# Essays in Applied Microeconomics

Inaugural-Dissertation

to obtain the degree of

Doktor der Wirtschaftswissenschaften (Dr. rer pol.)

submitted to the Faculty of Business Administration and Economics at the

Heinrich Heine University Düsseldorf

Presented by

**Simon Christian Schulten**

First Supervisor: Prof. Paul Heidhues, Ph.D.

Second Supervisor: Prof. Dr. Joel Stiebale

August 2023

*To Anna.*

# Contents

# Introduction

How do firms react to behavioral biases of their opponents, be it the behavioral biases of their customers, managers or rank and file employees? In this dissertation I study this question in three different contexts.

In Chapter 1, which is joint work with Paul Schäfer, we study consumer who makes mistakes when they repeatedly make purchases from the same exploitative firm. It is natural to think, that eventually consumers will learn to avoid their mistakes, which limits the firm's exploitation. A profit maximizing firm, however, has an incentive to undermine such learning. We study these learning dynamics and the firm's response in a multi-unit descending price auction with a simultaneous fixed price offer. In our panel of roughly 8 million bids by 280.000 bidders in 70.000 auctions, spread over two years, consumers often bid more than the simultaneously available fixed price. Depending on competing bids, an overbid may lead to paying more than the fixed price, which we call overpaying. Crucially, not every overbid leads to overpaying. We argue overpaying increases the saliency of the consumers' mistake by making it payoff relevant, which increases the likelihood that the consumer learns to avoid their mistake. Indeed, bidders who overpay subsequently overbid less often and are more likely to leave the market compared to bidders who similarly overbid, but did not overpay.

Having quantified the consumer learning dynamics, we turn attention to the firm incentives that derive from consumer learning. We show that the loss in future profits due to consumer learning makes overpaying undesirable and document a structural break in our data at which the firm eliminates overpayments – and the resulting consumer learning – through changes in how it runs its auctions. Specifically, we discuss how the firm most likely uses dynamic quantity increases *during* the auction to shut down overpaying. Interestingly, the firm has the ability to target only those auctions at risk of overpaying due to excess demand. We also discuss how the firm increases fixed prices only once a new CEO takes office after the end of our sample. Increasing the fixed prices rules out overbidding mechanically. Methodologically, we discuss identification of our treatment effects using causal graphs and show how these treatment effects identify a three type structural model of bidder behavior with learning dynamics.

In Chapter 2, which is joint work with David Zeimentz and Dennis Gottschlich, we study the effect of managerial overconfidence on investment cash-flow sensitivity, innovation, and CEO compensation using data from France, Germany, and the UK. Using self-collected stock trades and option exercises of C-suite directors, we revisit the canonical overconfidence classifications and discuss the portability of the approach from the US to Europe. Exploiting the fact that we observe managers at different stages of their professional life, we propose a novel classification to disentangle optimism and the overestimation of own managerial ability. We find similar effects of overconfidence in Europe. In particular, overconfident managers invest roughly 10 cents more when they have an additional Euro in cash flow and pursue more innovation as evidenced by roughly 30% more citation-weighted patents. With regard to overconfidence and compensation we find that overconfident managers receive smaller stock grants and smaller salary in Europe, whereas the US literature finds negative effects on options, bonus and total compensation. On the methodological side, we discuss the identification strategies of the existing literature using causal graphs.

In Chapter 3, which is single authored, I consider the puzzle that firms routinely remunerate employees with variable pay instruments like stock options or profit sharing. Standard principle agent models explain such contract features in settings where a CEO can influence the stock price or the probability of a good outcome through effort. It is, however, implausible to assume a single rank and file employee can move the stock price or profits. I explain the puzzling use of stock options for rank and file employees through a model where an an agent with endogenous optimism may acquire company-specific skills. The wage consists of a salary, stock options, and a bonus that is increasing in company-specific skill. Crucially, I assume the probability of the good state to be independent of the employees actions. Granting the optimal expectations agent stock options induces optimistic beliefs, which leads to the agent to accumulate a higher skill level to the benefit of the firm.

The common thread among the chapters of this dissertation is the aspect of the exploitative firm. In the context of compensation in Chapter 2 and 3 exploitation of the overconfidence and optimism bias is front and center. The realistic firm tailors the contract to the overconfidence or optimism of the agent for her own advantage. In Chapter 1 the explicit contracting context is missing, but the firm exploits overbids by customers to extract additional overpaying revenue. The situation is, however, more nuanced, as we present evidence that the initial exploitation is suboptimal due to consumer learning. In particular, consumer learn not to purchase with this exploitative firm anymore. Since the firms business model relies heavily on repeat purchases this is an effective deterrence to the exploitation the firm engaged in.

# 1 Managing Bidder Learning in Retail Auctions

Paul Schäfer, Simon Schulten

## 1.1 Introduction

Theory and evidence suggests that firms price strategically to exploit consumer biases (see, for example Della Vigna and Malmendier, 2006; Grubb, 2015; Grubb and Osborne, 2015; Heidhues and Kőszegi, 2018; Malmendier and Szeidl, 2020). Despite such evidence, a common critique is that consumers will learn to avoid being exploited. This argument, however, ignores the firm's incentives to inhibit consumer learning. In this paper, we empirically study consumer learning in retail auctions and document how a firm improves at designing the sales environment to impede consumer learning.

Despite its potential importance, we know little about whether and how firms are able to manage consumer learning. A measurement challenge arises when a firm takes actions to prevent consumer learning. While we can in principle observe the firm's actions, there is often little variation in those actions. Additionally, we cannot easily infer the prevented consumer learning from firm's actions alone. When a firm improves at managing consumer learning over time, we can overcome the measurement challenge. In such a case, researchers can collect data on consumer learning and on the firm's response.

In our case, the firm operates a televised multi-unit descending auction with uniform pricing and an online shop, where goods are sold at a fixed price. The auction starts at a high price that is lowered in increments over time. Bidders bid at the current price and the auction ends when all units are claimed. According to the uniform pricing rule, every winning bidder pays the lowest successful bid. Following the empirical literature, we call a bid that is higher than the fixed price an *overbid*.[1] When the auction price is higher than the fixed price, we call the auction *overpaid*. Crucially, overbidding does not imply overpaying, as overpaying requires that all bids in an auction are overbids (the lowest bid is higher than the fixed price).

We collect a bidder-level panel of the firm's multi-unit descending auction. The data spans over two years, detailing more than 8 million bids submitted by 280.000 bidders in approx. 70.000 auctions. The bidder level panel structure allows us to analyze consumer learning and the firm's reactions to it. In line with lab and field evidence many customers, at the beginning of our sample, overbid (e.g. Kagel and Levin, 2011; Malmendier and Lee, 2011; Gesche, 2022; Ocker, 2018).

In our auction format, overbidding is a necessary, but not a sufficient condition for overpaying. An auction ends in overpayment if and only if all bids are overbids. We exploit this fact to

[1] For a descriptive analysis of overbidding in auctions run by the same firm see Ocker (2018).

construct a natural treatment and control group design: treated are those bidders who overbid and — due to other consumers' bids - ended up overpaying, whereas bidders who similarly overbid but did not overpay end up in the control group.

Overpaying makes the consumers' mistake (overbidding) payoff relevant and arguably more salient. A salience effect of overpaying on future behavior compared to consumers who overbid but did not overpay is then naturally interpreted as evidence of consumer learning. We argue that theories explaining overbidding as optimal behavior are inconsistent with the observed behavioral changes following overpayment in our data. We find overpaying leads consumers to spend less in the future. More explicitly, it leads consumers to hand in fewer overbids and fewer bids overall, as well as increases the likelihood of refraining from bidding altogether.

We use a DAG as a convenient and precise way to codify causal knowledge about the data generating process. We derive our DAG from institutional knowledge about the way the firm plans and runs the auction, the auction rules (uniform pricing), and natural assumptions about bidder behavior. In addition, we discuss how such a DAG can be derived from, perhaps more familiar, structural equation models.[2]

Our analysis demonstrates a novel way to combine a traditional economic model, a directed acyclic graph (DAG) (Pearl, 2009; Imbens, 2020; Hünermund and Bareinboim, 2023), and the sufficient statistics approach (Chetty, 2009). We develop a three-type model of initial overbidders who may learn to become non-overbidders or dropouts. An overbidder who becomes a non-overbidder simply truncates her bidding function at the fixed price, so that she does not repeat her mistake. An overbidder who becomes a dropout ceases to participate in future auctions altogether. Learning not to overbid is a behavioral adjustment at the intensive margin, whereas dropouts represent the extensive margin. In our three-type model, conditional on bidding, learning not to overbid drives the observed behavioral change of bidders handing in fewer overbids. Dropouts, however, drive the negative effect on both overbids and non-overbids (fewer bids overall). We disentangle the two margins and find that an initial overbidder, who overpays, has a 4.2% chance of dropping out and a 7.2% chance of becoming a non-overbidder.

In the presence of consumer learning the firm faces a trade-off between extra overpaying revenues today and foregone revenue tomorrow. Back-of-the-envelope calculations demonstrate that the extraction of overpaying revenue is suboptimal in the beginning of our sample. In the second half of our sample we observe a structural break in the time series of overbidding and overpaying. Before the break, roughly 17% of all bids are overbids and 4% of auctions end in overpayment. After the break, overbidding prevalence is substantially reduced and practically no auction ends in overpayment. The structural break is accompanied by a small jump and a reversal in the trend of items sold per week, albeit statistically insignificant. Fixed prices remain unchanged at the break-point.

[2]Assigning casual meaning to a structural equation model turns it into a structural causal model.

Management explains that they induced the structural break in overpaying with a relatively minor, but targeted quantity increase. Auctions are simultaneously run by an on-screen auctioneer and a director, who is off-screen and has access to real-time demand (e.g., number of people watching, revenue by the minute). The director uses this information to "steer" the auctions and may, if necessary, increase the quantity *during* the auction.[3] Thus, the director is able to target quantity increases to auctions that are at risk of ending in overpayment, thus shutting down overpayment. Management also tells us of a new pricing policy implemented just after the end of our sample. This new pricing policy raises fixed prices and auctions routinely undercut these higher fixed prices with the starting bid, ruling out overbidding mechanically. Raising fixed prices (presumably) comes at the cost of lower online shop sales, so after an auction ends, fixed prices are lowered in a 24-hour sales discount to a small increment above the auction price. This gives customers who missed the auction the chance to purchase at a comparable price, rather than the very high regular fixed price.[4]

By documenting that our firm does not take into account consumer learning in the beginning of our sample, we contribute to the literature on non-optimizing firms (see, for example Cho and Rust, 2010; DellaVigna and Gentzkow, 2019; Hortaçsu and Puller, 2008; Hortaçsu et al., 2019). In contrast to these papers, our firm is adaptive as it gradually improves at managing consumer learning, by increasing its scope of maximization (Doraszelski et al., 2018). First, the firm optimizes its quantity choice in a targeted way, but leaves fixed prices unchanged. Later, it optimizes over quantities and fixed prices jointly. Interestingly, the last change was only implemented by a new CEO, in line with the idea of learning through noticing (Hanna et al., 2014).

Our paper is related to the empirical literature that studies how consumers learn when they trigger a fee. This literature documents that (some) consumers learn to avoid triggering fees, but this learning effect depreciates over time (Haselhuhn et al., 2012; Agarwal et al., 2013). Ater and Landsman (2013) find consumers who switch their banking plans after paying an overage fee are more likely to switch to plans with larger allowances. In contrast to consumers who switch plans but did not pay overage fees, fee-switchers *increase* rather than decrease their monthly payments, suggesting that consumer learning can be detrimental to consumers.

Our finding that consumers give up on bidding rather than simply adjust their bid function to avoid overbidding also indicates a complex, and possibly non-rational, response to feedback. A transaction that leaves the bidders worse off than the reference point (i.e., overpaying) results in a negative transaction utility (Thaler, 1999) and may reduce future market participation through several channels: updated beliefs about the utility from market participation (Backus

[3] In its terms and conditions the firm reserves the right to increase quantity after the auction has started. Quantity decreases are not mentioned in the terms and conditions, but there is an option to fully cancel an auction when demand is so low that the price would have to approach 0 for the auction to end.

[4] Another interpretation of setting a rebate to the auction price is that the auction serves as a mechanism to discover a "reasonable" fixed price.

et al., 2021), updated beliefs about their abilities (Seru et al., 2010) and antagonizing consumes (Anderson and Simester, 2010). Gesche (2022) documents customer complaints in the eBay feedback system when they pay more at auction than the fixed price offered by the same seller. In our case, management confirms anecdotally that some consumers who overpay call the hotline to complain.

To the best of our knowledge, we are first to illustrate how firms actively suppress or at least slow down consumer learning. If consumers bid less than their value of the good, trade is efficient and consumer attrition decreases consumer welfare. Under this assumption, the firm's quantity targeting policy increases consumer welfare by minimizing attrition compared to the initial exploitation of overpaying. On the other hand, if consumers bid more than their value, they may pay an auction price above their value, particularly in overpaid auctions (for a survey of above value bidding see Kagel and Levin, 2011, Section 1.2). Consumer attrition may then increase consumer welfare. In this case, shutting down consumer learning through targeted quantity increases may harm consumer welfare. In any case, targeted quantity increases are preferred in terms of consumer welfare over the pricing policy of increased fixed prices. The firm's quantity targeting decreases the auction price - at least in auctions that would have ended in overpayment otherwise.

Our dynamic considerations complement the existing literature.[5] In a static analysis, relatively few behavioral buyers suffice to generate a price impact in an auction compared to a fixed price market (Malmendier and Szeidl, 2020). Considering a dynamic setting, however, adds downsides to the exploitation incentive. In our application, exploiting overpaying causes some consumers to drop out of the market, so the firms loses those consumers' lifetime value. The firm faces a trade-off between revenue maximization in a single auction (in a static sense) and customer retention across auctions. Controlling the learning opportunities that the customer has alleviates this trade-off for the firm. In our data, the firm can remove the learning stimulus altogether by changing fixed prices, thus resolving the trade-off.

We demonstrate the need to disentangle customer attrition (extensive margin) from strategic learning on the platform (intensive margin) and provide an empirical approach to that end. Customer attrition has recently been studied in the context of eBay auctions with buy-it-now option (Backus et al., 2021) and in the context of ending a session of chess on a won or lost game (Avoyan et al., 2021). Both papers analyze psychological reasons for consumer attrition, although they do not study behavioral mistakes.

The paper proceeds as follows. In Section 1.2 we discuss the rules of the multi-unit descending auctions and further institutional details. In Section 1.3 we report on data collection. In Section 1.4 we describe our data including the empirical evidence on firm behavior. The model of firm

[5] Heidhues and Kőszegi (2018) discuss a lack of dynamic analysis in behavioral industrial organization, and call for more research on the topic.

incentives is laid out in Section 1.5. Section 1.6 discusses our empirical strategy. We report estimates of the bidders' learning response in Section 1.7. Section 1.8 concludes.

## 1.2 The Multi-Unit Auctions

The seller uses a televised multi-unit descending price auction embedded in hour-long shows to sell consumer goods. Each show consists of several auctions for similar products such as home textiles, men's watches, or jewelry. The average auction lasts about 11 minutes. The seller broadcasts auction shows 20 hours a day, via TV (bids submitted by phone) or online (websites, several apps). At any given time only one auction is held and bids are submitted into the same auction through different channels.

Bidders can also purchase every product up for auction through an online shop at a fixed price. The online shop is available on the same website and the apps that also broadcast the live stream of the auction shows. We therefore view the fixed price as the relevant point of comparison for bids and auction price.

The auction rules ensure that only people who bid above the fixed price (overbid) also pay above the fixed price (overpay). At the beginning of each auction, the auctioneer announces the (initial) number of items to be sold and the auction's starting bid[6]. This starting bid is then gradually lowered over time in discrete increments. Bidders can enter the auction at the current bid and claim one or more units of the good. The auction ends when all units are claimed. All bidders pay the lowest successful bid (uniform pricing rule). Because of this uniform pricing rule, an auction is only overpaid if all bids are overbids.

Shipping costs apply to the fixed price and the auction in the same way. For that reason, we ignore shipping costs in our analysis. Additionally, bidders who bid by phone have to pay a flat fee of one Euro. Since research on shipping costs suggests that this fee is likely (at least partially) ignored, we do not include this fee when we calculate overpaying (Hossain and Morgan, 2006). Furthermore, if customers actually internalise the phone fee, we erroneously assign some bidders to the control group and hence, our approach is conservative.

## 1.3 Data

We scrape data on bids and products from the seller's website from October 20, 2016, to January 3, 2019. Since data is removed from the website after some time, we run the scraping script in hourly intervals.[7]

First, we access the schedule in the TV programming section of the website. This schedule gives us information on the show level, such as time and date, product category, and the

[6] In their terms and conditions, the firm reserves the right to increase the number of units to be sold even after an auction has started.

[7] Due to a small coding error we did not collect auction shows at 6, 10 and 11 pm. Other than that, we observe all shows and within shows all auctions and bids that took place.

auctioneer running the show. Second, we collect auction-level data by going through the list of all planned auctions. This list contains an auction ID that we use to scrape bids and bidder nicknames from a separate part of the website. Third, we collect product information from the online shopping section of the website. Most importantly, we collect the fixed price of each product at the time of the auction.[8]

This data collection yields a bidder level panel of 8.48 million bids in more than 69000 auctions spanning over 2 years and 2 months. We use this raw data to calculate several variables, including the auction price (the minimum of the bids), bidder history variables that capture typical behavior and past experience on the auction platform, and dummies indicating whether a bid is an overbid or leads to an overpayment (overpaid).

Table 1 reports summary statistics broken down to the show category level. Naturally, product categories differ with respect to price and quantity. For example, there are more than 7 times as many items sold in the household category than in the watches category. Nevertheless, the two categories are responsible for similar shares of the firms revenue at 20.5% and 18.3% of total revenue made in the auctions. Unfortunately, we could not obtain purchased quantity for the fixed price sales channel, but discussions with the firms management revealed that revenue s overwhelmingly made in the auctions.

| Show Category | Items Sold | Auction Price | Fixed Price | Revenue Share |
|---|---|---|---|---|
| Beauty & Wellness | 2512517 | 10.646 | 15.863 | 0.170 |
| Leisure & Collecting | 20808 | 54.854 | 92.173 | 0.007 |
| Household | 2583827 | 12.497 | 17.625 | 0.205 |
| DIY & Gardening | 666292 | 13.534 | 21.569 | 0.057 |
| Home Textiles | 705398 | 11.917 | 16.718 | 0.053 |
| Fashion & Accessory | 987338 | 22.674 | 29.302 | 0.142 |
| Jewelry | 674742 | 42.865 | 56.709 | 0.183 |
| Watches | 338449 | 85.259 | 119.726 | 0.183 |

Table 1: Descriptives by Show Category

## 1.4 Break in Overbidding and Overpaying

We observe a structural break in the empirical overbidding and overpaying rates in our data. In Panel A of Figure 1 we plot the probability to overpay conditional on overbidding aggregated to weekly averages. Initially, the probability to overpay given one has overbid is roughly 23%, so overbidders are likely to pay for their mistake. Subsequently, we observe a sharp decline in the conditional probability to overpay given one has overbid from roughly 23% to essentially 0%. We determine the exact date of the structural break with a QLR test (Kleiber and Zeileis, 2008).[9] To illustrate, we add a linear trend on both sides of the structural break.

Table 2 reports summary statistics split by the structural break, since that is the defining

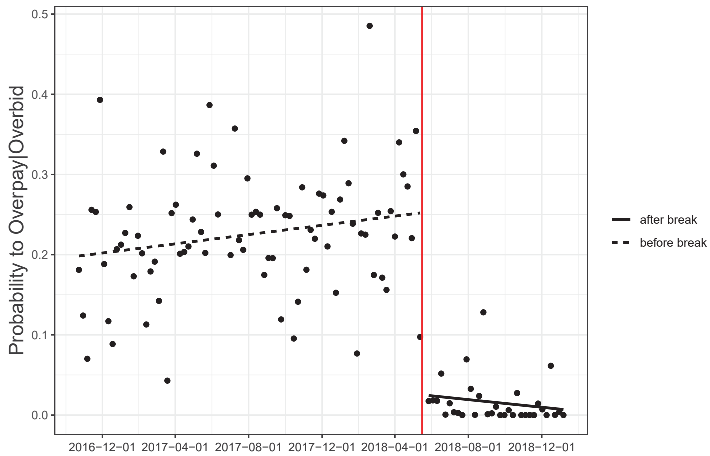
[8]We also collected product ratings, but those are quite sparse at this retailer, so we do not use them.

[9]The most probable break-point is the day with the highest individual test statistic, in our case, the 16th of May 2018. We plot the time series of test statistics in Figure 13 in Appendix A.8.

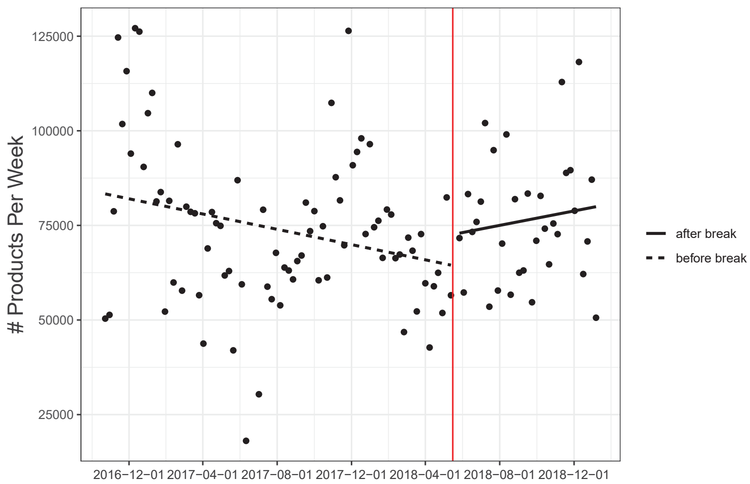
feature of our data. Before the break, 17% of all bids are overbids. While the overall probability of overpaying is small at 4%, the probability that this behavioral mistake becomes payoff-relevant is substantially higher. After the break, the probability of overbidding collapses to just below 10%, which lowers the probability of overpaying to essentially 0%. Together, these statistics indicate that before the break consumer learning is a lot more likely than after the break.

The sharp decline in overpaid auctions coincides with a discrete increase in the number of products sold in each weak (Panel B in Figure 1). The number of products primarily increased because the seller conducted more auctions. Fixed prices do not change at the structural break (Figure 15 and 14 in Appendix A.8).

Figure 1: Structural Break




(a) Probability to overpay given one has overbid (weekly averages).




(b) Number of Products sold each week.

