researchers to compare outcomes of individuals who did and who did not experience an exogenous shock during childhood. However, those comparisons only provide the so-called ITT. To obtain the average causal effect, they have to be scaled upwards unless compliance is perfect. Two-sample IV estimators make the scaling possible, even if the information about compliance stems from a different data source than the ITT estimate.

References

- ALMOND, D. (2006): “Is the 1918 Influenza pandemic over? Long-term effects of in utero Influenza exposure in the post-1940 US population,” *Journal of Political Economy*, 114(4), 672–712.
- ALMOND, D., L. EDLUND, AND M. PALME (2009): “Chernobyl’s subclinical legacy: prenatal exposure to radioactive fallout and school outcomes in Sweden,” *Quarterly Journal of Economics*, 124(4).
- ANGRIST, J. D., AND A. B. KRUEGER (1992): “The effect of age at school entry on educational attainment: an application of instrumental variables with moments from two samples,” *Journal of the American Statistical Association*, 87(418), 328–336.

⁵This can be done by using either a non-parametric Wald estimator or the parametric two-stage least squares estimator (2S2SLS). The 2S2SLS-estimator has been developed by Angrist and Krueger (1992) and Arellano and Meghir (1992); Inoue and Solon (2010) adjust the estimator for use in small samples.

- ANGRIST, J. D., AND J.-S. PISCHKE (2008): *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.
- ARELLANO, M., AND C. MEGHIR (1992): "Female labour supply and on-the-job search: an empirical model estimated using complementary data sets," *The Review of Economic Studies*, 59(3), 537–559.
- CAMACHO, A. (2008): "Stress and birth weight: evidence from terrorist attacks," *The American Economic Review*, pp. 511–515.
- CURRIE, J., S. DELLAVIGNA, E. MORETTI, AND V. PATHANIA (2010): "The Effect of Fast Food Restaurants on Obesity and Weight Gain," *American Economic Journal: Economic Policy*, 2(3), 32–63.
- FRÖLICH, M. (2007): "Nonparametric IV estimation of local average treatment effects with covariates," *Journal of Econometrics*, 139(1), 35–75.
- IMBENS, G. W., AND J. D. ANGRIST (1994): "Identification and estimation of local average treatment effects," *Econometrica: Journal of the Econometric Society*, pp. 467–475.
- INOUE, A., AND G. SOLON (2010): "Two-sample instrumental variables estimators," *The Review of Economics and Statistics*, 92(3), 557–561.
- LINDEBOOM, M., F. PORTRAIT, AND G. J. VAN DEN BERG (2010): "Long-run effects on longevity of a nutritional shock early in life: the Dutch Potato famine of 1846–1847," *Journal of Health Economics*, 29(5), 617–629.
- MACCINI, S., AND D. YANG (2009): "Under the weather: Health, schooling, and economic consequences of early-life rainfall," *The American Economic Review*, pp. 1006–1026.
- MOGSTAD, M., C. BRINCH, AND M. WISWALL (2014): "Beyond LATE with a discrete instrument: Heterogeneity in the quantity-quality interaction of children," Discussion paper, Working Paper.
- PEARL, J. (2000): *Causality: models, reasoning and inference*, vol. 29. Cambridge University Press.
- VAN DEN BERG, G. J., P. PINGER, AND J. SCHOCH (2014): "Instrumental Variable Estimation of the Causal Effect of Hunger Early in Life on Health Later in Life," *The Economic Journal*, forthcoming.

Appendix

Figure 1: Relationship between contextual variable (Z), treatment (D), outcome (Y) and unobserved characteristics (V, U)

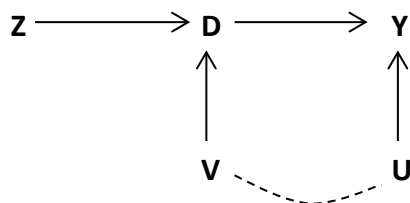


Figure 2: Probability to report hunger conditional on famine experience at respective age (reprinted from van den Berg, Pinger, and Schoch, 2014)