### 1.4.1 Back-of-the-Envelope Calculation

If overpaying causes a demand response, the firm faces a trade-off between extracting overpaying revenue today and foregone revenue tomorrow. Naturally, extracting overpaying revenue comes at little to no operational cost, so for the purpose of our back-of-the-envelope calculation we

Table 2: Estimated probabilities of overpaying and overbidding.

| | before break (N = 4573854) | after break (N = 1960103) |
|---|---|---|
| **overbid** | | |
| probability | 0.17 | 0.093 |
| average amount | 3.6 | 4.4 |
| median ammount | 1.1 | 1.1 |
| **overpaid** | | |
| probability | 0.039 | 0.0014 |
| average amount | 2.5 | 2.9 |
| median amount | 1.1 | 0.6 |
| **overpaid\|overbid** | | |
| probability | 0.23 | 0.015 |
| **auctions** | | |
| average duration (minutes) | 11.25 | 11.9 |
| average product price | 27.8 | 28.2 |
| average auction price | 21.3 | 20.8 |

assume it is profit. Foregone revenue, however, does not equate foregone profit, so we need to adjust foregone revenues with the gross margin. Our back-of-the-envelope calculation suggests that extracting overpaying revenue is suboptimal if

$$\text{Foregone Revenue} \cdot \text{Gross Margin} > \text{Avg. Overpayment}.$$

In our data, overpaying revenue is small at 2.50€ on average (see Table 2). From the available balance sheets, we calculate gross margin at 0.256 and 0.268, in 2014 and 2015 respectively.[10] To be conservative, we use gross margin at 0.256. Thus, if bidders drop out for one year, the foregone revenue threshold that renders extraction of overpaying revenues suboptimal is 9.77€.

Naturally, foregone revenue may be driven by the intensive and the extensive margin. We calculate a *lower* bound by only considering revenue lost from customers dropping out of the customer base. To quantify revenue lost from customer loss, we assume the customer abstains from bidding for at least one year. Thus, multiplying average annual spending with the probability of losing a customer due to overpaying, $\epsilon$, gives the revenue loss due to overpaying attrition. In our data annual average spending is 360€ in 2017 and 383€ in 2018. This is in line with annual spending of over 300€ as claimed in investor presentations by the firm.[11]. Thus, extracting overpaying revenue is suboptimal if

$$\text{Annual Spending} \cdot \text{Gross Margin} \cdot \epsilon > \text{Avg. Overpayment}.$$

Plugging in the numbers yields an epsilon-threshold of $\epsilon > 0.027$. Note that we are using the

[10]The balance sheet data needed to calculate gross margins for later years are unavailable due to a change in reporting format.

[11]see https://www.1-2-3.tv/uploads/files/2013_06_123tv%20Company%20Profile.pdf https://www.1-2-3.tv/uploads/files/2012_10_%20123tv%20Das%20Unternehmen.pdf https://www.1-2-3.tv/uploads/files/PM_123tv_2014_07_01.pdf, accessed 12.01.2022

conservative numbers for gross margin (0.256) and annual spending (360€) to arrive at this number.

We emphasize that the back-of-the-envelope calculations are conservative for two more reasons. First, using the annual spending number assumes a lost customer would have stopped purchasing after one year. Given the high rate of repeat purchases, however, it is unlikely that customers only stay one year. Second, assigning the full overpaying revenue to profit may be problematic. In our conversation with the firms management it was hinted that return rates may be higher for overpaid items: Customers can simply return their purchase at the overpaid auction price and - if they so choose - repurchase the item at the fixed price. This would undo the overpaying revenue for the seller and, even worse, impose costs on the seller who has to deal with the returned item. Unfortunately, we are neither able to calculate counterfactual auction prices, nor do we have data on average customer lifetime or return rates.[12]

## 1.5 Firm Incentives

The televised auctions are run jointly by an auctioneer, who is on-screen, and a director, who is off-screen. The director has access to rich data such as the number of viewers and revenue broken down by the minute. The director has the authority to "steer" the auctions progress in a number of ways, but crucially he may increase the quantity even after the auction has started.[13] Management is also aware that overpaying may lead to unhappy customers, since those customers sometimes call the hotline to complain. We take this information seriously and use them in our model to capture the trade-off between overpaying today and customer retention.

We assume every bidder has a latent bid, which is the result of some mapping from values to latent bids. Whether a bidder hands in her latent bid depends on her behavioral type. We assume three types of bidders: overbidders, non-overbidders, and dropouts. An overbidder simply submits his latent bid. A non-overbidder, however, never submits an overbid by truncating the bidding function at the fixed price. A dropout, as the name suggests, does not participate in the auction.

Consider an initial overbidder, whose overbid leads to overpayment and thus makes the mistake salient. In our model, learning happens through type changes. An initial overbidder may learn not to repeat her mistake by becoming a non-overbidder. We denote the probability of learning at this intensive margin by $\iota$. Besides learning not to overbid, we also allow for learning at the extensive margin. There are multiple reasons why a bidder may cease to participate in the auctions. These include updated beliefs about the utility from market participation (Backus

[12]We note, however, that the fixed price is lagging the auction in terms of revenue: management told us that they make less than 10% of the revenue through the fixed price offering.

[13]Typical director tasks, for example, include sequencing of the camera feeds or when to show certain TV overlays.

et al., 2021), updated beliefs about ability (Seru et al., 2010)[14] and consumer antagonism (Anderson and Simester, 2010; Gesche, 2022). We denote the probability of learning at the extensive margin by $\epsilon$

Figure 2 visualises how latent bids, $\beta_{i,t}$, map into actual bids. Even though our analysis does not rely on functional form assumptions, for ease of illustration we suppose that latent bids are uniformly distributed in Figure 2. An overbidder simply hands in her latent bid, which is indicated by the 45° line. Thus, the distribution of actual bids submitted by overbidders is also the uniform distribution. A non-overbidder, on the other hand, never hands in a latent bid larger than the fixed price $p_t$. Instead, we assume that non-overbidders exactly bid the fixed price in the auction, which is why the bid distribution has a mass point at $p_t$.

We model intensive margin learning narrowly, in the sense that bidders who experience the consequences of their mistake avoid said mistake in the future. The literature provides evidence for such a learning dynamic, both in controlled lab experiments on auctions as well as in field settings other than auctions. In the lab, bidders adjust their bid in the direction that would have been better in the past (Neugebauer and Selten, 2006). In the field, customers who pay a fee avoid the action that triggered the fee (Haselhuhn et al., 2012; Agarwal et al., 2013; Ater and Landsman, 2013).

Suppose there are $N = o_t + s_t$, bidders in the auction at time $t$, of which $o_t$ are overbidders and $s_t$ are non-overbidders. All bidders have unit demand and the same latent bid $\beta > p$ for simplicity. Then, the auction price is a function of the number of overbidders and non-overbidders and the auction quantity.

We present a simplified version of our empirical model in Section 1.6.1 in Appendix A.1 to illustrate the seller's incentives arising from changes in type from overbidder $o_t$ to non-overbidder $s_t$ or dropouts. We simplify the analysis by assuming there are $N = o_t + s_t$ bidders in the auction at time $t$ and that all bidders have unit demand and the same latent bid $\beta$.[15] To make the case interesting, we assume that the latent bid is larger than the fixed price $\beta > p$. Then, the auction price is a function of the number of overbidders and non-overbidders and the auction quantity.

$$p_a(q_t, o_t, s_t) = \begin{cases} \beta & q_t \leq o_t \\ p & o_t < q_t \leq o_t + s_t \\ 0 & o_t + s_t < q_t \end{cases}$$

Choosing a sufficiently small quantity, $q_t < o_t$, will ensure that all bids in the auction

[14]In our context, bidders may think they are irredeemably bad or unlucky at bidding, so they should stop bidding altogether.

[15]We make the simplifying assumptions here to clearly state the main point, but we drop them in our empirical analysis.

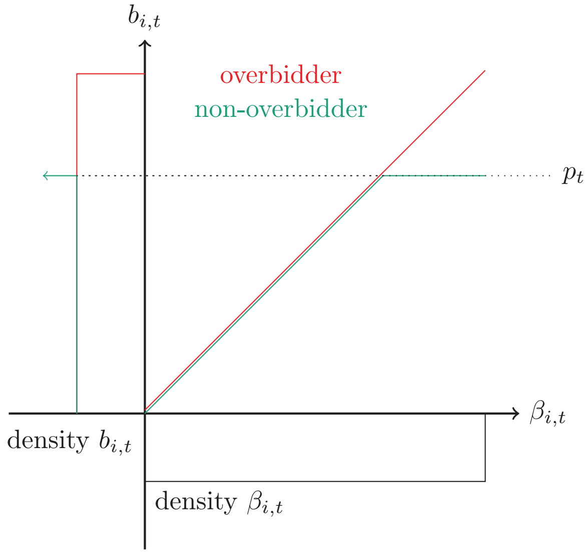



Figure 2: Bids as a function of latent bids. Marginal distribution of bids and latent bids for uniformly distributed latent bids.

are overbids and thus the auction ends in overpayment at $p_a = \beta > p$. A larger quantity, $o_t < q_t \leq o_t + s_t$ will ensure that non-overbidders also bid in the auction and thus the auction price will realize exactly at the fixed price $p_a = p$. Finally, setting quantity larger than the number of participants leads to an auction price of 0, which is never optimal.

Appendix A.1 presents this model in more detail. Here we restrict ourselves to discussing the cases where the firm views fixed prices as exogenously given and chooses quantity and when both variables are chosen simultaneously. This approach is motivated by the fact that fixed prices where only changed after a new CEO took office, whereas they remained unchanged at the structural break in our data. This is in line with a model of learning through noticing where a decision maker may not fully optimize because he fails to notice an important feature of the optimization problem (Hanna et al., 2014).

**Policy 1: Choosing Quantities** As discussed above, the director may increase quantity during the auction, while taking fixed prices as given. Offering extra units in the auction leads to a downward move along the demand curve in each auction, which lowers prices.[16] Overpaying may also be reduced without increasing overall quantity, by shifting quantity from an non-overpaid auction to an overpaid auction, where both auctions sell the same good.

We find that in the case of exogenous fixed prices, the seller's optimal quantity choice depends on how large overpaying $\beta - p$ is compared to the discounted revenue lost due to extensive margin learning $\epsilon$ as given by Observation 1.1.[17]

**Observation 1.1.** *Suppose the seller can only choose quantity $q$ and the fixed price $p$ is*

[16] While we observe the number of units sold in each auction, we do not observe the number of units that were originally planned for the auction. Unfortunately, this means we do not know which auction increased supply dynamically during the auction.

[17] The proof is in Appendix A.2.

*exogenously given. If* $\beta - p \geq \frac{\delta}{1-\delta} \cdot p \cdot \epsilon$ *the profit maximizing quantity is* $q_t = o_t, \ \forall t$. *If* $\beta - p < \frac{\delta}{1-\delta} \cdot p \cdot \epsilon$ *the profit maximizing quantity is any* $q_t \in (o_t, o_t + s_t], \ \forall t$.

If overpaying is larger than the revenue loss due to extensive margin learning, then the seller chooses a small quantity and the auctions end in overpaying, $p_a = \beta > p$. If the reverse is true, the seller prefers not to extract the extra revenue through overpaying to preserve future demand and the auctions end at the fixed price. In this setting with exogenous fixed prices, the sellers can only shut down learning by lowering the auction price and thus foregoing the extra revenue. Note that the optimal decision does not depend on intensive margin learning as taking away the opportunity to exploit overbidding tomorrow is not an effective deterrent against exploitation today.[18]

**Policy 2: Choosing Fixed Prices and Quantities** Since the seller makes most revenue in the auctions, the fixed-price outside option primarily acts as a reference price. Consequently, the seller can use adjustments of this reference price as a second instrument to avoid overpaying.

After our sample ends, the seller raised fixed prices. Initially, the fixed price is set (presumably very) high and the auction high bid always undercuts the fixed price. After an auction ends the fixed price is lowered to the auction price plus a small increment for 24 hours. This strategy gives potential customers who missed the auction the chance to purchase the good at a price below the recommended retail price. For example, an item may be offered at its recommended retail price of €30, while the auction starts at €20. The auction price may realize at €12, and the fixed price falls to €15 for 24 hours.[19] Since we did not collect data after the policy change, we cannot check if the fixed prices rose on average. By construction, the new policy, however, makes overbidding and overpaying impossible.

**Observation 1.2.** *If sellers can choose* $p$ *and* $q$*, they maximize their profits by setting* $p > \beta$ *and* $q_t \in (o_t, o_t + s_t], \ \forall t$.

Observation 1.2 states that the firm can circumvent revenue losses due to learning, by setting high fixed prices. Instead of addressing the cause of consumer learning (high auction prices) the firm can remove the stimulus (overpaid auctions) by adjusting the fixed price. This policy increases prices without changing quantity and, thus, surplus is redistributed to the firm.

## 1.6 Empirical Strategy

We take a sufficient statistics approach to quantify the extensive and intensive margin learning rates. We present a structural equation model of how consumers bid at auction and how the firm runs the auction. We derive treatment effects that we can estimate from our data. We explain

[18] Observation 1.1 closely resembles our back-of-the-envelope calculation in Section 1.4.1

[19] The short-term rebate on the fixed price may be seen as the price discovery aspect of an auction.

how to represent our structural equation model as a Directed Acyclic Graph (DAG)[20]. We use the DAG representation of our model to show that our treatment effects are identified. The DAG representation allows us to derive the conditional independence assumption and the control variables required for identification from our empirical model using Pearl (2009)'s back-door criterion. We discuss interpretation of our treatment effects and the assumptions needed to recover extensive and intensive margin learning rates, $\epsilon$ and $\iota$. The strength of this approach is that the treatment effects are valid under weaker assumptions and we only need additional assumptions on a part of our model to back-out the structural parameters of interest.

### 1.6.1 Empirical Model

We use a general version of our model from Section 1.5. This model includes bid heterogeneity and firm behavior as a function of exogenous shocks. We model bidder learning parametrically and avoid parametric assumptions on firm behavior and the latent-bid distribution.

We introduce new notation to describe bidder behavior. We focus on a specific bidder $i$ whose first overbid is at time $t \in \{1, ..., \infty\}$. We observe this bidder from their first overbid until the end of our sample. Since this time differs between bidders, we aggregate bidder's outcomes over a standardized period of time. We assume that a bidder i's own bids today have no effect on bids she faces in the future, so that we can treat different auctions as independent observations. Thus, the stable unit treatment value assumption is satisfied and, consequently, the behavior of all other bidders is uncorrelated across auctions. To set notation, we collect all other bidders in the set $J_t = \{1, ..., N_t\}$.

Our model uses exogenous shocks to model empirically relevant sources of heterogeneity. We assume that in each period, a fixed price shock $\tilde{p}_t$, an auction quantity shock $\tilde{q}_t$ and an individual specific time-varying shock $v_{i,t}$ realize as independent draws from a continuous distribution. We assume that these shocks are independent from each other. We model individual specific unobserved heterogeneity with the time-constant variable $u_i$, which in our model realizes before any choices are made and is independent across individuals. Auction-specific characteristics are denoted by $A_t$.

As before, we separate the bidding process into latent (unmodeled) bids and a bidder type (overbidder, non-overbidder, dropout) that determines the submission of these bids. We refrain from modelling the process of mapping valuations into latent bids because of the highly complicated nature of our dynamic auction. For example, bidding would depend on the observed pace of other bidders bids being submitted, which we neither observe, nor do we think it adds much to our analysis. We are mainly interested in whether bidders learn to avoid a specific mistake or drop out of the market. Our approach assumes that other differences in the bidding process between auctions are independent of the learning margins we describe.

[20]Sometimes also called Causal Graph

We let the individual latent bid depend on auction-specific characteristics $A_t$, the individual specific time-varying shock $v_{i,t}$ and the time-constant unobserved heterogeneity $u_i$. These shocks are i.i.d. from a continuous distribution. We denote the latent bid of the individual in question by $\beta_{i,t} = \beta(A_t, u_i, v_{i,t})$ and the set of latent bids by the other bidders by $\beta_{-i,t} = \{\beta(A_t, u_j, v_{j,t})\}$, where $j \neq i$. Together, the latent bids by bidder i and competing bidders $j \neq i$, $\beta_{i,t}$ and $\beta_{-i,t}$, represent the the latent demand of all bidders in auction $t$.

The dependence of $\beta_{it}$ on auction characteristics, $A_t$, models that different individuals might be interested in different auctions. A special case of this is that most individuals do not participate in an auction. In this case their latent bid is 0. The dependence on $u_i$ models that different bidders might differ in the amount they usually bid. The time-varying individual specific shock $v_{it}$ models the main source of heterogeneity for a specific bidder across auctions.

According to the model in Section 1.5 and our discussion with management, the firm targets its quantities and fixed prices to latent bidder demand. While our empirical analysis focuses on the period before the structural break when the firm likely does not behave optimally, we still allow for the fixed price $p_t$ and the auction quantity $q_t$ to depend on latent bids $\beta_{it}$ and $\beta_{-it}$. Since we do not specify the parametric form of this dependence, we allow for optimal as well as non-optimal firm behavior in our empirical analysis. Further the firm might tailor fixed-prices to auction characteristics, e.g. the type of products on sale. We model this by letting the fixed price and the auction quantity depend on auction-characteristics $A_t$. Thus, auction quantity ($q_t$) and fixed-price ($p_t$) may depend on these quantities as well as their specific exogenous shock ($\tilde{q}_t$ and $\tilde{p}_{st}$).[21]

$$p_t = c(\tilde{p}_t, A_t, \beta_{it}, \beta_{-it})$$
$$q_t = d(\tilde{q}_t, A_t, \beta_{it}, \beta_{-it})$$

As in Section 1.5, we assume that a bidder's bid depends on their type $\theta_{i,t}$ and their latent bid $\beta_{it}$. The bidder's type at time $t$ is $\theta_{i,t} \in \{o, s, d\}$, where we denote overbidders by $o$, non-overbidders by $s$, and dropouts by $d$. Since we consider bidders after their first overbid, we only select overbidders. Overbidders always bid their latent bid $\beta_{i,t}$, while non-overbidders wait until the price drops below the fixed price, that is they bid $\min\{\beta_{i,t}, p_t\}$. Bidders who drop out always bid 0. We summarize this behavior in the following bid function.

[21]Think of these shocks as supply shocks that are not captured by auction characteristics $A_t$.

$$b_{i,t} = f(p_t, \beta_{i,t}, \theta_{i,t}) = \begin{cases} \beta_{i,t} & \text{if } \theta_{i,t} = o \\ \min\{\beta_{i,t}, p_t\} & \text{if } \theta_{i,t} = s \\ 0 & \text{if } \theta_{i,t} = d \end{cases}$$

In our empirical model, we allow for heterogeneity in learning rates. Specifically, we assume that learning rates are time constant, but may differ between individual bidders. Overpaying turns overbidders into non-overbidders with probability $\iota_i$ (intensive margin), and makes them drop out with probability $\epsilon_i$ (extensive margin). We allow for dependence between these treatment effect parameters and bidder specific shocks $u_i$. In Section 1.6.4, we assume, for technical reasons, that learning rates are a function of individual characteristics, $u_i$. This assumption is natural as it still allows bidders of different individual characteristics, $u_i$, to exhibit different learning rates.

We express the auction's outcome from the perspective of bidder $i$ in terms of two order statistics of all other bids (the set $\beta_{-i,t}$): the $q_t$-highest and the $q_t - 1$-highest rival bid, which we denote by $b(q_t)$ and $b(q_t - 1)$, respectively. The $q_t$-highest rival bid determines if bidder $i$ wins, and the $q_t - 1$-highest rival bid influences the auction price. The auction ends when all products are sold, and the lowest successful bid determines the price. Bidder $i$ loses the auction if all $q_t$ units are sold to bidders in $J_t$, that is, bidder $i$ loses if $b_{i,t} < b(q_t)$. Conversely, bidder $i$'s bid is successful if $b_{i,t} > b(q_t)$. In this case, there are $q_t - 1$ units that remain for the competing bidders included in $J_t$. The lowest successful bid is then either by the bidder in question or the lowest successful bid by the other bidders ($b(q_t - 1)$). If bidder $i$ places a winning bid, the auction price is $\min\{b_{i,t}, b(q_t - 1)\}$. We define an overbid as a bid that is strictly larger than the fixed price $b_{i,t} > p_t$. Similarly, we call a bid a non-overbid when it is strictly smaller than the fixed price $b_{i,t} < p_t$. An auction is overpaid when all bids are overbids and that means that the auction price is higher than the fixed price, or $\min\{b_{i,t}, b(q_t - 1)\} > p_t$.

While we use our parametric assumptions on bidding behavior to interpret our treatment-effects, we do not need parametric assumptions to estimate these effects. For this purpose, we summarise our model as a system of non-parametric structural equations. Each structural equation expresses a left-hand side variable in terms of other variables and exogenous shocks. This model is non-parametric in the sense that we do not use any functional form assumptions on the right-hand side.

$$p_t = c(\tilde{p}_t, A_t, \beta_{it}, \beta_{-it}) \tag{1}$$

$$q_t = d(\tilde{q}_t, A_t, \beta_{it}, \beta_{-it}) \tag{2}$$

$$\beta_{i,t} = \beta(A_t, u_i, v_{i,t}) \tag{3}$$

$$b(q_t) = f(q_t, p_t, \beta_{-i,t}, \theta_{-i,t}) \tag{4}$$

$$overbid_{i,t} = g(\beta_{i,t}, b(q_t), b(q_t - 1), p_t, \theta_{i,t}) \tag{5}$$

$$non - overbid_{i,t} = v(\beta_{i,t}, b(q_t), b(q_t - 1), p_t, \theta_{i,t}) \tag{6}$$

$$overpaid_{i,t} = v(overbid_{i,t}, b(q_t), p_t) \tag{7}$$

We briefly go through each equation and relate it to the previous discussion. We explicitly introduced Equation 1 to 3 in the preceding section. Equation 1 and 2 describe the information set of the firm when setting auction parameters. One model that fits these equations is the simplified model from Section 1.5. Since these equations do not assume any structure, they nest all (optimal and non-optimal) firm policies that condition fixed prices and quantities on a signal of latent demand, $\beta_{i,t}$ and $\beta_{-i,t}$, and auction-specific characteristics, $A_t$. Equation 4 summarizes the order statistics of the rival bidder's bids, that determines whether a bid is a winning bid. Equations 5 to 7 apply the auction's rule to this section's expression of bidder's successful bids ($b_{i,t}$). We will introduce parametric versions of Equations 5 to 7 in the next section. According to Equation 5, an overbid is an observed variable (as governed by the order statistic). We will refer to overbids regardless of observation status as latent overbids throughout the text.

### 1.6.2 Interpretation of Treatment Effects

Recall that we consider the first overbid for each bidder in our sample (if there is any). That means we look at overbidders and would like to quantify type changes to non-overbidder or dropout. A bidder who changes type from overbidder to dropout just leaves the market. Consequently, we should observe fewer overbids, as well as fewer non-overbids in this case. A bidder who changes from overbidder to non-overbidder type, however, just avoids overbids in the future and bids at the fixed price whenever the latent bid is a latent overbid. Hence, bidders who become a non-overbidder bunch at the fixed price. We exclude the fixed price by focusing on strictly defined non-overbids. Thus, a first starting point to test for extensive margin learning is to estimate the treatment effect of overpaying on strict non-overbids.

To see this in more detail, reconsider Figure 2 from Section 1.5, where we depict bids as a function of latent bids and bidder type. As an example we depict a marginal uniform distribution of latent bids underneath the x-axis. The figure also shows the resulting marginal density of bids to the left of the y-axis. Non-overbidders bid at the fixed price for all latent bids larger than the fixed price, so the density of bids has a mass point at the fixed price and has zero mass at bids above the fixed price. Strictly below the fixed price, the density is identical to that of overbidders. A dropout maps all latent bids to the zero bid (not depicted in Figure 2).

Figure 2 illustrates what we can learn about latent changes in type from changes in the observed bid distribution. Bids strictly below the fixed price (strict non-overbids) are submitted by overbidders and non-overbidders and bids strictly above the fixed price (strict overbids) are only submitted by overbidders. Latent bids below the fixed price ($p_t$) directly translate into observed bids. Latent bids above the fixed price bunch at the fixed price for non-overbidders and directly translate into observed bids for overbidders. Bids at the fixed-price are composed of bunched overbids by non-overbidders and latent bids at the fixed price by both non-overbidders and overbidders.

Since strict overbids are only submitted by overbidders a decrease in these bids indicates a reduction in overbidders. Intensive as well as extensive margin learning can cause such a decrease. Since strict non-overbids are submitted by overbidders as well as non-overbidders, a decrease in these bids indicates extensive margin learning (overbidders leaving the auction). Bids directly at the fixed price increase when there are more non-overbidders and decrease when there are more overbidders. Consequently these bids increase with intensive margin learning and decrease with extensive margin learning. We focus on treatment effects of overpaying on strict overbids and strict non-overbids, in order to recover the extensive and intensive margin learning parameters, $\epsilon_i$ and $\iota_i$.

The only way to observe a non-overbid is a latent non-overbid ($\beta_{i,t} < p_t$), which is successfully submitted ($\beta_{i,t} > b(q_t)$), either by an overbidder or a non-overbidder. Thus, the treatment effect of overpaying on non-overbids in the next period is the expected extensive margin learning parameter scaled by the probability of a successful strict non-overbid. We calculate this effect conditional on $u_i$. This conditioning renders $\epsilon_i$ and the latent bid independent and thus allows the factorization. Since we are interested in learning as a response to overpaying, we also look at the most narrow way to avoid overpaying, which is learning not to overbid. In our model only overbidders submit overbids. Thus, a type transition from overbidder to non-overbidder, as well as dropping out reduces overbidding. Consequently, the treatment effect of overpaying on overbidding in the next period is given by the sum of learning at both margins multiplied by the probability of a latent successful overbid.

**Proposition 1.1.** *The treatment effects of overpaying in $t$ on the number of non-overbids and number of overbids in $t+1$ take the following form.*

$$E[TE^{t+1}_{non-overbid}|u_i] = -E[\epsilon_i|u_i]\mathbb{P}(p_{t+1} > \beta_{i,t+1} > b(q_{t+1}))|u_i)$$

$$E[TE^{t+1}_{overbid}|u_i] = -E[\epsilon_i + \iota_i|u_i] \cdot \mathbb{P}(\beta_{i,t+1} > p_t \wedge \beta_{i,t+1} > b(q_{t+1}))|u_i)$$

Proposition 1.1 reminds us that the treatment effect of overpaying on the number of overbids is a function of learning rates at the extensive margin and intensive margin. To disentangle the two margins it would be sufficient to divide the treatment effect by the probability of observing a latent overbid (non-overbid) and taking the difference of the two treatment effects. It turns out that the probability of observing a latent overbid (non-overbid), is simply the potential outcome of the untreated.[22] In Subsection 1.6.4, we explain how the analysis is complicated by the fact that we need to pool observations in order to estimate the treatment effects. Before we get to pooling of observations we explain, however, how a shift in the latent bid distribution as a result of overpaying would change our results.

### 1.6.3 Shift in Latent Bid Distribution

Bidders may react to overpaying in more general ways than the two adjustments we impose, truncating the bid function at the fixed price or dropping out. Suppose, for example, that a bidder, who overpays, may, in addition to truncating their bid function at the fixed price, increase the shading of their bid. Shading one's bid as a reaction to overpaying is a natural adjustment: the auction price a bidder has to pay is weakly lower than her bid and so shading said bid reduces payments.

We show an example of such a bid shading adjustment in Figure 3. Instead of simply truncating, the additional bid shading shifts the bid function downward (in red). This is, of course, a departure from the behavioral type change to non-overbidder we impose (in blue), so it's worth reviewing how this would impact our results. The treatment effect of overpaying on future overbids is unchanged since the bidder truncates her bid function at the fixed price, $p_{f,t}$, as before. A difference may arise when we consider the treatment effect on non-overbids, which are strictly smaller than the fixed price. Some latent bids that are mapped to the fixed price are now handed in below the fixed price due to the additional bid shading. This is the interval marked with + in Figure 3. On the other hand, some bids close to the lowest competing bid in the same auction, $\beta^{q_t}$ are pushed below this lowest competing bid and thus are not handed-in in the auction — these bids are losing because of the bid shading. In Figure 3 these latent bids are in the interval marked with $-$.

How well our approach using two behavioral types to capture learning works in the presence of unmodeled bid shading depends on the relative size of the intervals marked with + and $-$. In Figure 3, we depict the case of a linear bid function (and a uniform distribution of latent bids). As it turns out, the intervals + and $-$ cancel exactly. In this case, imposing the assumed changes in behavioral types is a good approximation of this more general adjustment of truncating and shading the bid function.

[22]We can calculate this potential outcome of the untreated from our regression in Section 1.7 by setting the treatment dummy to 0 and all other variables to the sample mean.

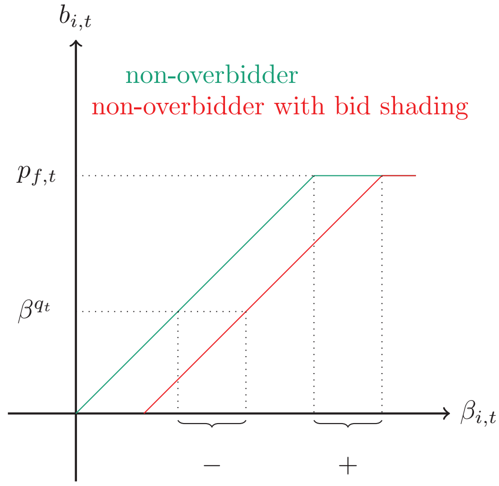



Figure 3: Truncating and shading bids as response to overpaying. Some latent bids are shaded out of the auction (−) and some that were otherwise mapped to the fixed price are now shaded in to the non-overbid region below the fixed price (+).

Proposition 1.2 discusses the effect of bid shading on our estimation strategy for arbitrary latent bid distributions. Here, the difference in shaded, $\mathbb{P}'$ and non-shaded bid distributions, $\mathbb{P}$, may be negative or positive. If bid shading shifts more bids below the fixed price than it pushes out of the auction, the difference of the shaded and non-shaded distributions is positive and thus we would underestimate the extensive margin using our two-type approximation. If, on the other hand, bid shading prices more latent bids out of the auction than it pushes bids below the fixed price, the difference in bid distributions is negative, $\mathbb{P}'(.) - \mathbb{P}(.) > 0$,. In this case, our approach mis-attributes some of the bid shading to dropouts and thus we would overestimate the extensive margin.

**Proposition 1.2.** *If non-overbidders shift their distribution of latent bids compared to overbidders, the treatment effect of overpaying on non-overbids in the next period is given by,*

$$E[TE^{t+1}_{non\text{-}overbid}] = -E[\epsilon_i|u_i] \cdot \mathbb{P}(p_{t+1} > \beta_{i,t+1} > b(q_{t+1})|u_i)$$
$$+E[\iota_i|u_i] \cdot \underbrace{\left[\mathbb{P}'(p_{t+1} > \beta_{i,t+1} > b(q_{t+1})|u_i) - \mathbb{P}(p_{t+1} > \beta_{i,t+1} > b(q_{t+1})|u_i)\right]}_{Shift\ in\ latent\ bid\ distribution},$$

*where $P'$ is a probability calculated from the latent bid distribution of non-overbidders and $P$ is a probability calculated from the latent bid distribution of overbidders.*

The proof of this result is in Appendix A.3.

In the eyes of the firm's management, these differences are somewhat muted. In one case the firm loses a sale because a consumer drops out of the market. In another case, the firm loses a sale because a consumer shades her bid so much that she does not bid in the auction anymore — effectively dropping out. In particular, a consumer who always bid very low, say 1€, is observationally indistinguishable from a consumer who does not take part in the auction.

### 1.6.4 Pooling of Observations Over a Period of Time

Our analysis, so far, focuses on treatment effects of overpaying in $t$ for outcomes of interest in $t+1$, i.e. for the next auction. Bidders do not generally take part in every auction, so we need to pool observations over a period of time in order to estimate treatment effects. To calculate a pooled treatment effect, we simply sum over the treatment effect in $t+k$, where $k$ covers the period of time over which we are pooling observations. The derivation of the period $t+k$ treatment effect is largely similar to the treatment effect in $t+1$, but for the possibility of subsequent treatments.

Subsequent treatments may occur in both the treatment and the control group. In the control group, subsequent treatments are likely since overbidders did not have the opportunity to learn in period $t$ (they are the control after all). In the treatment group, subsequent treatments may happen, whenever bidders do not learn from their initial treatment — either by chance or because their learning rate is small. The subsequent treatments in control and treatment group bias our results in opposite directions: including treated bidders in the control group biases downwards, while multiple treatments in the treatment group increases the likelihood of finding a learning effect.

In Proposition 1.3 we show that the treatment effects are attenuated rather than exacerbated by the possibility of subsequent treatments due to pooling over a period of time. The intuition behind Proposition 1.3 is that subsequent treatments are more likely to occur in the control group than in the treatment group. This is the case, as a bidder in the treatment group can only receive an additional treatment if he fails to learn from the first treatment. No such condition applies for subsequent treatments of the control group and, thus, our results are conservatively estimated.

**Proposition 1.3.** *Let $p_{l,k}$ denote the probability that a bidder changes his type from overbidder to dropout because of a treatment in subsequent periods $t+1$ to $t+k$. Similarly, let $p_{s,k}$ denote the probability that a bidder changes type from overbidder to non-overbidder due to a treatment in periods $t+1$ to $t+k$. Then, the treatment effect of the initial treatment in period $t$ on non-overbids and overbids in period $t+k$ is given by the following expressions.*

$$E[TE^{t+k}_{non\text{-}overbid}|u_i] = E\big[-\epsilon_i E[p_{l,k}|\epsilon_i,\iota_i,u_i]|u_i\big]\mathbb{P}(p_{t+k} > \beta_{i,t+k} > b(q_{t+k})|u_i)$$
$$E[TE^{t+k}_{overbid}|u_i] = E[-(\epsilon_i+\iota_i)E[p_{l,k}+p_{s,k}|\epsilon_i,\iota_i,u_i]|u_i]\mathbb{P}(\beta_{i,t+k} > p_{t+k} \wedge \beta_{i,t+k} > b(q_{t+k})|u_i)$$

The proof of this result is in Appendix A.4. Note that the $t+k$ period treatment effects in Proposition 1.3 are identical to the corresponding $t+1$ period treatment effects in Proposition 1.1, but for the scaling factor $E[p_{l,k}|\epsilon_i,\iota_i,u_i]$ and $E[p_{l,k}+p_{s,k}|\epsilon_i,\iota_i,u_i]$, respectively. Since these factors are conditional probabilities, they are (at least weakly) smaller than 1 and thus the treatment effects in later periods are attenuated rather than exacerbated by subsequent treatments due to pooling of observations.

To back out the extensive and intensive margin learning parameters, $\epsilon_i$ and $\iota_i$, from the pooled treatment effects, we assume that individual heterogeneity in learning rates is a function of the unobserved individual characteristics, $u_i$. This restricts learning rates to be homogeneous for individuals with the same unobserved individual characteristics. Formally, we denote $E[\epsilon_i|u_i] = \epsilon_u$ and $E[\iota_i|u_i] = \iota_u$ and we call it heterogeneous-learning-rates assumption. We need to make this technical assumption since our pooled treatment effects do not factorize otherwise. The heterogeneous-learning-rates assumption is natural, since it still allows heterogeneity in learning rates along the heterogeneity in individual characteristics, $u_i$.

Recall, that the treatment effect in Proposition 1.1 is just the conditional expectation of the extensive margin learning parameter, $E[\epsilon_i|u_i]$, scaled by the probability to observe a latent non-overbid. Untreated individuals are still participating in the auctions and submit their overbids, so dividing our treatment effect by the potential outcome of the untreated $E[non-overbid_t^{t+1}(0)|u_i]$, yields the conditional expectation of the extensive learning parameter.

$$\frac{E[TE_{non-overbid}^{t+1}|u_i]}{E[non-overbid_t^{t+1}(0)|u_i]} = \frac{-E[\epsilon_i|u_i]\mathbb{P}(p_{t+1} > \beta_{i,t+1} > b(q_{t+1}))|u_i)}{\mathbb{P}(p_{t+1} > \beta_{i,t+1} > b(q_{t+1}))|u_i)} \tag{8}$$

$$= -E[\epsilon_i|u_i] \tag{9}$$

It remains to account for the pooling of observations to estimate the effect. In Proposition 1.3, we provide expressions for the treatment effect in some subsequent auction $t+k$. Summing over these treatment effects gives us the treatment effect that we estimate from pooled data. Proposition 1.1 provides these sums over the treatment effects on non-overbids and overbids divided by the appropriate potential outcome as shown in Equation 8. The proof is in Appendix A.5.

**Lemma 1.1.** *Suppose individuals are treated at time $t \in \{1, ..., \infty\}$ and we aggregate our treatment effects over the following $k \in \{1, ..., \infty\}$ periods. Then the treatment effects divided by the potential outcomes are given by the following expressions:*

$$\frac{\Sigma_{m=0}^k E[TE_{non\text{-}overbid}^{t+m}]}{\Sigma_{m=0}^k E[non\text{-}overbid_t^{t+m}(0)|u_i]} = \frac{\Sigma_{m=0}^k E[-\epsilon_i E[p_{l,m}|\epsilon_i, \iota_i, u_i]|u_i]}{\Sigma_{m=0}^k E[E[p_{l,m}|\epsilon_i, \iota_i, u_i]|u_i]}$$
$$\frac{\Sigma_{m=0}^k E[TE_{overbid}^{t+m}]}{\Sigma_{m=0}^k E[overbid_t^{t+m}(0)|u_i]} = \frac{\Sigma_{m=0}^k E[-(\epsilon_i + \iota_i) E[p_{l,k} + p_{s,k}|\epsilon_i, \iota_i, u_i]|u_i]}{\Sigma_{m=0}^k E[E[p_{l,k} + p_{s,k}|\epsilon_i, \iota_i, u_i]|u_i]}.$$

Unfortunately, the expressions in 1.1 do not immediately simplify because the probabilities of subsequent treatment ($p_{l,k}$ and $p_{s,k}$) are functions of the corresponding learning parameters

($\epsilon_i$ and $\iota_i$). Consider, for example, the expression $E[\epsilon_i E[p_{l,m}|\epsilon_i, \iota_i, u_i])|u_i]$, where $E[\epsilon_i|u_i]$ does not factor out as $\epsilon_i$ and $E[p_{l,m}|\epsilon_i, \iota_i, u_i]$ are dependent.

Assuming that $\epsilon_i$ is a function of the individual characteristics $u_i$ solves this problem and allows us to recover $\epsilon_i$ from the pooled treatment effects. This assumption effectively means that we are restricting individual heterogeneity in learning rates $\epsilon_i$ and $\iota_i$ to be captured by heterogeneity in individual characteristics $u_i$. In other words, we have to assume that learning rates are homogeneous for every value of individual characteristics, $u_i$. As a short hand for this assumption we write $E[\epsilon_i|u_i] = f(u_i) = \epsilon_u$. Corollary 1.1 shows that using this assumption our expressions from Proposition 1.1 simplify and we can recover the learning parameters $\epsilon_u$ and $\iota_u$ from our treatment effects.

**Corollary 1.1.** *Under the heterogeneous-learning-rates assumption, heterogeneity in learning is restricted to the heterogeneity in individual characteristics, that is, $E[\epsilon_i|u_i] = \epsilon_u$ and $E[\iota_i|u_i] = \iota_u$. Then, we can recover the learning parameters $\epsilon_u$ and $\iota_u$ from the pooled treatment effects.*

$$\frac{\Sigma_{m=0}^{k} E[TE_{non\text{-}overbid}^{t+m}]}{\Sigma_{m=0}^{k} E[non\text{-}overbid_t^{t+m}(0)|u_i]} = \epsilon_u$$
$$\frac{\Sigma_{m=0}^{k} E[TE_{overbid}^{t+m}]}{\Sigma_{m=0}^{k} E[overbid_t^{t+m}(0)|u_i]} = \epsilon_u + \iota_u$$

Note that the probabilities of subsequent treatments $p_{l,k}$ and $p_{s,k}$ depend on the time period we pool over. Thus, violations of the assumption in Corollary 1.1 should lead to incongruous results, when we pool over different time periods. Indeed, in Section 1.7 we show that our results do not depend on the pooling period used.

### 1.6.5 Identification of Treatment Effects

After clarifying the connection between estimable treatment effects and the learning parameters of our underlying model in the previous section, we now turn to the identification of those treatment effects. Recall that a bidder is assigned to the treatment (control) group when his first overbid (did not) lead to overpayment and that whether an overbid leads to overpayment is entirely determined by the rival bidders in that auction. A remaining concern with this design is selection into treatment. This is an issue if, for example, bidders who learn well select into watch auctions, while bidder who do not learn well spread out evenly over all product categories. For a rigorous analysis of this logic, we represent our empirical model in Equation 1 to 7 as a causal Directed Acyclic Graph (DAG).[23]

The DAG representation illustrates the causal relationships implied by the structural equations model and allows us to compute a set of control variables to satisfy the conditional

[23]To put it precisely, we interpret our Structural Equation Model (SEM) in Section 1.6.1 as a Structural Causal Model (SCM).

independence assumption required for identification. In our case, identification depends in part on unobserved bidder characteristics, so we conclude with a discussion of how we can implement our identification strategy using past bidder behavior as a proxy for this unobserved variable.

In a DAG, a directed edge (an arrow) indicates a causal relationship, that is, the node where the arrow originates is a cause of the node that the arrow points to. For example, if we draw an arrow from $overpaid_{i,t}$ to $overbid_{i,t+1}$, we show that our model allows for a causal effect of overpaying on overbidding in a subsequent auction. In our context, the fact that DAGs do not contain any cycles has an economic interpretation: bidders are myopic. Otherwise, future auctions would influence bidding behavior in today's auction, which would lead to a cycle in our graph. This assumption is in line with other behavioral economics auction papers such as Malmendier and Lee (2011). We try to explain the theory on DAGs as we go along.[24]

To generate a DAG from our empirical model in Section 1.6.1, we go through each equation and draw an edge from each right-hand side variable to each left-hand side variable.[25] We leave out exogenous shocks for ease of exposition[26] and draw boxes around variables that are observable. The procedure results in the DAG depicted in Figure 4.

We use $y_{i,t+1}$ as a stand-in for the outcomes we are interested in: revenue and number of overbids and non-overbids in subsequent auctions. We focus on time period $t$ and display arrows pointing from $t$ to $t+1$ only in a stylized way. In particular, the path $u_i \rightarrow y_{i,t+1}$ abstracts from the fact that this causal relationship is again channeled through the bidding process. This simplification is without loss of generality, since $u_i$ is the only connection between behavior in $t$ and $t+1$. We also abstract from bidding behavior before $t$. We restrict our data set to behavior after the first overbid in $t$. This restriction selects only bidders who are overbidders in $t$ and thus, there is no remaining variance in the bidders behavioral type, $\theta_{i,t}$, and we can omit it form the DAG.

In a DAG, the paths where all arrows point from the treatment to the outcome variable are called front-door paths, or causal paths. This is the causal relationship of interest, in our case $overpaid_{i,t} \rightarrow y_{i,t+1}$. There are also paths from the treatment to the outcome, where at least on arrow points in the opposite direction, called back-door paths. In our case, all other paths from treatment to outcome are back-door paths, since every path other than the causal path starts with an arrow pointing to $overpaid_{i,t}$ (instead of originating from it). The main idea of proving identification in a DAG is to select control variables to block all back-door paths.

Panel A of Figure 4 shows the origin of our causal graph. The arrows in blue encode institutional knowledge about our setting. For example, the director of the auction sees latent demand and can choose quantity accordingly, so there are arrows from the latent bids $\beta_{i,t}$ and

[24]For a gentle introduction, see chapter 3 of Cunningham (2021).

[25]See, for example, Peters et al. (2017) for a more complete treatment of the connection between DAGs and Structural Causal Models.

[26]This is without loss of generality because these shocks are exogenous by assumptions, so no edges point to these nodes.

$\beta_{-i,t}$ to quantity $q_t$. Similarly, the seller incorporates that auction characteristics will have an impact on demand, quantity and fixed prices when planning the auction so there are arrows from auction characteristics, $auction_t$, to latent demand, $\beta_{i,t}$ and $\beta_{-i,t}$, quantity, $q_t$, and the fixed price $p_{f,t}$.

The arrows in violet depict the auction rules, namely uniform pricing and our definitions of overbidding and overpaying. The order statistic $\beta_{-i,t}^{(q_t)}$ determines winning bids and what price winning bidders have to pay. As bidder i has to beat the $q_t$-highest rival bids to win the auction, the order statistic has arrows incoming from $q_t$ and $\beta_{-i,t}$. We only observe winning bids and consider initial overbids for each bidder, so there is an arrow from the order statistic to $overbid_{i,t}$. Together with the fixed price it is determined whether the overbid leads to overpayment.

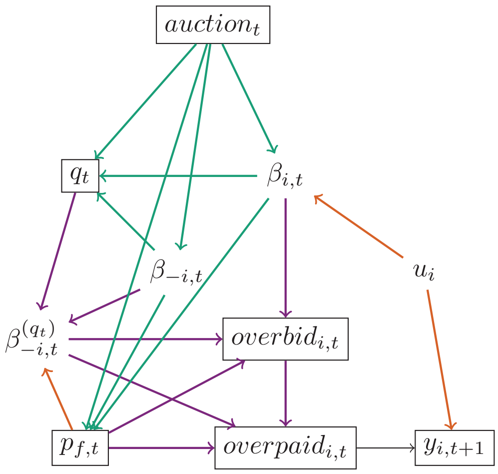
Finally, the arrows in orange depict substantial economic assumptions. Considering only initial overbidders $i$ means we can omit the behavioral type of those bidders from the DAG. Rival bidders, however, may be non-overbidders, so the fixed price has an influence in whether rival bidders hand in their latent bids. Thus, we draw an arrow from the fixed price $p_{f,t}$ to the order statistic of handed in bids $\beta_{-i,t}^{(q_t)}$. Finally, we assume that latent bids today and latent bids tomorrow are connected by individual characteristics. Since we abstract away the bidding process in $t+1$, we end up with arrows from individual characteristics $ind.char_i$ to latent bids $\beta_{i,t}$ and our outcomes of interest $y_{i,t+1}$.

Our effect of interest is the black arrow from $overpaid_{i,t}$ to outcome $y_{i,t}$. Threats to identification are posed by, so-called, back-door-paths, which are paths that start with an arrow going into our treatment indicator $overpaid_{i,t}$ and go to the outcome $y_{i,t+1}$, but not through the the direct arrow. The back-door paths consist of two patterns: confounders (e.g. $\leftarrow auction_t \rightarrow$) and colliders (e.g. $\rightarrow overbid_{i,t} \leftarrow$).[27] A back-door path through a confounder is blocked if we control for that confounder. A back-door path through a collider is blocked if we do not control for that collider, instead it is undesirably opened (in DAG lingo) if we control for the collider (cf. bad control problem Angrist and Pischke (2009)).

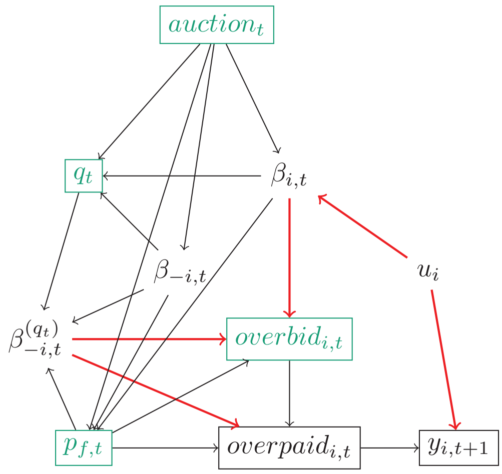
In Panel B of Figure 4, an adjustment set that blocks confounder paths is highlighted in blue. That is, we block all confounder paths if we include the set $\{overbid_{i,t}, A_t, q_t, p_t\}$ as control variables in our regression. We control for auction price and quantity directly and we operationalize auction characteristics using fixed effects such as weekday, week, hour, product category and auctioneer fixed effects. Additionally, we condition on $overbid_{i,t}$ as we only look at bidders first overbids. This conditioning on the first overbid, however, is not without drawbacks. Indeed, Panel B in Figure 4 delineates the collider path that is opened by restricting the analysis to first overbids: $\beta_{-i,t}^{(q_t)} \rightarrow overbid_{i,t} \leftarrow \beta_{i,t}$. That is, by conditioning we on $overbid_{i,t}$ we leave open the possibility that bidders with high latent bids select into similar auctions and that this drives treatment.

[27] Readers interested in graph theory will recognize these patterns as forks and inverted forks.

Fortunately, the same graph also shows that the collider path also passes through individual bidder characteristics $u_i$. In fact, all back-door paths go through $u_i$, so we could block them all by simply conditioning on $u_i$. While this is an elegant solutions, it is complicated by the fact that $u_i$ is unobserved. Thus, we have to rely on proxies that are, by definition, imperfect. The variable $u_i$ mainly determines the height of a bidder's latent bid. Thus, variables such as the average amount of a bidder's past behavior and experience in the auctions are very informative about individual characteristics. We calculate bidder history variables, both for bidder $i$ and the rival bidders (see Appendix A.7).




(a) Origin of the DAG




(b) Collider Path

Figure 4: Panel (a) shows the origin of the DAG: Seller planning and running the auction (blue), Auction rules (uniform pricing, violet), Economic Assumptions (orange). Panel (b) shows that the collider path (red) opened by conditioning on $overbid_{i,t}$ (blue) also goes through $ind.char_i$.

We formalize our empirical strategy with the back-door criterion (Theorem 3.3.2 in Pearl (2009)). As we have shown $\{ind.char_i, auction_t, p_t, q_t, overbid_{i,t}\}$ or $\{ind.char_i\}$ block all back-door paths. Thus the causal effects of overpaying on future overbids and future non-overbids are identified and can be computed by controlling for these variables. This statement is equivalent to the statement that our potential outcomes are independent conditional on $\{ind.char_i, auction_t, p_t, q_t, overbid_{i,t}\}$ or $\{ind.char_i\}$.

## 1.7 Regression Results

We adjust for overbidding by restricting the sample to the first overbid for any customer. These initial overbids can be in an auction that ends below or above the fixed price. Bidders whose initial overbid was in an overpaid auction overpay and are in our treatment group. Bidders whose initial overbid was not in an overpaid auction do not overpay and form the control group. We follow these bidders for 90 days after their first overbid and count the number of overbids and non-overbids during that period. We exclude data after the structural break, because overbids are much less likely due to firm policy after the break (see Figure 1).

To estimate the treatment effects in Proposition 1.1, we run the regression in Equation 1.7 on the sample of bidders first overbids.

$$Y_{i,t+k} = \beta_1 overpaid_{i,t} + \beta_2 p_t + \beta_3 q_t + H_{i,t}B_1 + H_{-i,t}B_2 + A_t + \eta_{i,t+k}$$

Here, $overpaid_{i,t}$ indicates treatment when the first overbid of bidder $i$ in auction $t$ lead to overpayment. Following our analysis of the causal graph in Section 1.6, we control for the auction price $p_t$, auction quantity $q_t$ and a set of history controls as proxies of individual characteristics. We include bidder history variables, such as the average bid before the first overbid, both for bidder $i$, $H_{i,t}$, as well as the rival bidders in the same auction, $H_{-i,t}$. The full set of bidder history controls is described in detail in Appendix A.7. We use weekday, week, hour, product category, and auctioneer fixed effects to capture auction characteristics, $A_t$. Treatment is assigned at the auction level, so we cluster standard errors at the auction level (Abadie et al., 2017).

The variable $Y_{i,t+k}$ is a stand-in for revenue and the number of overbids and number of non-overbids in a $k$ day long period. We use 0 to 90 and 90 to 180 days after the first overbid to provide estimates with a varying time frame. Finding similar results should reinforce confidence in the assumptions needed to recover the learning rates from the pooled treatment effects as argued in Corollary 1.1. We also report regressions excluding the history controls.

Table 1.7 shows the results of our revenue regressions. We find overpaying reduces revenue by roughly 9.95€ in the first 90 days after treatment. Compared to the revenue threshold we calculate in Section 1.4.1 this effect is sufficient to determine that the extraction of overpaying revenues was indeed suboptimal. The results for the period 90-180 days after the treatment are qualitatively similar, suggesting additional revenue loss of 7.31€, albeit statistically insignificant (see Appendix A.9).

Table 3: Overpaying Reduces Future Revenue

| | Revenue | Revenue |
|---|---|---|
| Overpaid | −5.079 | −9.955** |
| | (3.981) | (4.852) |
| Num.Obs. | 115 295 | 71 261 |
| R2 | 0.043 | 0.100 |
| Counterfactual Mean | 138.206 | 160.407 |
| Bidder History | No | Yes |
| Window | 0-90 | 0-90 |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Overpaying reduces revenue in the 90 days after treatment by 9.95€. This revenue effect is larger than the revenue threshold we calculate in Section 1.4.1 of 9.77€ and we find further revenue effect for the period of 90-180 days after treatment (see Appendix A.9.

Table 4: Overpaying Reduces #Overbids and #Non-Overbids

| | # Overbids | # Overbids | # Non-Overbids | # Non-Overbids |
|---|---|---|---|---|
| Overpaid | −0.171*** | −0.190*** | −0.185* | −0.288** |
| | (0.032) | (0.034) | (0.110) | (0.127) |
| Num.Obs. | 115 295 | 71 261 | 115 295 | 71 261 |
| R2 | 0.074 | 0.131 | 0.071 | 0.154 |
| Cf. Mean | 1.433 | 1.62 | 5.796 | 6.854 |
| Bidder History | No | Yes | No | Yes |
| Window | 0-90 | 0-90 | 0-90 | 0-90 |

* p < 0.1, ** p < 0.05, *** p < 0.01

The negative impact of overpaying on number of overbids indicates treated bidders repeat their mistake less often than untreated. Larger negative effect on number of non-overbids is evidence of adjustment at extensive margin.

We also consider the number of overbids and non-overbids in our regression analysis, as this allows us to back-out the underlying learning rates from our three-type model. Table 1.7 reports the results of these regressions. The first row reports the causal effect of overpaying on future overbids and non-overbids. The row labelled counterfactual mean reports the fitted value for the regression with all variables set at their means and overpaid set to zero. With the full set of controls we find that overpaying decreases overbids in the following 90 days by $-0.19$ (compared to a counterfactual mean of 1.62) and non-overbids by $-0.288$ (compared to a counterfactual mean of 6.8). Results for the time period 90-180 days after treatment are similar (see Appendix A.9).

We assess our strategies of using bidder histories to proxy for $u_i$ by looking at coefficient movements when adding these variables. Since our proxies have a good theoretical justification (high bids in the past are likely a good indicator of a tendency for high bids), our estimates should move closer to the truth when controlling for these proxies. Thus, if the magnitude of our estimates increases when we add the proxies, it should increase even more if we could actually control for $u_i$ (see Oster, 2019, for a formalisation of this argument). Adding history controls (our proxies) increases the magnitude of our coefficient estimates. We take this as evidence that our identification strategy works well.

We recover the extensive and intensive margin learning rates as laid out in Corollary 1.1. Since we have to pool auctions to make estimation feasible, the expressions in the Corollary rely on the time period we aggregate over. Using two different time periods affords us a plausibility check: if the results are consistent it reinforces our confidence in the validity of our assumptions.

We report the backed out learning rates in Tables 5. We repeat the exercise using a Poisson regression model to complement the linear regression (the corresponding regression Tables are in Appendix A.10). The results are quite consistent between the two periods of time and the different regression methods, affirming our confidence in the heterogeneous-learning-rates assumption. We estimate the average overbidder has an extensive margin learning parameter of approximately 4% and an intensive margin learning parameter of approximately 7%. In other words, overpaying causes roughly 4% of overbidders to drop out of the market due to overpaying. This suggests extracting the extra overpaying revenue in the beginning of our sample was indeed suboptimal (the extensive margin threshold in Section 1.4.1 is 2.7%). This reinforces our conclusion from the revenue regressions that extracting overpaying revenue is indeed suboptimal.

Table 5: Backed out Learning Rates

| Linear model | $\epsilon_u$ | $\iota_u$ |
|---|---|---|
| 0-90 | 0.042 | 0.075 |
| 90-180 | 0.041 | 0.062 |

| Poisson model | $\epsilon_u$ | $\iota_u$ |
|---|---|---|
| 0-90 | 0.039 | 0.075 |
| 90-180 | 0.033 | 0.066 |

## 1.8 Conclusion

We find evidence for extensive as well as intensive margin learning as overpaying decreases future overbids and non-overbids. We use our model to recover extensive and intensive margin learning from these treatment effects. The causal effects imply that roughly 4% of bidders who overpay learn to drop out of the market (extensive margin), and about 7% of bidders who overpay learn not to overbid (the intensive margin).

A simple economic model teaches us how the firm should react to learning at these margins. We model an increase in firm sophistication by a broader scope of optimization: initially, the firm behaves sub-optimally, exploiting extra revenue from overpaying. Then, the firm behaves optimally along the quantity-adjustment margin, but treats fixed prices as exogenous, even though it has control over fixed prices. Finally, the firm chooses fixed prices and quantity optimally. If fixed prices are exogenous and extensive margin learning is high, the firm offers quantities that prevent overpaying. On the other hand, if the firm endogenously chooses fixed prices, it sets them high enough to prevent overbidding entirely, while still achieving a high price in the auction.

We make a number of observations that are in line with our model. First, we observe a period with overpaying in the market, followed by a sudden elimination of overpaying and increased quantity. This is in line with the initially suboptimal behavior of the firm and a policy change to target optimal quantity dynamically in the auctions, while keeping fixed prices unchanged. Second, we document a policy change after our sample ends. The new policy increases fixed prices and undercuts these higher prices with the auction's starting bid, and thus, ruling out overbidding by definition. This policy is in line with our model with endogenous fixed prices.

We find that strategic learning and leaving the market are roughly equally likely. This finding unites the literature on learning (Haselhuhn et al., 2012; Agarwal et al., 2013; Ater and Landsman, 2013) and customer retention (e.g. Seru et al., 2010; Backus et al., 2021; Anderson and Simester, 2010). From the perspective of the learning literature, consumers try to avoid the action that had negative consequences: they avoid overbidding because they overpaid. According to the literature on customer retention, they might also leave the market. Bidders might leave the market because they learn about their abilities as bidders or the value of participating in auctions. They can also become angry and leave the market.

That the firm is shaping this learning process is a novel reason for the persistence of consumer biases. The previous literature finds that firms can exploit consumer biases because consumers forget what they have learned (Agarwal et al., 2013), or new naive consumers replace experienced ones (Wang and Hu, 2009; Augenblick, 2016). We document that biases may also persist because firms make learning harder, when it is profit maximizing to do so. Our results on firms shaping consumer learning can explain market design choices and suggests possible avenues for regulating this behavior.

Complementing previous results, we find that when biased consumers learn, market-like institutions might be preferable to multiple single-unit auctions. Malmendier and Szeidl (2020) argue that firms want to sell several goods in individual auctions to fish for fools. In single-unit auctions, the highest bidder (likely upward biased) sets the price, whereas, in markets (and in the market-like auction we study), a larger share of biased buyers is needed to influence the price. According to Malmendier and Szeidl (2020) choosing individual auctions maximizes period profits. We show, however, that this may cost the firm customers because more individual auctions end overpaid. Consequently, sellers should be more likely to choose markets when bidders learn.

Firms can shape consumer learning in two ways: ways that benefit and ways that harm consumers.[28] According to our model, consumers are worse off when firms can change reference prices. In this case, the firm can remove the learning stimulus without benefiting the consumer. If the reference prices are exogenous, the firm prevents consumer learning through lower prices, which is in the interest of consumers.

In our setting, reference price regulation can constrain a firm's harmful ways of shaping consumer learning. For example, a regulator could mandate a minimum revenue share through sales at fixed prices. While the practical implementation of such a policy is uncertain, it diminishes a firm's ability to raise fixed prices. Consequently, firms have to shape consumer learning through higher quantities, which benefits the consumer. There are already other types of reference price regulation. In Germany, for example, firms that advertise undercutting a reference price need to offer that reference price for a sufficient amount of time.[29]

We provide a foundation for further research on customer retention and learning in platform

[28]We do not model consumer preferences. Consequently, our only criterion for welfare analysis is that a lower price for the same quantity is good for consumers.

[29]`https://www.frankfurt-main.ihk.de/recht/uebersicht-alle-rechtsthemen/wettbewerbsrecht/unlauterer-wettbewerb/irrefuehrende-werbung/mondpreise-5196206` accessed: 2.02.2022.

markets. While we study policies specific to our context (higher quantities and higher fixed prices), these policies suggest a general pattern. Consumers learn from negative experiences. Consequently, the firm can reduce the number of negative experiences (higher quantities) or make existing negative experiences less salient (higher fixed prices). Further, more general research can build on our work and map features of existing markets into these two categories.

# References


Abadie, A., Athey, S., Imbens, G., and Wooldridge, J. (2017). When Should You Adjust Standard Errors for Clustering? Technical report, National Bureau of Economic Research, Cambridge, MA.

Agarwal, S., Driscoll, J. C., Gabaix, X., and Laibson, D. I. (2013). Learning in the Credit Card Market. *SSRN Electronic Journal*.

Anderson, E. T. and Simester, D. I. (2010). Price stickiness and customer antagonism. *Quarterly Journal of Economics*, 125(2):729–765.

Angrist, J. D. and Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Ater, I. and Landsman, V. (2013). Do customers learn from experience? Evidence from retail banking. *Management Science*, 59(9):2019–2035.

Augenblick, N. (2016). The sunk-cost fallacy in penny auctions. *Review of Economic Studies*, 83(1):58–86.

Avoyan, A., Khubulashvili, R., and Mekerishvili, G. (2021). Behavioral market design for online gaming platforms. *Available at SSRN 3690620*.

Backus, M., Blake, T., Masterov, D. V., and Tadelis, S. (2021). Expectation, Disappointment, and Exit: Evidence on Reference Point Formation from an Online Marketplace. *Journal of the European Economic Association*.

Chetty, R. (2009). Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods. *Annual Review of Economics*, 1(1):451–488.

Cho, S. and Rust, J. (2010). The flat rental puzzle. *The Review of Economic Studies*, 77(2):560–594.

Cunningham, S. (2021). *Causal Inference: The Mixtape*. Yale University Press, New Haven.

Della Vigna, S. and Malmendier, U. (2006). Paying Not to Go to the Gym. *American Economic Review*, 96(3):694–719.

DellaVigna, S. and Gentzkow, M. (2019). Uniform pricing in us retail chains. *The Quarterly Journal of Economics*, 134(4):2011–2084.

Doraszelski, U., Lewis, G., and Pakes, A. (2018). Just starting out: Learning and equilibrium in a new market. *American Economic Review*, 108(3):565–615.

Gesche, T. (2022). Reference-price shifts and customer antagonism: Evidence from reviews for online auctions. *Journal of Economics & Management Strategy*, 31(3):558–578.

Grubb, M. D. (2015). Overconfident consumers in the marketplace. *Journal of Economic Perspectives*, 29(4):9–36.

Grubb, M. D. and Osborne, M. (2015). Cellular service demand: Biased beliefs, learning, and bill shock. *American Economic Review*, 105(1):234–71.

Hanna, R., Mullainathan, S., and Schwartzstein, J. (2014). Learning through noticing: Theory and evidence from a field experiment. *The Quarterly Journal of Economics*, 129(3):1311–1353.

Haselhuhn, M. P., Pope, D. G., Schweitzer, M. E., and Fishman, P. (2012). The impact of personal experience on behavior: Evidence from video-rental fines. *Management Science*, 58(1):52–61.

Heidhues, P. and Kőszegi, B. (2018). Behavioral Industrial Organization. In *Handbook of Behavioral Economics: Applications and Foundations*, volume 1, pages 517–612. Elsevier, Amsterdam and Boston.

Hortaçsu, A., Luco, F., Puller, S. L., and Zhu, D. (2019). Does strategic ability affect efficiency? evidence from electricity markets. *American Economic Review*, 109(12):4302–42.

Hortaçsu, A. and Puller, S. L. (2008). Understanding strategic bidding in multi-unit auctions: a case study of the texas electricity spot market. *The RAND Journal of Economics*, 39(1):86–114.

Hossain, T. and Morgan, J. (2006). ... plus shipping and handling: Revenue (non) equivalence in field experiments on ebay. *Advances in Economic Analysis & Policy*, 5(2).

Hünermund, P. and Bareinboim, E. (2023). Causal inference and data fusion in econometrics. *The Econometrics Journal*, page utad008.

Imbens, G. W. (2020). Potential outcome and directed acyclic graph approaches to causality: Relevance for empirical practice in economics. *Journal of Economic Literature*, 58(4):1129–1179.

Kagel, J. H. and Levin, D. (2011). Auctions: A survey of experimental research, 1995-2010. *Handbook of experimental economics*, 2:563–637.

Kleiber, C. and Zeileis, A. (2008). *Applied Econometrics with R*. Springer New York, New York, NY.

Malmendier, U. and Lee, Y. H. (2011). The bidder's curse. *American Economic Review*, 101(2):749–787.

Malmendier, U. and Szeidl, A. (2020). Fishing for fools. *Games and Economic Behavior*.

Neugebauer, T. and Selten, R. (2006). Individual behavior of first-price auctions: The importance of information feedback in computerized experimental markets. *Games and Economic Behavior*, 54(1):183–204.

Ocker, F. (2018). "Bid more, pay less"–overbidding and the Bidder's curse in teleshopping auctions. *Electronic Markets*, 28(4):491–508.

Oster, E. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. *Journal of Business & Economic Statistics*, 37(2):187–204.

Pearl, J. (2009). *Causality*. Cambridge university press.

Peters, J., Janzing, D., and Schölkopf, B. (2017). *Elements of causal inference : foundations and learning algorithms*. The MIT press, Cambridge (MA) ;London.

Seru, A., Shumway, T., and Stoffman, N. (2010). Learning by Trading. *Review of Financial Studies*, 23(2):705–739.

Thaler, R. H. (1999). Mental accounting matters. *Journal of Behavioral Decision Making*, 12(September 1998):183–206.

Wang, X. and Hu, Y. (2009). The effect of experience on Internet auction bidding dynamics. *Marketing Letters*, 20(3):245–261.

# 2 Managerial Overconfidence in Europe

David Zeimentz, Simon Schulten, Dennis Gottschlich

## 2.1 Introduction

A bulk of empirical literature shows that CEO overconfidence is an important determinant of firm behavior.[30] These findings almost exclusively rely on evidence from the US.[31] Differences in corporate culture, corporate governance, and investor control, however, may affect the type of individuals becoming CEO and to which degree their personal traits affect firm outcomes. It is, therefore, non-obvious that the findings from the US are transferable to other cultural or institutional settings, and evidence for other domains is scarce.

We narrow this gap in the literature by analyzing the effect of overconfidence on (i) investment behavior, (ii) innovation, and (iii) compensation packages in Europe and compare our results to the findings from the US. To do so, we construct a novel data set on director transactions with firm stocks and stock options, so-called directors' dealings, that allows us, in combination with compensation data, to construct the canonical overconfidence measures as proposed by Malmendier and Tate (MT, 2005).[32] We collect these data on directors' dealings from the financial authorities in France, Germany, and the UK. Our sample ranges from 2008 to 2020.

Moreover, we propose a novel classification approach to distinguish the overestimation of own ability from a skill-independent optimism about the firm's prospects, by exploiting the increase in decision-making power a member of the executive board experiences when being promoted to CEO. The canonical overconfidence measures rely on the CEO's failure to diversify by not exercising stock options or actively increasing the exposure to firm-specific risk by net-purchasing company stock, indicating optimistic beliefs about the firm's future prospects.[33] We classify a CEO as an *optimist*, i.e. having upward biased beliefs about the firm's prospects independent of ability, when she net-purchases stocks before and during her tenure as CEO. When a CEO reveals optimistic beliefs only after being appointed CEO, we classify her as *overconfident*, i.e. having upward biased beliefs about ability. Being able to observe CEOs trading behavior before

[30] Overconfidence is, for example, predictive for investment behavior (Malmendier and Tate, 2005; Ben-David et al., 2007), merger decisions (Malmendier and Tate, 2008), innovation (Hirshleifer et al., 2012; Zavertiaeva et al., 2018; Galasso and Simcoe, 2011), financial misreporting (Schrand and Zechman, 2012), corporate failure (Leng et al., 2021), dividend policy (Deshmukh et al., 2013), hiring decisions (Campbell et al., 2011; Campbell, 2014), debt level (Hackbarth, 2008; Malmendier et al., 2022), accounting conservatism (Ahmed and Duellman, 2013), stock price crashes (Kim et al., 2016), optimistic earning forecasts (Hribar and Yang, 2016), earnings smoothing Bouwman (2014), and CEO compensation (Otto, 2014; Humphery-Jenner et al., 2016).

[31] To the best of our knowledge, Leng et al. (2021) are the only exception using the revealed preference overconfidence measures by Malmendier and Tate (2005). They study overconfidence in the UK.

[32] The existing literature relies on three US data sets to construct the overconfidence measures. Early contributions use a data set provided by Hall and Liebman (1998) of CEOs of 477 large US companies from 1980 to 1994 (e.g. Malmendier and Tate, 2005, 2008; Galasso and Simcoe, 2011). Recent papers rely on the Compustat Execucomp data set following 2006 and extend their sample up to 1992 using data from Thompson Reuters (e.g. Malmendier et al., 2011, 2022).

[33] A risk-averse CEO has a strong incentive to insure against firm-specific risk as both her personal wealth and the market's perception of her human capital are tightly linked to the firm's performance. Note that top-level managers are legally prohibited from hedging financially against firm-specific risk.

and during their tenure as CEO, is a key advantage of our data set.

The exact interpretation of the overconfidence measures proposed by Malmendier and Tate (2005) varies in the ensuing literature. The majority of the literature follows Malmendier and Tate (2005) and interprets the measures as having favorable beliefs about own ability and a strong self-attribution of firm success (e.g. Malmendier and Tate, 2008; Galasso and Simcoe, 2011; Humphery-Jenner et al., 2016; Campbell, 2014; Malmendier et al., 2022). Some papers, however, give the measures an interpretation that resembles our definition of optimism (e.g. Otto, 2014; Campbell et al., 2011; Malmendier et al., 2011; Hribar and Yang, 2016; Deshmukh et al., 2013; Banerjee et al., 2015). Although both variants are in line with the classification idea, optimism and overconfidence can have differential effects, especially in domains where ability matters. Our classification approach provides a framework to differentiate between those two interpretations.

Adopting the empirical strategies of the existing literature, we find mainly similar effects of CEO overconfidence on firm behavior. First, we find supporting evidence for the hypothesis by Malmendier and Tate (2005) that an overconfident CEO's investment decisions are sensitive to the availability of cash. They explain their findings in the following way: An overconfident CEO overestimates the profitability of investments relative to the market and regards company stock as undervalued. External financing, therefore, is perceived as unduly costly compared to internal financing. This induces an overconfident CEO to invest more when there are abundant internal funds. As in prior work (e.g. Thomas et al., 2010), without controlling for CEO overconfidence our European data shows no sign of the investment-cash flow sensitivity puzzle that has been well-documented in the US. However, ignoring CEO overconfidence in an investment-cash flow regression introduces an omitted variable bias under the premise that overconfidence affects both investment and cash flow. We illustrate this using a directed acyclic graph derived from the theoretical model of Malmendier and Tate (2005). Specifically, we find that an overconfident CEO invests 10 cents more per dollar of cash than her unbiased counterpart. These estimates are statistically indistinguishable from the results of Malmendier and Tate (2005).

Furthermore, we also find a positive link between CEO overconfidence and citation-weighted patent counts (e.g. Galasso and Simcoe, 2011; Hirshleifer et al., 2012), a typical indicator of innovation. Galasso and Simcoe (2011) explain their results with a career-concern model. Caring about her outside option on the labor market, a CEO invests in innovation to signal high ability, which is associated with successful innovation. An overconfident CEO, who overestimates her ability and hence the likelihood of successful innovation, is willing to invest at a higher cost of innovation. This induces higher investment levels, which directly translates into more innovation. Using the empirical model of Galasso and Simcoe (2011), we find a positive total effect of overconfidence on patent citations. The total effect captures both the direct effect of a change in the effectiveness of investments and the indirect effect of a change in R&D spending. An overconfident CEO significantly increases citations by 30%. Controlling for R&D stock, we still find a positive effect on patent citations, but the effect is insignificant at the 10% level. In this case, the coefficient can be interpreted as a change in patent citations per Euro of R&D stock. Under the assumption that the patent production functions of overconfident and non-overconfident CEOs are concave and do not intersect, these two results imply a higher R&D productivity of overconfident CEOs. In both specifications, we cannot reject the null that the estimated coefficients are similar to those from Galasso and Simcoe (2011).

Finally, we find different effects of overconfidence on CEO compensation packages in Europe than in the US. Otto (2014) predicts theoretically and finds empirically that overconfident CEOs earn fewer incentive payments and obtain a lower total compensation. An overconfident CEO overestimates the value of incentive payments, due to her optimistic beliefs about the market's future evaluation of the firm. Under the assumption that the CEO's risk aversion is not outweighed by her overconfidence, a lower level of incentive payments is sufficient to induce incentive-compatible behavior and participation of the CEO at unchanged levels of fixed payments. We find that an overconfident CEO receives 68.4% more stock options, 41.6% fewer stocks, 6.1% more bonus payments, 14.6% less in base salary, and 3.5% lower compensation in total. Only the negative effects on stock options and base salary are significant at the 10% level and the remaining effects are not statistically different from zero. Nevertheless, we cannot reject the null hypothesis that the effect on total compensation is similar to the result of Otto (2014).

On the methodological side, we contribute to the literature by visualizing the identification strategies of Malmendier and Tate (2005), Galasso and Simcoe (2011), and Otto (2014) using causal graphs (Pearl, 2000). We derive a causal graph from the career-concern model in Galasso and Simcoe (2011) and discuss the critique that a failure to diversify might constitute a costly signal to potential investors. The signaling critique implies a causal relationship between the overconfidence measure and investor beliefs. Innovation activities, however, also affect investor beliefs about the firm, thus rendering investor beliefs a bad control (or collider in causal graph terms). Controlling for investor beliefs, then, breaks identification in the model of Galasso and Simcoe (2011), while the model is identified when we do not control for investor beliefs in a regression.

We find similar effects in Europe despite that institutional differences, in particular differences in corporate governance and corporate finance between Europe and the US, may affect the interplay of CEO overconfidence and firm behavior.[34] Corporate governance determines the strategic leeway of directors and thereby determines the limits to which personal traits of executives affect firm decisions (Hambrick, 2007). Crossland and Hambrick (2011) document substantial differences in corporate governance between countries. Further, the decision-making

[34]Due to missing exogenous variation, we cannot identify the effects of certain differences in institutional settings. In this paper, we merely document the effects of overconfidence on firm behavior in Europe and leave the question of specific causes of differences open for future research.

power of executives may be curtailed by investors. Investor control does vary with the form of financing and there are differences in the preferred form of financing between countries. In the UK, firms predominantly rely on banks as corporate creditors (Marshall et al., 2016).

The paper is structured as follows. The Sections 2.2 and 2.3 describe the classification approaches and the data sets. We analyze the effect of overconfidence on investment behavior, innovation, and compensation packages in Sections 2.4, 2.5, and 2.6. Section 2.7 concludes.

## 2.2 Classification

The seminal paper by Malmendier and Tate (2005) proposes an empirical strategy to infer overconfidence of CEOs from a combination of data on compensation and trading behavior. The key idea is that CEOs are typically highly exposed to company-specific risk. First, their compensation is often tied to firm success through conditional bonuses, stock or stock option payments. Second, the value of a CEO's human capital is tied to company success, as both success and failure will be attributed in no small part to the CEO's ability to manage the firm.[35] Since CEOs are legally prohibited from hedging against this risk, a risk-averse CEO should generally reduce exposure to company risk when such actions are available. Malmendier and Tate (2005) argue that a failure to do so can be most convincingly interpreted as overconfidence since an overconfident CEO overestimates the chances of success and thus underestimates company idiosyncratic risk.

One prominent action a CEO can take to reduce her exposure to company risk is exercising her stock options. A risk-averse CEO should thus exercise stock options as soon as they are vested, provided exercising the stock options is profitable beyond some threshold. Malmendier and Tate (2005) calibrate the threshold using the Hall and Murphy (2002) framework, assuming a constant relative risk aversion parameter of three and that two-thirds of the CEO's wealth is invested in the company. A CEO is then indifferent between exercising and holding a stock option that is 67% in-the-money.[36] A CEO who fails at least twice to exercise a stock option that is 67% in-the-money in the vesting year is assumed to be overconfident. The resulting overconfidence measure is called *Holder67*.

The second approach to classifying CEOs as overconfident, which also exploits the failure to diversify, is to consider late exercises of stock options. Malmendier and Tate (2005) posit that a CEO, who fails to exercise a stock option until its expiration year, holds the belief that the company's stock will appreciate. They call a CEO who fails to exercise a stock option until the expiration year a *Longholder*.

The third approach does not rely on the failure to diversify. Instead, it builds on deliberate behavior that increases exposure to company risk instead of decreasing it. Malmendier and

[35] Galasso and Simcoe (2011) even argue that the CEO's human capital may be firm-specific in no small part.
[36] This means that the current stock price is 67% higher than the options exercise or strike price.

Tate (2005) consider long-time CEOs whose tenure lasts at least ten years in their sample. They classify a CEO as overconfident if she net-purchases company stock in the first five years of her tenure and only use the remaining years, which are not used to classify the CEO, in their estimation to guard against endogeneity problems. Malmendier and Tate (2005) call this measure *Net-Buyer*. Our data also allows us to construct all three approaches to measuring overconfidence, *Holder67*, *Longholder*, and *Net-Buyer*.

Using our European sample, we use all three approaches to measuring overconfidence, *Holder67*, *Longholder*, and *Net-Buyer*. While we get sufficiently many observations using the option-based approaches *Holder67* and *Longholder*, the original definition of *Net-Buyer* seems too demanding for our data. Therefore, we propose a modified *Net-Buyer* dummy, where a CEO is classified as overconfident when she net-purchases company stock in more years during her tenure than she net-sells. We conservatively calculate net-purchases since we exclude stock grants from purchases and consider a balance of purchases and sales not to be a net-purchase. We only classify CEOs whose tenure is five or more years in our sample.[37] We follow the focus on option-based measures in the existing literature, but report results based on the *Net-Buyer* classification in the Appendix.

A strength of our European sample is that it is not restricted to CEOs. This allows us to observe managers at different stages of their professional lives. We propose a novel classification approach to distinguish between two types of overconfidence based on the fact that we observe managers before and during their tenure as CEO.[38]

### 2.2.1 Overconfidence and Optimism

Overconfidence is a collective term for several closely related psychological biases[39]. Moore and Healy (2008) define one variant of overconfidence as the "overestimation of one's actual ability, performance, level of control, or chance of success."[40] This definition conflates two conceptually different yet intimately linked and economically relevant versions of overconfidence. On the one hand, the definition alludes to the ability to affect the likelihood of success and the control over that likelihood. On the other hand, it mentions that the chance of success could be overestimated independently of ability or control. We refer to the former as *overconfidence* and to the latter as *optimism*. The extant literature often does not distinguish between these two versions of overconfidence when interpreting the canonical overconfidence measures.

We propose a novel approach to disentangle overconfidence from optimism to distinguish between the biases in the analysis. The key idea is that the overestimation of own ability

[37] The 5-year requirement corresponds to the 25th percentile of CEO tenure in our data for those CEOs that we can classify as *Net-Buyer*.

[38] We also observe CEOs after their tenure if they remain in a position discharging managerial duty (e.g. on the supervisory board). The number of observations is small, so we do not exploit this interesting variation.

[39] See Moore and Healy (2008) for an overview.

[40] The literature has also used earnings forecasts and surveys to measure overprecision, i.e., optimistic beliefs about the certainty of events. Our data does not permit these approaches, so we omit a discussion here.

(overconfidence) relies on being in charge, while an optimism bias is independent of the power of being in charge. Accordingly, a manager who overestimates her own ability is observationally equivalent to an unbiased manager when they have no influence over the outcome. An *optimist*, however, will believe that the probability of success of a decision she has no control over is higher than it actually is, leading her to act differently from an unbiased but otherwise identical manager.

One source of variation in the control a manager has over firm outcomes stems from the fact that we observe CEOs and other C-suite executives alike. CEOs have substantial influence on top-level decisions as they set the agenda and possess the de-facto decision-making power in the firm. In comparison the influence a single C-suite executive has over the firm's overall performance is rather limited compared to the influence of the CEO.

An overconfident CEO erroneously believes that her own ability is higher than it actually is. However, during her appointment as a regular C-suite executive, her control over the firm's performance is indirect and limited as the CEO sets the agenda. Thus, an overconfident C-suite executive does not net-purchase company stock.[41] With the promotion to CEO, however, she gains the control over firm outcomes she was lacking as a regular C-suite executive. Thus, we expect an overconfident CEO to net-purchase company stock during her tenure as CEO. We use this pattern to call a CEO overconfident, who does not net-purchase company stock during her time as C-suite executive, but starts net-purchasing during her tenure as CEO.

**Classification 2.1** (Overconfidence)**.** *An eventual CEO is classified as overconfident if she does not net-purchase company stock during her time as C-suite executive, but does net-purchase company stock during her tenure as CEO.*

Our optimism classification uses the same general idea. An optimist believes that the probability of success is higher than it actually is, irrespective of her own ability and level of control over the firm's performance. Thus, an optimist will view the firm's prospects too favorably, even during her time as a C-suite executive. Thus, we expect an optimistic CEO to net-purchase company stock during her tenure as C-suite executive *and* her tenure as CEO.

**Classification 2.2** (Optimism)**.** *An eventual CEO is classified as an optimist if she net-purchases company stock during her time as C-suite executive and also during her tenure as CEO.*

Note that the two measures are mutually exclusive, as an optimist can not also be classified as overconfident and vice versa. Further, the data requirements for the two measures are identical, so whenever we are able to classify a CEO as overconfident (or not overconfident), we can also

[41]We implicitly assume that an overconfident C-suite member does not have optimistic beliefs about the likelihood of being promoted to CEO. We think that this is a reasonable assumption as CEO successions are typically highly debated in the financial press.

apply the optimism classification. Thus, we can include both dummy variables in our regression to compare optimistic and overconfident CEOs to the left-out category of unbiased CEOs.

The classification idea for overconfidence and optimism relies on an increase in control over firm outcomes due to the ascension to CEO. It is quite plausible and the literature also finds that CEOs have a strong influence (e.g., Malmendier and Tate, 2005; Bertrand and Schoar, 2003). The extent of CEO influence may, however, vary from company to company or even from CEO to CEO.[42] For example, Adams et al. (2005) find that firms run by a powerful CEO have higher variance in their stock returns.[43]

Table 6: Correlation Matrix.

| | Optimism | Overconfidence | Longholder | Holder67 |
|---|---|---|---|---|
| Optimism | 1 | | | |
| Overconfidence | -0.111 | 1 | | |
| Longholder | 0.191 | -0.038 | 1 | |
| Holder67 | 0.007 | 0.03 | -0.123 | 1 |

Note: Correlation coefficients are rounded to the third decimal.

Table 6 reports correlations between these measures. As expected, *optimism* and *overconfidence* have a negative correlation. The *optimist* correlates positively with *Longholder* and *Holder67*, while the *overconfident* correlates negatively with the *Longholder* dummy. Interestingly, the *Longholder* dummy correlates negatively with the *Holder67*, even though one might expect a positive correlation. The magnitudes of all of these correlations are in line with the correlations reported in Malmendier and Tate (2005).

### 2.3 Data

Our data set is compiled from four sources. The first source is the BoardEx Core Reports Europe and UK that provide information on director compensation packages, characteristics, and positions. We construct our panel from this data. We complement the compensation data with stock prices downloaded from Yahoo! Finance. Besides compensation packages, we require information on director transactions with company stocks or stock options, so-called directors' dealings. Falling under disclosure rules, we gather this data from the websites of the financial authorities in France, Germany, and the UK using web-scraping techniques. Finally, we use data

[42] Similar to the modified *Net-Buyer*, we restrict attention to CEOs with at least four years of tenure for the classification of optimism and overconfidence. This corresponds to the 25th percentile for the CEOs that we observe as executives in the same firm before becoming CEO.

[43] Oracle comes to mind as a counterexample to the assumption that CEOs have higher control than other executives. At Oracle, founder and Chairman of the Board Larry Ellison officially has the title of Chief Technical Officer, rather than Chief Executive Officer. To the extent that we did not identify a distinct CEO for a given company year we manually looked up who was CEO and assigned a CEO dummy to the plausibly powerful manager. We manually classified 33 individuals as CEO.

on firm financials from Bureau van Dijk's Amadeus database. Appendix B.2 provides details on the data preparation.

### 2.3.1 BoardEx

BoardEx is a commercial data provider of business network and business leader information. Their database comprises detailed information on board members, C-suite executives, senior leaders, and professional advisors. The data is gathered from publicly available sources such as regulatory filings, annual reports, proxy statements, company websites, press, and regulatory newswires. We use their Core Reports data for Europe and the UK.

The basic structure of our data set, a panel of CEOs within companies, is constructed from start and end dates of role descriptions indicating that the person is the principal executive instance of the company.[44] The data further contains information on director characteristics such as date of birth, gender, and education[45] as well as previous and other current positions within the company. We also construct board size measures from this data. Specifically, we construct an efficient board size dummy which takes the value one when the executive and the supervisory board together comprise more than four but less than twelve individuals. The main value of the data lies in the detailed information on compensation packages, including salaries, bonuses, stock grants, option grants, and other cash payments such as relocation costs and fringe benefits. To construct the option-based overconfidence measures we use (besides information on directors' dealings) information on strike prices, vesting dates, and expiry dates of option packages. The data also contains information on total stock and option holdings. Table 7 provides summary statistics.

### 2.3.2 Directors' Dealings

In 2014, the European Commission unified the financial transparency rules of the member states in the Market Abuse Regulation No. 596/2014. Following Article 19, persons discharging managerial responsibilities, meaning top-level management and individuals exerting supervisory functions, are obligated to publish transactions with financial instruments linked to their company via the responsible authorities. National transparency rules have been effective since the early 2000s.[46]

We gathered data on directors' dealings from the publication websites of the national

[44]We infer the principal executive instance from job titles and manually verify the ambiguous cases.

[45]We only have access to information on the type of degree and the institution attended, but no further information on the field of studies.

[46]The predating regulations are article L. 621-18-2 of the code monétaire et financier (France), §15a of the Wertpapierhandelsgesetz (Germany), and section 96(A) of the Financial Services and Markets Act 2000 (UK). Since the Directive 2003/6/EC of the European Commission, transparency rules are similar across member states.

Table 7: CEO Summary Statistics.

| Variable | N | Mean | Median | SD | Min. | Max. |
|---|---|---|---|---|---|---|
| Holder67: No, Number of Firms = 148 | | | | | | |
| Age | 377 | 53 | 53 | 7.2 | 32 | 76 |
| Tenure | 360 | 5.8 | 4 | 5.5 | 1 | 32 |
| Gender | 377 | | | | | |
| ... F | 20 | 5% | | | | |
| ... M | 357 | 95% | | | | |
| Total Compensation (€T) | 371 | 2093 | 1060 | 2709 | 19 | 21324 |
| Salary (€T) | 371 | 527 | 412 | 409 | 0 | 2200 |
| Bonus (€T) | 371 | 468 | 178 | 722 | 0 | 4946 |
| Stocks (€T) | 371 | 753 | 0 | 1748 | 0 | 16551 |
| Options (€T) | 371 | 273 | 0 | 865 | 0 | 6880 |
| Other (€T) | 371 | 2.8 | 3 | 1.7 | 0 | 8.2 |
| Stock Ownership | 377 | 0.016 | 0.0021 | 0.032 | 0.000063 | 0.17 |
| Vested Options | 377 | 0.0016 | 0 | 0.0044 | 0 | 0.028 |
| Holder67: Yes, Number of Firms = 355 | | | | | | |
| Age | 1801 | 54 | 54 | 7.2 | 29 | 80 |
| Tenure | 1728 | 8.3 | 7 | 5.4 | 1 | 34 |
| Gender | 1801 | | | | | |
| ... F | 67 | 4% | | | | |
| ... M | 1734 | 96% | | | | |
| Total Compensation (€T) | 1666 | 1228 | 637 | 1929 | 0 | 39190 |
| Salary (€T) | 1666 | 407 | 335 | 278 | 0 | 2646 |
| Bonus (€T) | 1666 | 242 | 71 | 382 | 0 | 3416 |
| Stocks (€T) | 1666 | 345 | 0 | 810 | 0 | 9017 |
| Options (€T) | 1666 | 170 | 0 | 1085 | 0 | 29514 |
| Other (€T) | 1666 | 2.5 | 2.8 | 1.7 | 0 | 9.4 |
| Stock Ownership | 1801 | 0.022 | 0.0063 | 0.035 | 0.000064 | 0.18 |
| Vested Options | 1801 | 0.0091 | 0.0017 | 0.017 | 0 | 0.2 |

*Notes:* This table reports summary statistics of CEO characteristics and compensation packages split by overconfidence status (Holder67). The unit of observation is CEO-year combinations, and we restrict the sample to observations that enter at least one of the regressions of Sections 2.4, 2.5, or 2.6. Tenure is the number of years an individual served as CEO within the company. Total compensation is the sum of the individual components of the compensation package listed below. Stocks is the value of shares awarded as evaluated at the closing stock price of the annual report date. Options is the value of stock options grants as evaluated by a generalized Black-Scholes option pricing model. Other is annual ad hoc cash payments such as relocation costs and fringe benefits. Stock ownership is the share of company stocks held by the CEO at the beginning of the year, normalized by market capitalization. Vested options are the number of options held by the CEO that are exercisable within the first six months of the year relative to the total shares outstanding.

authorities in the beginning of 2021 and mid of 2022 (see Table 8).[47,48] Each financial authority maintains a database of notifications of information with disclosure requirements. These databases can be accessed via a search interface on their website. The search results provide URLs to websites, each providing detailed information on single transactions of directors. The information is either provided directly on the website or in a PDF file which can be downloaded. In the first step, we download all websites or PDF files that document transactions. This means that we observe *all* directors' dealings that have been filed with and published by the national financial authorities. In total, we recover 158.142 files. In the second step, we extract the required information from the HTML or the PDF files. We recover information on the company name, director name, director position, transaction date, publication date, prices, quantities, type of financial instrument, and type of transaction.

The information extraction process is entirely automated for the files obtained from the French and the German authorities. This is possible because individual dealings have been reported consistently in a few distinct formats even before the unification of reporting standards in 2014. In the case of the UK, the extraction process is challenging as reporting has not been standardized before the EU regulation in 2014. We manually extract the information for these cases. To reduce the workload[49] (24.015 files from investegate.co.uk and 15.613 files from the FCA that could not be extracted automatically) we only consider files for which we can match the corresponding company to our remaining data sources and for which we can find the surname of any CEO of the company during our sample period in the document.[50]

We manually categorize the type of financial instrument into "stocks", "options" and "other"

[47]We implement the web-scraping in `python`. In the case of the websites of the French and the German authorities we mainly rely on the packages `Beautiful Soup`, `Selenium`, `Requests`, `pandas`, and `PyPDF2`. We switched to the `Scrapy` framework to download the websites of the financial authority of the UK, but still relied on the aforementioned packages to extract information from the websites.

[48]Note that we rely on an external source (investegate.co.uk) to obtain data on dealings in the UK for the years of 2008 to 2013. The reporting system of the FCA underwent a technical change in 2013. As a result, only directors' dealings after 2013 are directly accessible via their current search engine. The FCA provides a list of disclosures and websites to access the information before 2013, but the information on directors' dealings seems to be incomplete. We, therefore, use data obtained from investegate.co.uk for the years 2008 to 2013. We contacted the FCA to verify that investegate.co.uk is an official data provider of regulatory information.

[49]Another difficulty of the UK publications is that the HTML code of the standard form is not standardized either, rendering it difficult to automate information extraction. To mitigate the risk of potential extraction errors, we check the results for completeness of the essential information. In particular, we check that the executor name, the company name, the type of transaction, the type of financial instrument, the price, and the quantities are available. We added any missing information manually for incomplete cases for which we can match the company and find the surname of any CEO in the file.

[50]We consider documents in which the surname of any CEO of the associated company is a substring of or has a Damerau-Levenshtein string distance smaller than one to any word in the document. With this approach, we mitigate the possibility of missing relevant documents due to typing errors. For details on the company matching procedure see below.

Table 8: Meta Data on Scraped Directors' Dealings.

| **Country** | **Authority** | **Website** | **Date** | **# Docs.** | **Years** |
|---|---|---|---|---|---|
| FR | AMF | bdif.amf-france.org | 28.01.2021 | 58.459 | 2009-2021 |
| DE | BaFin | unternehmensregister.de | 30.04.2021 | 23.379 | 2011-2021 |
| UK | FCA | data.fca.org.uk | 15.09.2022 | 52.289 | 2013-2021 |
| UK | FCA | investegate.co.uk | 08.09.2022 | 24.015 | 2008-2013 |

and the type of transaction into "buy", "sell", "exercise" and "other". Further, we generate a company identifier and person-company identifier to harmonize differences in spellings of companies and individuals within companies.

Figure 5 provides an overview of the number of directors' dealings over time. Note that we manually extract all files before 2013 and some files between 2014 and 2016 in the case of the UK.[51] Because we only consider CEOs and companies that we can match to the remaining sources, the number of director dealings' is substantially lower for this period. After 2016, the graph shows transactions of all persons discharging managerial responsibilities and of individuals associated with unmatched companies.

### 2.3.3 Amadeus Data and Yahoo! Finance

We use data on firm financials from Bureau van Dijk's (BvD) Amadeus database. The Amadeus database relies on information retrieved from official annual accounts. Bureau van Dijk unifies the information provided in the financial statements to render comparisons across countries possible, despite differences in reporting standards. The patent data is obtained from the European Patent Office and is matched on the company-year level by BvD. We construct variables so that they resemble the variables of the papers from which we adapt our empirical models as closely as possible.[52] We follow the recommendations on preparing Amadeus data from Gopinath et al. (2017). For further details on the data preparation, see Appendix B.2.

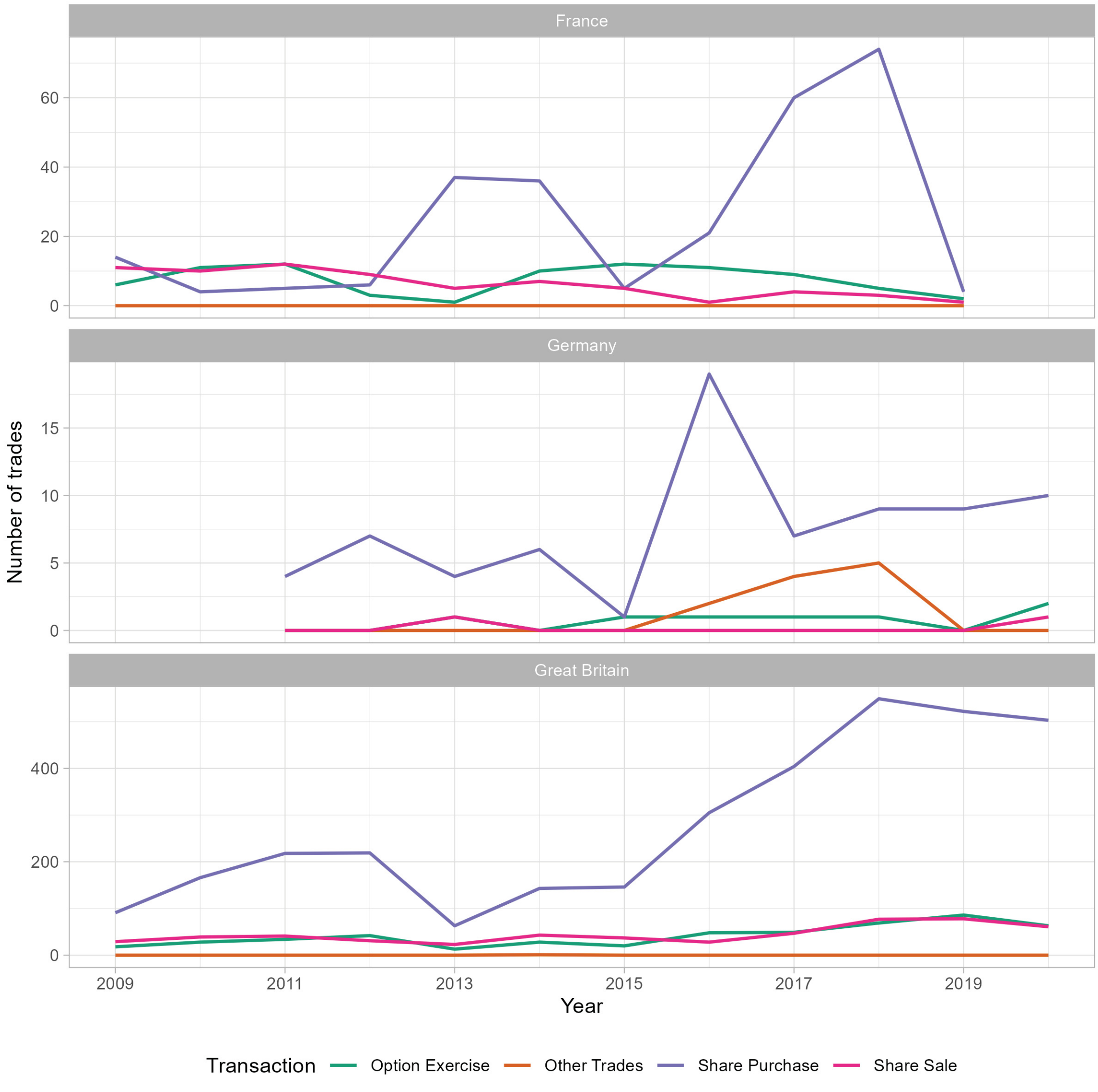
The variables used in the analyses below are defined as follows. Investment is the change in tangible fixed assets plus depreciation and amortization. Our measure of cash flow is cash flow as defined by BvD excluding extraordinary and other profits or losses. Both investment and cash flow are normalized by total assets.[53] Total assets will also be used as a measure of company size and operating revenue as a measure of sales. Leverage is defined as the ratio of long-term debt and total assets, and the market-to-book ratio is the sum of market capitalization and long-term debt divided by total assets. The capital-labor ratio is total assets divided by the number of employees. Information on Tobin's Q, market capitalization, earnings before interest and taxes (EBIT), and research and development expenditures are directly taken from the database. In Section 2.6 we normalize EBIT and R&D expenses by total assets. Further, we construct a R&D stock variable as in Hall (1990), who approximates the first period stock by the first period R&D expense divided by the sum of a depreciation rate of 15% and a R&D expenditure growth rate of 8%. Subsequent investments are depreciated over a 10-year period. We use SIC two-digit codes and Fama-French industry codes (constructed from SIC primary

[51] The unified publication standards were introduced in 2014 and have been mandatory since 2016. The new format has been taken up gradually by companies.

[52] Our empirical models are based on Malmendier and Tate (2005), Galasso and Simcoe (2011), and Otto (2014).

[53] Malmendier and Tate (2005) normalize investment and cash flow by capital in their preferred specification but also point out that a normalization by total assets does not change the results. In our sample, capital seems to be prone to outliers. We, therefore, normalize the variables by total assets.

Figure 5: Number of Directors' Dealings per year, by Nature of Transaction and Country.




*Source:* Scraped data about directors' dealings, own illustration.
*Notes:* This figure shows the number of directors' dealings over time split by type of financial instrument and reported separately for each country. Here we show only those dealings corresponding to any of the CEOs entering one of our main regressions.

codes) as industry classifications.

Our measure of innovation is citation-weighted patent counts, i.e. the sum of forward citations of all patents granted within one year divided by the average number of citations across companies within a year.[54] Further, we use the average of citation-weighted patent counts over the ten years preceding our sample as a control variable (as in Blundell et al., 1999). Table 9 provides summary statistics.

[54]We add one citation to each patent to differentiate between a patent with no citations and no patent at all. The normalization by the average number of citations within a year deals with the truncated nature of patent citations.

Table 9: Descriptive Statistics Amadeus Data.

*Panel A: Investment-Cash Flow Sample*

| Variable | N | Mean | Median | SD | Min. | Max. |
|---|---|---|---|---|---|---|
| Holder67: No, Number of Firms = 120 | | | | | | |
| Investment (€M) | 273 | 90 | 8.4 | 263 | -0.0079 | 2426 |
| Investment (Norm. Assets) | 273 | 0.07 | 0.051 | 0.068 | -0.0012 | 0.35 |
| Cash Flow (€M) | 273 | 142 | 20 | 416 | -2300 | 3875 |
| Cash Flow (Norm. Assets) | 273 | 0.078 | 0.094 | 0.14 | -0.49 | 0.3 |
| Total Assets (€M) | 273 | 1542 | 200 | 3733 | 3.2 | 32063 |
| Tobin's Q | 273 | 1.5 | 1.2 | 1.1 | 0.24 | 6.1 |
| Board Size | 273 | 8.1 | 8 | 2.5 | 3 | 17 |
| Executive Board Size | 273 | 3.1 | 3 | 1.2 | 1 | 8 |
| Supervisory Board Size | 273 | 5 | 5 | 2.5 | 1 | 16 |
| Efficient Board Size | 273 | 0.93 | 1 | 0.26 | 0 | 1 |
| Holder67: Yes, Number of Firms = 308 | | | | | | |
| Investment (€M) | 1319 | 56 | 4.4 | 229 | -88 | 3245 |
| Investment (Norm. Assets) | 1319 | 0.065 | 0.046 | 0.066 | -0.0039 | 0.34 |
| Cash Flow (€M) | 1319 | 95 | 11 | 286 | -352 | 3712 |
| Cash Flow (Norm. Assets) | 1319 | 0.059 | 0.092 | 0.14 | -0.5 | 0.3 |
| Total Assets (€M) | 1319 | 1142 | 129 | 3103 | 2.6 | 28934 |
| Tobin's Q | 1319 | 1.5 | 1.2 | 1.1 | 0.23 | 6.1 |
| Board Size | 1319 | 7.5 | 7 | 2.3 | 2 | 18 |
| Executive Board Size | 1319 | 3 | 3 | 1.2 | 1 | 9 |
| Supervisory Board Size | 1319 | 4.5 | 4 | 2 | 0 | 15 |
| Efficient Board Size | 1319 | 0.96 | 1 | 0.2 | 0 | 1 |

*Panel B: Innovation Sample*

| Variable | N | Mean | Median | SD | Min. | Max. |
|---|---|---|---|---|---|---|
| Holder67: No, Number of Firms = 55 | | | | | | |
| Total Citations | 157 | 103 | 0 | 297 | 0 | 1783 |
| Total Citations (Demeaned) | 157 | 13 | 0 | 40 | 0 | 215 |
| RD Stock (€M) | 157 | 1272 | 30 | 3526 | 0.5 | 21053 |
| Capital/Labor | 157 | 605 | 259 | 2200 | 34 | 21492 |
| Sales (€M) | 157 | 8918 | 363 | 15274 | 0.21 | 58770 |
| Holder67: Yes, Number of Firms = 136 | | | | | | |
| Total Citations | 645 | 6.5 | 0 | 67 | 0 | 874 |
| Total Citations (Demeaned) | 645 | 0.84 | 0 | 8.1 | 0 | 109 |
| RD Stock (€M) | 645 | 234 | 7.6 | 1862 | 0.064 | 24150 |
| Capital/Labor | 645 | 306 | 223 | 331 | 16 | 4317 |
| Sales (€M) | 645 | 2518 | 35 | 8356 | 0.0073 | 47297 |

Table 9 – continued from previous page

| Panel C: Compensation Sample | | | | | | |
|---|---|---|---|---|---|---|
| Variable | N | Mean | Median | SD | Min. | Max. |
| Longholder: No, Number of Firms = 164 | | | | | | |
| Market Capitalization (€M) | 977 | 6.5 | 0.54 | 16 | 0.0025 | 147 |
| Leverage | 977 | 0.13 | 0.1 | 0.14 | 0 | 0.76 |
| Market-to-Book Ratio | 977 | 1.7 | 1.3 | 1.6 | 0.17 | 21 |
| Cash Flow (€M) | 977 | 741 | 47 | 1988 | -232 | 21208 |
| Cash Flow (Norm. Assets) | 977 | 0.11 | 0.11 | 0.12 | -0.87 | 0.95 |
| RD Expenses (€M) | 977 | 73 | 0 | 241 | 0 | 1799 |
| EBIT (€M) | 977 | 618 | 39 | 1819 | -7951 | 23371 |
| EBIT (Norm. Assets) | 977 | 0.072 | 0.08 | 0.14 | -1.2 | 0.64 |
| Stock Return | 977 | 0.14 | 0.14 | 0.35 | -1.5 | 2.2 |
| SD Return | 977 | 0.1 | 0.088 | 0.068 | 0 | 0.73 |
| Board Size | 977 | 9.2 | 8 | 3.7 | 3 | 24 |
| Supervisory Board Size | 977 | 6.3 | 5 | 3.8 | 1 | 22 |
| Executive Board Size | 977 | 3 | 3 | 1.4 | 1 | 9 |
| Longholder: Yes, Number of Firms = 336 | | | | | | |
| Market Capitalization (€M) | 1683 | 3.2 | 0.12 | 12 | 0.00063 | 130 |
| Leverage | 1683 | 0.13 | 0.067 | 0.31 | 0 | 10 |
| Market-to-Book Ratio | 1683 | 1.7 | 1.1 | 1.6 | 0.11 | 15 |
| Cash Flow (€M) | 1683 | 415 | 7.1 | 1956 | -2300 | 43491 |
| Cash Flow (Norm. Assets) | 1683 | 0.023 | 0.081 | 0.32 | -4.9 | 1.5 |
| RD Expenses (€M) | 1683 | 37 | 0 | 394 | 0 | 5894 |
| EBIT (€M) | 1683 | 278 | 4 | 1259 | -5716 | 17077 |
| EBIT (Norm. Assets) | 1683 | -0.032 | 0.049 | 0.31 | -4.9 | 1 |
| Stock Return | 1683 | 0.091 | 0.082 | 0.59 | -2.4 | 11 |
| SD Return | 1683 | 11 | 0.12 | 179 | 0.027 | 3000 |
| Board Size | 1683 | 7.9 | 7 | 3 | 3 | 32 |
| Supervisory Board Size | 1683 | 4.9 | 4 | 2.9 | 0 | 25 |
| Executive Board Size | 1683 | 3 | 3 | 1.2 | 1 | 9 |

*Notes:* This table reports summary statistics of company financials and characteristics split by overconfidence status (Holder67). The unit of observation is company-year combinations, and we report the statistics for each of the samples used in section Sections 2.4, 2.5, and 2.6 separately. Investment is the change in tangible fixed assets plus depreciation and amortization. Cash flow is profit after tax plus depreciation and amortization. Tobin's Q is market capitalization by total assets, and board size is the sum of executive and supervisory board members. Efficient board size is a dummy that takes the value one when the board size is between four and twelve. Total citations are the sum of all patent citations of a company in a year. Sales is operating revenue. Leverage is long-term debt divided by total assets. Market capitalization is calculated as the product of the number of outstanding shares and the end-of-year market price. The market-to-book ratio is the ratio of market capitalization and total assets. Stock return is yearly stock returns, and the standard deviation of stock returns is calculated based on monthly returns over the previous five years.

We obtain daily stock prices from Yahoo! Finance's API.[55] Stock prices are required to calculate the intrinsic value of option grants. Further, we use yearly stock returns as well as the standard deviation of monthly returns over the previous five years as control variables.

### 2.3.4 Merging of Data Sets

We need to match the BoardEx Core Reports to the Amadeus database on the company level and the BoardEx Core Reports and the directors' dealings on the company-person level. Both matching procedures consist of several steps, where each successive step is only performed on the set of remaining unmatched instances.

[55] BoardEx and BvD provide stock tickers for each company. We use this information to access the API with the `python` package `yfinance`.

The match between BoardEx Core Reports and the Amadeus database follows a three-step procedure. In the first step, companies are matched based on ISIN numbers.[56] In the second step, companies are matched, if they have the same company name.[57] For the remaining companies, we construct the best[58] fuzzy match based on the entire company name and the best fuzzy match between two single words of two company names. We select the correct fuzzy matches manually.

In the case of the match between the BoardEx Core Reports and the directors' dealings, we first match companies and then directors within a company. The company matching procedure is the same as with the Amadeus data. Directors are matched in a five-step procedure. We start with an exact match on the entire name[59] and proceed with an exact match on the first and the last name[60]. In the next two steps we fuzzy match on the entire name and on the first and the last name. We only consider fuzzy matches with a string distance smaller than four and select the best match in both cases. We manually verify that the fuzzy matches are correct. Finally, we compare (within a matched company) the remaining unmatched individuals of both data sources and match missed instances by hand.

### 2.4 Investment Cash Flow Sensitivity and Overconfidence

Under complete financial markets, an increase in cash flow should not increase a firm's investment since the investment decision is driven by the profitability of the investment projects and outside financing is available (Modigliani and Miller, 1958). One reason the irrelevance of the financing structure may break down is a disagreement between investors and managers caused by managerial overconfidence. An overconfident manager may perceive her investment projects as more profitable than they are and will thus view outside financing as too costly since she believes the firm is undervalued by rational investors. Given this disagreement between investors and managers, an increase in cash flow thus allows the overconfident manager to finance additional projects without relying on outside financing. In this section we follow the approach of Malmendier and Tate (2005) to investigate the effect of overconfidence on investment. We adopt the Pearl causal graph framework (see e.g. Pearl, 2000; Cunningham, 2021) to discuss the empirical strategy. The next paragraph provides a brief introduction to the conceptual framework.

In a causal graph (or directed acyclic graph, DAG) a node is a variable and an arrow represents a causal relationship between two variables. The causal graph is meant to depict all relevant causal relationships and provides a graphical representation of these causal assumptions.

[56]Note that the ISIN is not always available in both data sources.

[57]Before matching, we clean the company names, e.g. by removing corporate form abbreviations.

[58]The criterion for the best match is the smallest Damerau-Levenshtein string distance.

[59]Some directors delegate their finances to a wealth management firm. We identify wealth management firms and manually assign the corresponding director.

[60]Similarly to the company names, we clean director names before the matching procedure, e.g. we remove titles. Further, we identify a first and a last name for each individual.

The mathematical theory of causality that underpins this graphical approach (Pearl, 2000) facilitates automated non-parametric identification proofs (Hünermund and Bareinboim, 2023).

In a causal graph, a collection of arrows forms a path. The fundamental building block of any path are chains, $X \to Y \to Z$, forks, $X \leftarrow Y \to Z$, and inverted forks, $X \to Y \leftarrow Z$. A preceding node is referred to as a parent and a following node to as a child. A chain has all arrows in one direction and is called a causal path, as it is usually the effect of interest. Forks and inverted forks are nuisances to identification as they allow spurious correlation between the treatment, $X$, and the outcome, $Z$. They are referred to as backdoor-paths. It turns out that $X$ and $Z$ are independent in a fork conditional on $Y$, so controlling for $Y$ in a regression breaks the spurious correlation. This contrasts to an inverted fork, where $Y$ is called a collider as both arrows "collide" at $Y$. Here, $X$ and $Z$ are independent unconditionally, but conditioning on $Y$ makes $X$ and $Z$ dependent, thus introducing spurious correlation. In the economics literature collider bias is known as "bad control" problem (see Cinelli et al. (2022) for a treatment of so-called "bad controls" and collider bias).[61]

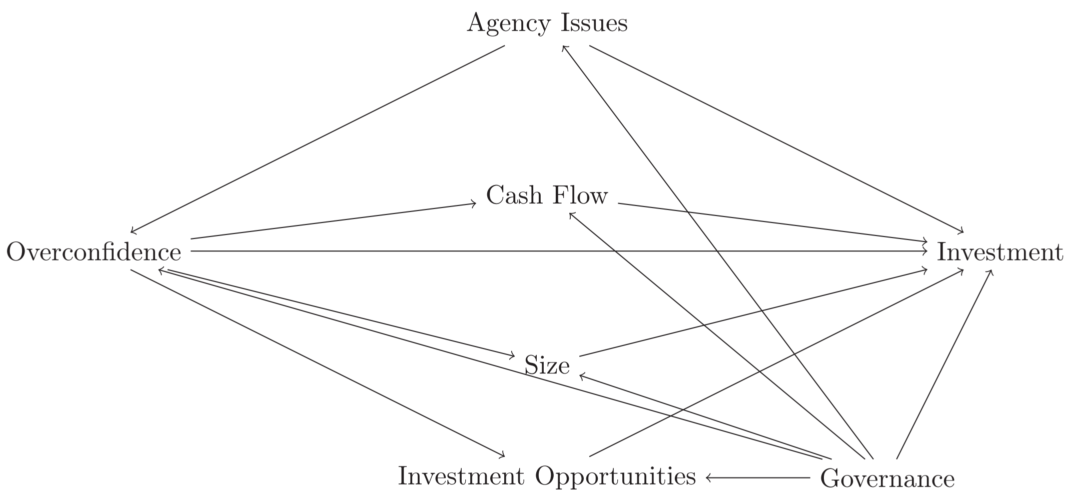
Figure 6 shows a DAG derived from the discussion of the causal relationships and the theoretical model in Malmendier and Tate (2005). Investment decisions are driven by the trade-off between the profitability of the investment opportunity and the cost of investment. The causal graph also allows firm size, cash flow, and agency issues to impact the investment decision directly.[62] Note that size and cash flow are generated by yesterday's investment, but this dynamic structure is de-emphasized by Malmendier and Tate (2005).

Corporate governance is assumed to have a direct impact on investment as well as indirect effects through the other variables. In particular, good corporate governance may be able to

[61] While the DAG in this section does not have a collider problem, the DAG in Section 2.5 on overconfidence and innovation does exhibit a collider.

[62] While Modigliani and Miller (1958) suggests cash flow should not matter, it is an empirical puzzle that it does seem to matter, at least in the US. The arrow from cash flow to investment allows for other explanations of investment cash flow sensitivity, including non-optimal firm behavior.

Figure 6: DAG derived from Malmendier and Tate (2005).



alleviate agency issues arising from misaligned incentives between the manager and the owner of the firm.

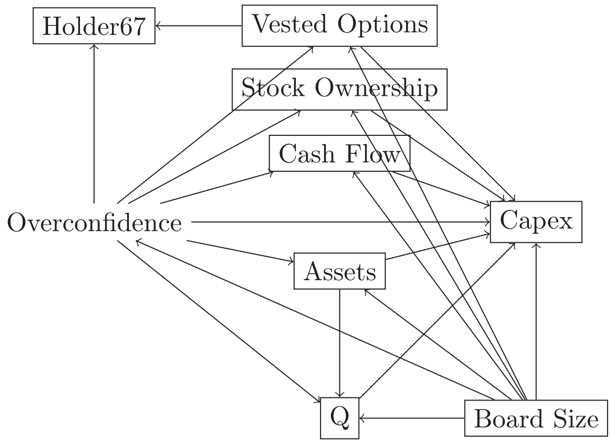
Overconfidence impacts the investment decision directly and indirectly, through cash flow, size, and investment opportunities. First, overconfidence could lead to different values in cash flow or investment opportunities. Because overconfident CEOs overestimate the probability of success, there is also a direct effect of overconfidence on investment, $Overconfidence \rightarrow Investment$. Overconfidence of the chosen CEO may be caused by agency issues and corporate governance. For example, an owner may select an overconfident CEO if she is worried an unbiased and risk-averse CEO would under-invest. Corporate governance may put restrictions on what a CEO can do and thus also restricts how effectively a given CEO overconfidence bias may play out.

Of course, many of these variables are rather conceptual in nature, and not directly measurable. In Panel a) of Figure 7 we replace the conceptual variables with their measured counterparts. Investment is measured by capital expenditure, size by total assets, and investment opportunities by Tobin's Q, which is the market value of assets divided by the book value of assets. This necessitates the introduction of an additional arrow, since Tobin's Q is directly computed from assets, $Assets \rightarrow Q$. Economic theory suggests that manager and owner incentives can be aligned by letting managers hold a share in equity through their compensation packages. Following Malmendier and Tate (2005), we operationalize this idea using managerial stock ownership (as a percentage of the company) and the number of vested options held by the managers to control for incentive alignment.

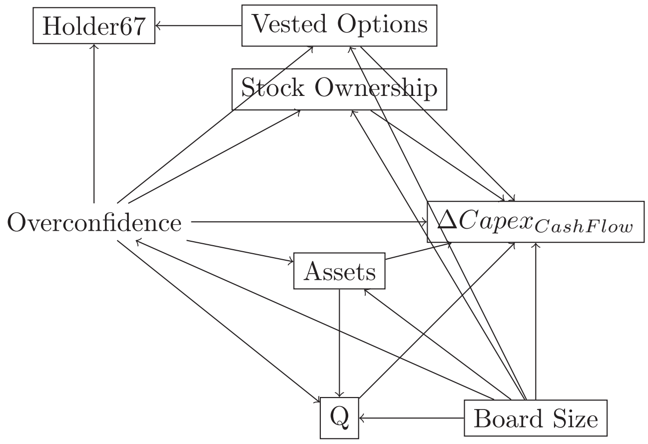
Importantly, the effect of interest is not the direct effect of overconfidence on capital expenditures. Rather, it is the interaction of cash flow and overconfidence: overconfident CEOs invest more when there is abundant cash for internal financing. Since DAGs are non-parametric they capture any interaction effects present in the data. To explicitly incorporate a parametric interaction effect, we can introduce $\Delta Capex_{CashFlow}$ to denote the causal effect of the interaction of CEO overconfidence and cash flow on capital expenditures. An arrow from $Holder67$ to $\Delta Capex_{CashFlow}$ then translates into the familiar interaction term in a linear regression model (Nilsson et al., 2021). Such a DAG is shown in Figure 7. Thus, the causal path of interest is $Overconfidence \rightarrow \Delta Capex_{CashFlow}$.

Overconfidence is, of course, not observed directly, but rather inferred from non-exercising of profitable options. Thus, the *Holder67* dummy is caused by overconfidence and vested options in tandem, $Overconfidence \rightarrow Holder67 \leftarrow Vested\ Options$. By controlling for vested options, stock ownership, number of independent directors, assets, and Tobin's Q the only path left open is $Holder67 \leftarrow Overconfidence \rightarrow \Delta Capex_{CashFlow}$. This means we indeed identify the effect of the *Holder67* classification on cash flow sensitivity of investment. The caveat is that *Holder67* is not a perfect measurement of overconfidence, but a proxy, and thus our estimate of the effect of overconfidence will only be as good as the proxy captures overconfidence in the first place.

Figure 7: DAG derived from Malmendier and Tate (2005).




(a) Measured Variables




(b) Interaction DAG

Panel (a) replaces unobserved variables in Figure 6 with measured variables. Panel (b) is the Interaction DAG that introduces the causal effect of Cash Flow on Capital Expenditure as $\Delta Capex_{CashFlow}$. This allows to depict the interaction effect of interest by the arrow $Overconfidence \rightarrow \Delta Capex_{CashFlow}$.

Note that random noise only affects the precision but not the consistency of the estimates. However, if there is a systematic misclassification, then an omitted variable could introduce a bias. Consider the following example: A CEO avoids exercising stock options to signal confident beliefs about the future firm performance to the market to reduce investment financing costs.[63] In this case, we would classify some truly non-overconfident CEOs as overconfident. Those CEOs, however, do not have to rely on internal financing as they can finance their projects via the market. The investment levels of those misclassified CEOs, therefore, should not be sensitive to the availability of cash. This means the estimate based on the imperfect CEO confidence measures provides a lower bound to the true effect. A potentially systematic misclassification, thus, only introduces an unfavorable bias when an omitted variable is correlated qualitatively in

[63] Malmendier and Tate (2005) explicitly state that stock-option exercises are not considered as a signal in the financial press.

the same way with the misclassification as with the outcome.

Following the analysis of the causal graph we estimate Equation 10, where $I_{it}$ denotes investment, $Q_{it-1}$ market value over book assets lagged, $CF_{it}$ cash flow, $\Delta_{it}$ the overconfidence dummy, $FIRM'_{it}$ a set of firm control variables, and $CEO_{it}$ a set of CEO control variables. The set of control variables comprises the share of stock ownership, the number of vested options, the size of the firm, and a dummy for efficient board size[64]. Further, the regression models may include year-fixed effects and firm-fixed effects. The effect of interest is the interaction of cash flow and the overconfidence dummy, *Holder67*.

$$\begin{aligned} I_{it} = \; & \alpha + \beta_1 CF_{it} + \beta_2 Q_{it-1} + \beta_3 CF_{it} \times Q_{it-1} + \delta_1 Holder67_{it} + \delta_2 CF_{it} \times Holder67_{it} \\ & + FIRM'_{it}\gamma_1 + CEO'_{it}\gamma_2 + CF_{it} \times FIRM'_{it}\gamma_3 + CF_{it} \times CEO'_{it}\gamma_4 \\ & + \tau_t + CF_{it} \times \tau_t + \lambda_i + \epsilon_{it} \end{aligned} \tag{10}$$

Table 10 presents our results and compares them to the results by Malmendier and Tate (2005).[65] In contrast to the US literature cash flow has a negative, but insignificant effect on investment in our data. This is in line with previous research on cash flow sensitivity in European countries. For example, Thomas et al. (2010) find no cash flow sensitivity for so-called *outsider* economies, in contrast to a substantial effect for *insider* economies. Countries with outsider economies have "large stock markets, dispersed ownership, strong outside investor rights, high disclosure level, and strong legal enforcement" (Thomas et al., 2010, p. 148). In their taxonomy the UK is an outsider economy, for which they find a null effect. One may be worried that the null effect would preclude a further analysis of overconfidence driving the cash flow sensitivity pattern. If overconfidence is indeed a causal factor behind investment cash flow sensitivity, not including overconfidence in the regression, however, amounts to omitted variable bias. This is also a point nicely illustrated with our DAG in Figure 6, where the backdoor path $Cash\ Flow \leftarrow Overconfidence \rightarrow Investment$ is left open by not controlling for overconfidence and thus breaks identification of the direct effect $Cash\ Flow \rightarrow Investment$.

The effect of interest is the interaction of cash flow and *Holder67*, which we analyse in Columns 2 and 3 of Table 10. The interaction term tells us whether overconfident CEOs invest more when cash flow is available to internally finance investment. We find a statistically significant positive effect of 0.096 and 0.101, depending on whether we include the interaction between industry fixed effects and cash flow. This means an overconfident CEO invests roughly 10 cents more than her non-overconfident counterpart when there is 1 Euro more in cash flow.

[64]Malmendier and Tate (2005) use the number of directors who are CEO at another company, which, unfortunately, is unavailable in our data. We use a dummy that takes the value one when the board size is between four and twelve.

[65]Following Malmendier and Tate (2005) we use trimming to deal with outliers in the variables investment, cash flow, Tobin's Q, size, stock ownership, and vested options. We trim at the 4% level.

Table 10: Investment Cash Flow Sensitivity and Overconfidence

| Dep. Var. | Investment | | | | | |
|---|---|---|---|---|---|---|
| | *Results: GSZ* | | | *Results: MT05* | | |
| Model: | (1) | (2) | (3) | (4) | (5) | (6) |
| *Variables* | | | | | | |
| Holder67 | | -0.014* | -0.014** | | -0.050* | -0.036 |
| | | (0.007) | (0.007) | | (0.026) | (0.029) |
| Cash Fl. × Holder67 | | 0.074* | 0.076* | | 0.234** | 0.172** |
| | | (0.042) | (0.044) | | (0.090) | (0.078) |
| Cash Fl. | -0.256 | -0.312* | -0.279 | 1.658*** | 1.704*** | 1.291*** |
| | (0.190) | (0.187) | (0.191) | (0.168) | (0.570) | (0.401) |
| Lag Q | 0.006 | 0.006 | 0.006* | - 0.005 | -0.009 | -0.0111 |
| | (0.004) | (0.004) | (0.004) | (0.002) | (0.049) | (0.032) |
| Stocks | 0.080 | 0.083 | 0.089 | -0.108 | -0.183 | 0.189 |
| | (0.114) | (0.115) | (0.115) | (0.567) | (0.705) | (0.556) |
| Options | -0.284 | -0.183 | -0.210 | 0.195 | 0.140 | 0.199 |
| | (0.213) | (0.208) | (0.219) | (0.120) | (0.134) | (0.128) |
| Size | -0.022*** | -0.022*** | -0.022*** | 0.047** | 0.054 | 0.043 |
| | (0.007) | (0.007) | (0.008) | (0.019) | (0.037) | (0.030) |
| Eff. Board Size | 0.012 | 0.011 | 0.011 | -0.004 | -0.007 | -0.013 |
| | (0.015) | (0.015) | (0.015) | (0.008) | (0.009) | (0.009) |
| Cash Fl. × Lag Q | -0.003 | -0.005 | -0.004 | 0.052*** | 0.065 | 0.065 |
| | (0.011) | (0.011) | (0.011) | (0.020) | (0.078) | (0.050) |
| Cash Fl. × Stocks | -0.886 | -0.872 | -0.969 | -0.575 | -0.690 | -1.114 |
| | (0.674) | (0.669) | (0.692) | (0.417) | (1.533) | (1.148) |
| Cash Fl. × Options | 0.654 | -0.089 | 0.048 | -0.461*** | -0.298 | -0.502*** |
| | (1.314) | (1.280) | (1.320) | (0.111) | (0.226) | (0.191) |
| Cash Fl. × Size | 0.029* | 0.029* | 0.027 | -0.171*** | -0.175** | -0.143*** |
| | (0.016) | (0.016) | (0.019) | (0.020) | (0.076) | (0.054) |
| Cash Fl. × Eff. Board Size | 0.097 | 0.109* | 0.110* | 0.036** | 0.044* | 0.059** |
| | (0.065) | (0.064) | (0.066) | (0.017) | (0.0261) | (0.023) |
| *Fixed-effects* | | | | | | |
| Firm (342) | Yes | Yes | Yes | Yes | Yes | Yes |
| Year (9) | Yes | Yes | Yes | Yes | Yes | Yes |
| Year × Cash Fl. | Yes | Yes | Yes | Yes | Yes | Yes |
| Industry × Cash Fl. | No | No | Yes | No | No | Yes |
| *Fit statistics* | | | | | | |
| Observations | 1,592 | 1,592 | 1,592 | 1,058 | 1058 | 1056 |
| Adjusted $R^2$ | 0.448 | 0.450 | 0.448 | 0.61 | 0.62 | 0.67 |

*Notes:* The dependent variable in all columns is investments, defined as the change in tangible fixed assets plus depreciation and amortization and normalized by lagged total assets. Cash flow is profit after tax plus depreciation and amortization, normalized by lagged total assets. Tobin's Q is lagged market capitalization by total assets. Stock ownership is the share of company stocks held by the CEO at the beginning of the year, normalized by market capitalization. Vested options are the number of options held by the CEO that are exercisable within the first six months of the year, normalized by the number of outstanding shares. Size is logged and lagged total assets. Efficient board size is a dummy that takes the value one when the board size is between four and twelve. Industry fixed effects are based on 12 Fama–French industry groups. *Signif. Codes:* ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered on the firm-level.

These estimates are close to the ones presented by Malmendier and Tate (2005), which are 0.234 and 0.172, respectively. The null hypothesis of identical estimates can not be rejected at conventional significance levels.[66] Thus, we cannot reject that overconfidence has the same effect on investment cash flow sensitivity in Europe.

## 2.5 Overconfidence and Innovation

The existing literature finds a positive link between CEO overconfidence and innovative activity measured in terms of patent citations (e.g. Galasso and Simcoe, 2011; Hirshleifer et al., 2012). The underlying mechanism is intuitive. Innovation activities are risky endeavors. The likelihood of success, however, is affected by the ability of the CEO. The market, therefore, regards successful innovation as a positive signal about the CEO's ability. In a competitive labor market, where CEOs are compensated according to the market's perception of their ability, the CEO trades off this expected gain from innovation, in terms of an increase in compensation, with the cost of innovation. Receiving a higher expected payoff, a more able CEO is willing to bear a higher investment cost, which induces a higher probability of innovation. An overconfident CEO, who overestimates her ability, is also willing to invest at higher cost levels, leading to a similar increase in the probability of successful innovation.[67]

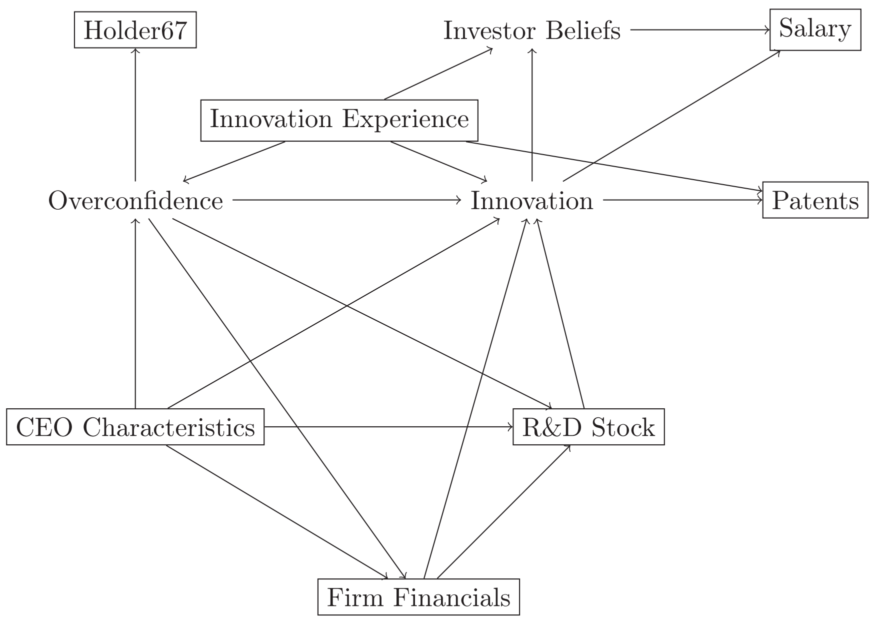
We adopt the empirical strategy from Galasso and Simcoe (2011) to estimate the effect of overconfidence on innovation and use the causal graph framework to clarify identification. Figure 8 visualizes their causal assumptions in a DAG. We are interested in quantifying the partial effect of overconfidence on patent citations along the causal path $Overconfidence \rightarrow Innovation \rightarrow Patents$. There are other partial effects one might want to consider to calculate a total effect of overconfidence, such as the path $Overconfidence \rightarrow R\&D\ Stock \rightarrow Innovation \rightarrow Patents$.

According to this causal model we need to control for CEO characteristics, firm financials, R&D expenditure, and innovation experience to estimate the causal impact of overconfidence on patents. CEO characteristics are operationalized by the socio-demographic variables age, tenure, number of vested options, and stock ownership of the CEO normalized by the number of shares outstanding of the company as in Galasso and Simcoe (2011). Further, we follow the authors in measuring the firm's innovation experience as the pre-sample mean of citation-weighted patents, a procedure introduced by Blundell et al. (1999). Pre-sample means are used instead of a firm fixed effects approach to relax the assumption of strict exogeneity that comes with firm fixed effects. It should be noted that this DAG is meant to capture the causal relationships within a single year and within a given sector. Thus, year and sector fixed effects are used in estimation.

[66] We run a z-test with the null that the estimated effects are identical. The test statistic is $\frac{\delta_2^{MT05}-\delta_2^{GSZ}}{\sqrt{SE^{MT05}+SE^{GSZ}}}$, which leads to Z values of 1.05 and 0.555. The corresponding p-values are 0.295 and 0.579.

[67] Galasso and Simcoe (2011) formalize this idea in a model that is based on the extension of the Holmstrom (1999) career-concern model by Aghion et al. (2013).

Figure 8: DAG derived from the economic theory in Galasso and Simcoe (2011).



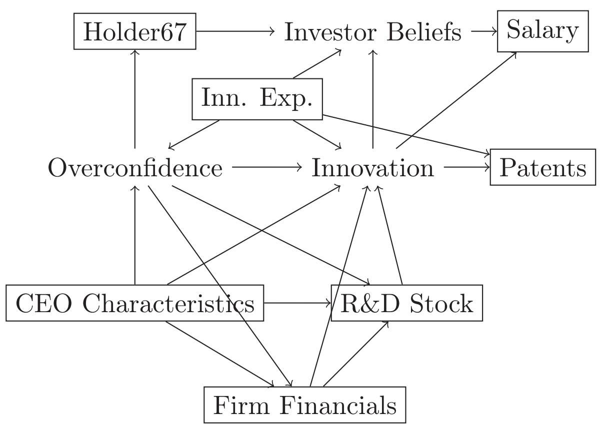

68

The DAG can also be used to debilitate the signaling critique described in Section 2.4. If the CEO's failure to diversify is not a consequence of overconfidence, but a deliberate action to signal a positive firm outlook to investors, then the *Holder67* measure has an effect on investor beliefs. Thus, the arrow $Holder67 \rightarrow Investor\ Beliefs$ is added in Figure 9. This makes investor beliefs a collider, $Holder67 \rightarrow Investor\ Beliefs \leftarrow Innovation$. Note that controlling for colliders introduces bias and researchers subscribing to the signaling critique should not control for investor beliefs according to the causal model in Figure 9.

Galasso and Simcoe (2011) employ a Poisson panel regression to test the empirical predictions. More specifically they use the following model

[68]The causal impact of year and sector would be represented by nodes that have no parents, i.e. no arrows point to them, but they are a parent to all other variables, i.e. an arrow points to all other nodes. It is therefore convenient to exclude them from the DAG for easier exposition.

Figure 9: The signaling critique in the DAG derived from the economic theory in Galasso and Simcoe (2011).



$$E[P_{it}] = exp(\delta Holder67_{it} + FIRM'_{it-1}\gamma_1 + CEO'_{it}\gamma_2 + \tau_t + \lambda_i) + \mu_i + \kappa_i \tag{11}$$

where $P_{it}$ is the citation-weighted patent count of firm $i$ at time $t$, $Holder67_{it}$ is the overconfidence measure, $FIRM_{it-1}$ is a vector of lagged firm control variables, $CEO_{it}$ is a vector of CEO control variables, $\lambda_i$ is the firm-level pre-sample mean of the outcome, $\tau_t$ is a time fixed effect, $\mu_i$ is an industry fixed effects, and $\kappa_i$ is a country fixed effect.

The regression results are summarized in Table 11. In column 1 we do not control for R&D stock. The coefficient, therefore, can be interpreted as a total effect, capturing both the direct effect of overconfidence on innovation and the indirect effect induced by a change in R&D spending. We find a positive and significant effect of overconfidence on citation-weighted patent counts. The coefficient of 0.265 suggests that an overconfident CEO obtains 30% more patent counts than her non-overconfident counterpart. Adding the control for R&D stock in column 2, the coefficient on overconfidence can be interpreted as the change in patent counts per Euro of investment. We still find a positive effect, but the effect is insignificant. Under the assumption that the patent production functions of overconfident and non-overconfident CEOs are concave and do not intersect, these two results together imply a higher R&D productivity of overconfident CEOs. Note that we cannot reject the hypothesis that the coefficients are identical to those of Galasso and Simcoe (2011) for both specifications.[69]

## 2.6 Compensation and Overconfidence

Standard contract theory suggests aligning incentives between owners and managers using equity, for example, stock options. These pay components are of uncertain value in the future, which the manager might impact through her effort in managing the company. An overconfident manager overestimates her ability to increase the likelihood of good company performance. This benefits the owner of the firm since incentivizing managerial effort is cheaper. In this section, we revisit the analysis of Otto (2014). We also expand the distinction between overestimation of one's ability (overconfidence) and optimism using our classification approach laid out in Section 2.2.

Otto (2014) provides a simple example of a more involved, dynamic model of executive compensation with CEO overconfidence. The core idea is that at the optimum the incentive constraint (and participation constraint) must be binding, under the assumption that overconfidence does not outweigh risk aversion.[70] Thus, the rational principal is able to properly incentivize managerial effort provision using fewer stock options, since the manager overvalues

[69] We run a z-test with the null hypothesis that the estimated effects are identical. The test statistic is $\frac{\delta^{GS11}-\delta^{GSZ}}{\sqrt{SE^{GS11}+SE^{GSZ}}}$, which leads to Z values of 1.63 and 1.61. The corresponding p-values are 0.104 and 0.106.

[70] If overconfidence outweighs risk aversion the effect of overconfidence on compensation ceases to be monotone negative. In this case, the participation constraint may be slack at the optimum and the extremely overconfident agent may receive *more* incentive claims.

Table 11: Innovation and Overconfidence

| Dep. Var. | Total Citations (Adjusted) | | | |
|---|---|---|---|---|
| | *Results: GSZ* | | *Results: GS11* | |
| Model: | (1) | (2) | (3) | (4) |
| *Variables* | | | | |
| Holder67 | 0.265*** | 0.157 | 0.543*** | 0.407*** |
| | (0.098) | (0.109) | (0.14) | (0.11) |
| ln(Lag Sales+1) | 0.863*** | -0.142 | 0.410*** | 0.054 |
| | (0.266) | (0.106) | (0.12) | (0.13) |
| ln(Lag C/L+1) | 0.283 | 0.517*** | 0.115 | 0.301** |
| | (0.204) | (0.128) | (0.23) | (0.15) |
| Stocks | -24.908 | -13.162*** | -2.107 | -2.455 |
| | (20.040) | (3.406) | (3.08) | (2.89) |
| Options | 0.976 | -0.789 | -0.789 | -0.458 |
| | (0.662) | (0.920) | (2.56) | (1.56) |
| ln(Lag RD Stock+1) | | 0.913*** | | 0.492*** |
| | | (0.249) | | (0.08) |
| CEO Controls | Yes | Yes | Yes | Yes |
| *Fixed-effects* | | | | |
| Country (2) | Yes | Yes | No | No |
| Industry (18) | Yes | Yes | Yes | Yes |
| Year (9) | Yes | Yes | Yes | Yes |
| Firm effects (BGV) | Yes | Yes | Yes | Yes |
| *Fit statistics* | | | | |
| Observations | 671 | 671 | 1,512 | 1,512 |
| Pseudo $R^2$ | 0.923 | 0.941 | NA | NA |

*Notes:* The dependent variable in all columns is citation-weighted patent counts, normalized by the mean number of citations across firms in this year. We add the value one before taking logs. Sales is logged and lagged operating revenue. The capital labor ratio is computed as the log of the difference between lagged total assets and lagged number of employees. Stock ownership is the share of company stocks held by the CEO at the beginning of the year, normalized by market capitalization. Vested options are the number of options held by the CEO that are exercisable within the first six months of the year, normalized by the number of outstanding shares. R&D Stock are constructed following Hall (1990). For details, see Section 1.3. We use a set of CEO controls, such as age, tenure, and their respective quadratic terms. Industry fixed effects are based on SIC two-digit industry codes. BGV (Blundell et al., 1999) firm effects are average citation-weighted patent counts over the ten years preceding our sample as a control variable. *Signif. Codes:* ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered on the industry-level.

this uncertain payoff component. Under a binding participation constraint, this overvaluation of incentive pay even leads to lower total pay.[71]

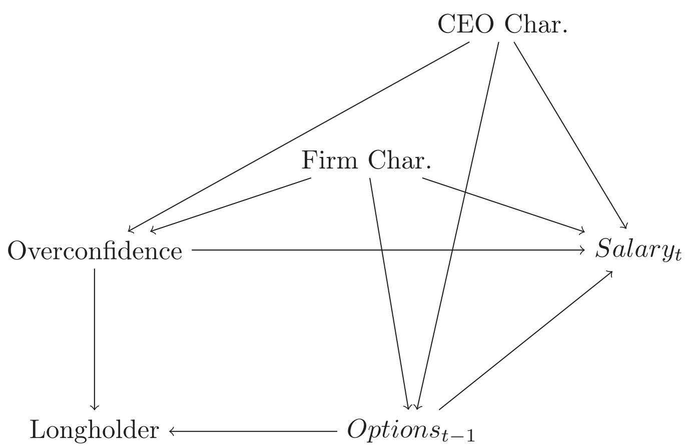
Depending on the specific firm decision managerial overconfidence may be value-increasing (see e.g. Section 2.5) or value-destroying (see e.g. Section 2.4). Thus, firms with different characteristics may have different overconfidence target levels when hiring a new CEO. For example, an R&D intensive firm might have a high or low target for managerial overconfidence, depending on whether good and bad projects can be easily distinguished and depending on how much discretion has to be left to the CEO. Consequently, it is important to control for firm characteristics, such as market capitalization, stock returns, the standard deviation of stock returns, leverage, market-to-book ratio, cash flow, R&D expenses, board size, and EBIT. It is similarly useful to control for CEO characteristics such as age, gender, and tenure if these are antecedents of overconfidence and compensation.[72]

Figure 10 shows a causal graph based on these arguments. CEO and firm characteristics are causes of the wage paid and managerial overconfidence. They are also causes of whether the CEO earned stock options in prior years. This is an important point since only those CEOs who have earned options in the past can become a *Longholder*. Since the sample is restricted to those who can be classified as overconfident (or not) in the sense of the *Longholder* dummy, it is implicitly conditioned on past option payment. According to the causal graph, the effect of interest is identified if overconfidence could be measured and all relevant firm and CEO characteristics are included in the regression. Since only the *Longholder* dummy is available as a proxy for overconfidence, identification rests on the assumption that this proxy works well.

We follow the approach by Otto (2014) and run the regression in Equation 12. Firm effects include market capitalization, stock returns, the standard deviation of stock returns, leverage, market-to-book ratio, cash flow, R&D expenses, board size, and EBIT. CEO characteristics

[71]For a survey on contract theory with behavioral agents see Koszegi (2014).
[72]If they are unrelated to overconfidence they may still increase the statistical power of the regression.

Figure 10: DAG derived from Otto (2014).



are age, gender, and tenure, which are included as fixed effects to allow non-linear effects. Additionally, year and firm fixed effects are included, so that estimates are calculated from within firm variation. In particular, this means the coefficient for *Longholder* is identified from firms who change from a non-*Longholder* to a *Longholder* CEO (or vice versa).

$$ln(Y_{it} + 1) = \alpha + \delta Longholder_{it} + FIRM'_{it}\gamma_1 + CEO'_{it}\gamma_2 + \tau_t + \lambda_i + \epsilon_{it} \quad (12)$$

Table 12 reports our regression results in Panel A and provides the results of Otto (2014) in Panel B. We find that a *Longholder* earns restricted stock grants that are roughly 41.7% smaller. This effect is significant at the 10% level (p-value of 0.08). In contrast to Otto (2014) we find a null effect on total compensation. This may be due to the positive effect of overconfidence on stock options packages, which are roughly 68% larger for *Longholder* CEOs, although this effect is not statistically significant at conventional levels (p-value of 0.13). Similarly, Otto reports that bonuses are significantly lower, whereas we find a null effect. Further, we find that a *Longholder* earns roughly 14.6% less in salary. Otto reports a similar effect, where a *Longholder* earns roughly 19.2% less in salary, although this time not statistically significant in his sample, but in ours. We reject the null that our estimates are the same as in Otto (2014) for each of the variable pay components as we find qualitatively opposite results with positive (negative) estimates for options and bonus (stocks). We do not reject the null for salary and total compensation.[73]

[73]The p-values of the z-test are 0.0076, 0.068, 0.028, 0.72, and 0.14 for Options, Stocks, Bonus, Salary, and Total, respectively.

Table 12: Compensation and Overconfidence

| Dep. Var. | ln(Options+1) | ln(Stocks+1) | ln(Bonus+1) | ln(Salary+1) | ln(Total+1) |
|---|---|---|---|---|---|
| *Panel A: Results GSZ* | | | | | |
| Longholder | 0.514 | -0.543* | 0.056 | -0.151* | -0.036 |
| | (0.345) | (0.313) | (0.275) | (0.082) | (0.075) |
| ln(Mkt Cap) | 0.460*** | 0.143 | 0.737*** | 0.110** | 0.310*** |
| | (0.130) | (0.118) | (0.121) | (0.051) | (0.052) |
| SD Return | 0.023** | 0.003 | -0.005 | -0.009*** | -0.007*** |
| | (0.011) | (0.011) | (0.009) | (0.003) | (0.003) |
| Leverage | -0.022 | -0.041 | -0.185 | -0.401*** | -0.379*** |
| | (0.134) | (0.135) | (0.190) | (0.099) | (0.105) |
| MtB Ratio | 0.005 | -0.005 | 0.007 | -0.051** | -0.064** |
| | (0.066) | (0.059) | (0.054) | (0.023) | (0.029) |
| Cash Fl. | -0.525* | 0.088 | 0.427* | 0.051 | -0.063 |
| | (0.267) | (0.134) | (0.249) | (0.144) | (0.128) |
| RD Expenses | 0.115 | -0.427 | 1.615** | 0.245 | 0.233 |
| | (1.101) | (0.702) | (0.690) | (0.335) | (0.389) |
| RD Missing | 0.111 | 0.044 | 0.457* | -0.022 | 0.040 |
| | (0.262) | (0.151) | (0.243) | (0.098) | (0.069) |
| Ex. Board Size | -0.082 | -0.074 | -0.106* | -0.026 | -0.037 |
| | (0.091) | (0.081) | (0.064) | (0.025) | (0.026) |
| EBIT | 0.086 | -0.099 | 0.072 | -0.085 | 0.019 |
| | (0.333) | (0.192) | (0.245) | (0.196) | (0.179) |
| Stock Return | 0.038 | 0.060 | 0.262** | 0.018 | 0.090* |
| | (0.106) | (0.061) | (0.110) | (0.034) | (0.049) |
| *Fixed-effects* | | | | | |
| Firm (445) | Yes | Yes | Yes | Yes | Yes |
| Year (9) | Yes | Yes | Yes | Yes | Yes |
| Age (47) | Yes | Yes | Yes | Yes | Yes |
| Tenure (35) | Yes | Yes | Yes | Yes | Yes |
| Gender (2) | Yes | Yes | Yes | Yes | Yes |
| *Fit statistics* | | | | | |
| Observations | 2,660 | 2,660 | 2,660 | 2,660 | 2,660 |
| Adjusted $R^2$ | 0.312 | 0.733 | 0.605 | 0.590 | 0.756 |
| *Panel B: Results Otto 2014* | | | | | |
| Longholder | -0.752** | 0.521 | -0.783*** | -0.213 | -0.201** |
| | (0.330) | (0.492) | (0.268) | (0.133) | (0.085) |
| ln(Mkt Cap) | 0.635*** | 0.244** | 0.980*** | 0.073 | 0.457*** |
| | (0.099) | (0.100) | (0.100) | (0.045) | (0.027) |
| SD Return | -0.338 | -0.377 | 4.126*** | -0.585 | 0.287 |
| | (1.134) | (1.178) | (0.941) | (0.468) | (0.315) |
| Leverage | -0.588 | 0.274 | 0.216 | 0.087 | 0.019 |
| | (0.504) | (0.475) | (0.469) | (0.073) | (0.117) |
| MtB Ratio | 0.014 | -0.002 | -0.072*** | 0.000 | 0.005 |
| | (0.019) | (0.024) | (0.022) | (0.011) | (0.008) |
| Cash Fl. | -0.122 | 0.382 | 0.673 | -0.167 | 0.283 |
| | (0.743) | (0.673) | (0.641) | (0.171) | (0.200) |
| RD Expenses | 0.043 | 2.785* | 3.819** | 0.417 | 0.206 |
| | (1.213) | (1.423) | (1.860) | (0.333) | (0.403) |
| RD Missing | 0.309 | -0.340 | 0.324 | -0.169 | -0.005 |
| | (0.503) | (0.380) | (0.314) | (0.119) | (0.093) |
| Ex. Board Size | 0.034 | -0.032 | -0.064** | 0.013 | -0.001 |
| | (0.033) | (0.040) | (0.029) | (0.012) | (0.007) |
| Independent | 0.130 | 0.757 | -0.044 | 0.204* | 0.065 |
| | (0.428) | (0.481) | (0.365) | (0.109) | (0.103) |
| EBIT | -0.055 | 0.769 | 7.166*** | 0.327* | 0.423* |

Table 12 – continued from previous page

| Dep. Var. | ln(Options+1) | ln(Stocks+1) | ln(Bonus+1) | ln(Salary+1) | ln(Total+1) |
|---|---|---|---|---|---|
| | (0.821) | (0.630) | (0.842) | (0.199) | (0.251) |
| Stock Return | -0.096 | 0.069** | 0.185 | -0.007 | -0.047 |
| | (0.069) | (0.029) | (0.138) | (0.009) | (0.029) |
| *Fixed-effects* | | | | | |
| Firm (443) | Yes | Yes | Yes | Yes | Yes |
| Year (9) | Yes | Yes | Yes | Yes | Yes |
| Age (47) | Yes | Yes | Yes | Yes | Yes |
| Tenure (34) | Yes | Yes | Yes | Yes | Yes |
| Gender (2) | Yes | Yes | Yes | Yes | Yes |
| *Fit statistics* | | | | | |
| Observations | 5,777 | 5,777 | 5,777 | 5,777 | 5,777 |
| Adjusted $R^2$ | 0.482 | 0.582 | 0.533 | 0.672 | 0.766 |

*Notes:* The dependent variables are logged compensation components, indicated in the column headers. We add the value one before taking logs. Market capitalization is calculated as the product of the number of outstanding shares and the end-of-year market price. Standard deviation of stock returns is calculated based on monthly returns over the previous five years and is divided by 100. Leverage is long-term debt divided by total assets, Market-to-book ratio is the ratio of market capitalization and total assets. Cash flow is profit after tax plus depreciation and amortization, normalized by lagged total assets. R&D expenses are normalized by total assets. R&D Missing is a dummy indicating whether R&D expenses are missing. Executive board size is the number of executive directors on the board. EBIT is normalized by total assets. *Signif. Codes:* ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered on the firm-level.

We repeat the estimation exercise using the *overconfidence* and *optimism* classification. Table 13 reports the results. We find negative estimates for *optimists* throughout, though they are only statistically significant for options and total salary at conventional significant levels. For the *overconfident* CEOs we only find a significant effect for the coefficient on stock options, with most of the other estimates showing negative effects, although they are all statistically indistinguishable from zero. The effects on stock options are quite strong at -2.4 and -3.8, respectively, indicating a -90.9 percent lower compensation in options (and analogously, a -97.76 percent lower option grant for an *overconfident* compared to non-overconfident CEOs). We also find a statistically significant effect on total compensation for *optimists*, at -0.85. This corresponds to total compensation being -57 percent lower than that of their unbiased counterparts.

## 2.7 Conclusion

The empirical literature inspired by the classification idea of Malmendier and Tate (2005) provides evidence for the effects of CEO overconfidence on firm behavior in the US. Are there similar effects in other countries, despite differences in corporate governance and corporate finance that affect the strategic leeway of CEOs? In this paper, we provide evidence that CEO overconfidence has similar effects on investment behavior and innovative activity but differential effects on CEO compensation packages in Europe. We construct a novel data set on directors' dealings in France, Germany, and the UK that allows us, in combination with data on CEO compensation, to construct the canonical CEO overconfidence measures. In addition to CEOs,

Table 13: Compensation and Overconfidence - Overconfidence and Optimism

| Dep. Var.<br>Model: | ln(Options+1)<br>(1) | ln(Stocks+1)<br>(2) | ln(Bonus+1)<br>(3) | ln(Salary+1)<br>(4) | ln(Total+1)<br>(5) |
|---|---|---|---|---|---|
| *Variables* | | | | | |
| Overconfidence | -2.360*** | 0.184 | -0.980 | -0.201 | -0.321 |
| | (0.784) | (0.998) | (0.702) | (0.277) | (0.294) |
| Optimism | -3.729** | -1.774 | -0.517 | -0.460 | -0.827* |
| | (1.554) | (1.087) | (1.065) | (0.422) | (0.471) |
| ln(Mkt Cap) | 0.310* | 0.048 | 0.626*** | 0.213* | 0.370*** |
| | (0.180) | (0.286) | (0.223) | (0.115) | (0.125) |
| SD Return | 43.956 | 9.261 | 26.576 | 0.991 | 11.179 |
| | (28.948) | (24.925) | (27.976) | (12.876) | (12.010) |
| Leverage | -0.418 | -1.069 | -1.264** | 0.556* | 0.363 |
| | (0.367) | (0.775) | (0.587) | (0.317) | (0.394) |
| MtB Ratio | 0.163 | 0.254 | 0.154 | -0.099 | -0.025 |
| | (0.142) | (0.258) | (0.165) | (0.092) | (0.119) |
| Cash Fl. | 0.224 | 0.580 | 0.900** | 0.278 | 0.291 |
| | (0.357) | (0.444) | (0.385) | (0.326) | (0.371) |
| RD Expenses | 1.475 | 4.913 | 3.091 | -3.132 | -3.827 |
| | (3.897) | (3.404) | (5.035) | (3.124) | (4.599) |
| RD Missing | -0.718 | 0.569 | 1.156** | -0.111 | 0.343 |
| | (0.561) | (0.646) | (0.546) | (0.138) | (0.230) |
| Ex. Board Size | -0.075 | 0.004 | 0.044 | -0.029 | -0.013 |
| | (0.101) | (0.117) | (0.112) | (0.029) | (0.037) |
| EBIT | -0.326 | -0.153 | -0.065 | -0.249 | -0.429 |
| | (0.771) | (0.496) | (0.591) | (0.331) | (0.359) |
| Stock Return | 0.108 | 0.160 | 0.560*** | -0.035 | 0.033 |
| | (0.141) | (0.152) | (0.185) | (0.103) | (0.091) |
| *Fixed-effects* | | | | | |
| Firm (156) | Yes | Yes | Yes | Yes | Yes |
| Year (9) | Yes | Yes | Yes | Yes | Yes |
| Age (46) | Yes | Yes | Yes | Yes | Yes |
| Tenure (12) | Yes | Yes | Yes | Yes | Yes |
| Gender (2) | Yes | Yes | Yes | Yes | Yes |
| *Fit statistics* | | | | | |
| Observations | 836 | 836 | 836 | 836 | 836 |
| Adjusted $R^2$ | 0.300 | 0.730 | 0.678 | 0.705 | 0.754 |

*Notes:* The dependent variables are logged compensation components, indicated in the column headers. We add the value one before taking logs. Market capitalization is calculated as the product of the number of outstanding shares and the end-of-year market price. Standard deviation of stock returns is calculated based on monthly returns over the previous five years and is divided by 100. Leverage is long-term debt divided by total assets, Market-to-book ratio is the ratio of market capitalization and total assets. Cash flow is profit after tax plus depreciation and amortization, normalized by lagged total assets. R&D expenses are normalized by total assets. R&D Missing is a dummy indicating whether R&D expenses are missing. Executive board size is the number of executive directors on the board. EBIT is normalized by total assets. *Signif. Codes:* ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered on the firm-level.

we observe the trading behavior of all C-suite executives, the supervisory board, and top-level managers who are obligated to disclose transactions due to European transparency rules. As CEOs are typically hired within the firm, we can exploit this additional source of information and the change in the decision-making power of individuals promoted to CEO to construct novel measures of overconfidence and optimism. Moreover, we visualize identification strategies of the existing literature using directed acyclic graphs.

Even though we find no evidence for cash flow sensitivity in Europe (as in Thomas et al., 2010), we still find that the investment level of an overconfident CEO increases with cash-flow (Malmendier and Tate, 2005, as in). Specifically, we find that an overconfident CEO increases investment by 10 cents per Euro of available cash relative to their non-overconfident counterpart. Our estimates are not significantly different from those of Malmendier and Tate (2005).

We also find a positive effect of CEO overconfidence on innovation as measured by citation-weighted patent counts (similar to Galasso and Simcoe, 2011; Hirshleifer et al., 2012). In our data, an overconfident CEO is associated with 30% more patent citations. Controlling for R&D stock, we still find a positive effect, albeit insignificant. In both cases, we cannot reject the null that the estimates are similar to those of Galasso and Simcoe (2011). Under the premise that the patent production functions of an overconfident and non-overconfident CEO do not cross, the results imply that overconfident CEOs are (on average) better innovators.

Moreover, our data suggests that CEO overconfidence affects CEO compensation packages in Europe differently than in the US. Otto (2014) illustrates in a theoretical model that overconfident CEOs should receive less variable pay and less total compensation.[74] He finds that this negative effect on incentive pay goes through options and bonuses (negative, significant), but not through stocks (positive, insignificant). In our data, we find that there is no significant difference in total compensation between overconfidence and non-overconfident CEOs. However, we cannot reject the null hypothesis that our estimate is similar to the one of Otto (2014). Furthermore, we find that overconfident CEOs receive more options and higher bonuses (both insignificant) but lower salaries and fewer stocks (both significant). We also investigate the effects of overconfidence and optimism on CEO compensation packages. These results mirror the findings by Otto (2014) more closely than those based on the canonical *Longholder* measures. We find significantly lower stock option grants (significant) and total compensation (only significant for optimistic CEOs).

Our novel approach to classifying CEOs as overconfident and optimistic exploits the change in decision-making power after being promoted to CEO. We define overconfidence as the overestimation of one's own ability (similar to Malmendier and Tate, 2005) and optimism as having upward biased beliefs about firm success (independent of own ability). The canonical *Net-buyer* measure relies on the idea that only an overconfident CEO is willing to actively increase exposure to company-specific risk by net-buying company stock. We classify a CEO

[74]The results rely on the assumption that the CEO's overconfidence does not outweigh her risk aversion.

as *overconfident* when we classify her as a *Net-Buyer* only during her tenure as CEO, and as *optimistic*, when we classify her as a *Net-Buyer* before *and* during her tenure as CEO. The classification requires a panel over a long time horizon, as only CEO-years of CEOs that are appointed during the sample period can be used in the analysis. Although our data set ranges from 2008 to 2020, we (currently) lack the statistical power to conduct reliable analyses on the effect of optimism and overconfidence on investment behavior and innovative activity. We are looking forward to revisiting these analyses in a few years.

We believe that there are some promising applications of our classification approach, e.g., in the analysis of compensation structures or the firms' innovation strategies. It is, however, challenging to identify situations in which overconfidence and optimism have qualitatively different effects on firm behavior. Most decisions within firms are based on beliefs about some probability of success. An overconfident CEO's self-serving beliefs induce similarly upward biased beliefs about the probability of success as those of an optimistic CEO who overestimates the probability straight away. The behavior of both types is, therefore, indistinguishable in most situations. Promising candidates are situations where the CEO's ability or effort is of particular importance. Consider, for example, a situation in which a CEO can invest in low-risk-low-yield and high-risk-high-yield projects resembling incremental and breakthrough innovation. The latter requires in-depth industry knowledge or great management skill to be successful. An overconfident CEO, therefore, is more likely to invest in breakthrough innovation, whereas an optimist invests relatively more in incremental innovation. Both types, however, invest more than their unbiased counterpart. Another example is compensation structures. Overconfident CEOs have self-serving beliefs about their impact on firm performance. Their effort, thus, will react more sensitively to changes in incentive pay than that of an optimistic type. Under the assumption that the overconfidence bias and the optimism bias are of similar size, one can show that overconfident CEOs should receive more incentive pay than their optimistic counterparts. This, however, requires measures of the degree of the biases. In a similar spirit, one could investigate the change in compensation structures of overconfident and optimistic CEOs when being promoted to CEO. The compensation structure of an overconfident CEO should change more drastically after the promotion than that of an optimistic type. Our classification approach may, therefore, be useful to explain firm behavior in domains where the CEO's effort or ability is particularly important.

# References


Adams, R. B., Almeida, H., and Ferreira, D. (2005). Powerful CEOs and Their Impact on Corporate Performance. *Review of Financial Studies*, 18(4):1403–1432.

Aghion, P., van Reenen, J., and Zingales, L. (2013). Innovation and Institutional Ownership.

*American Economic Review*, 103(1):277–304.

Ahmed, A. S. and Duellman, S. (2013). Managerial Overconfidence and Accounting Conservatism. *Journal of Accounting Research*, 51(1):1–30.

Banerjee, S., Humphery-Jenner, M., and Nanda, V. (2015). Restraining Overconfident CEOs through Improved Governance: Evidence from the Sarbanes-Oxley Act. *Review of Financial Studies*, 28(10):2812–2858.

Ben-David, I., Graham, J., and Harvey, C. (2007). Managerial Overconfidence and Corporate Policies. NBER Working Paper 13711, National Bureau of Economic Research, Cambridge, MA.

Bertrand, M. and Schoar, A. (2003). Managing with Style: The Effect of Managers on Firm Policies. *Quarterly Journal of Economics*, 118(4):1169–1208.

Blundell, R., Griffith, R., and van Reenen, J. (1999). Market Share, Market Value and Innovation in a Panel of British Manufacturing Firms. *The Review of Economic Studies*, 66(3):529–554.

Bouwman, C. H. (2014). Managerial Optimism and Earnings Smoothing. *Journal of Banking and Finance*, 41(1):283–303.

Campbell, T. C. (2014). CEO Optimism and the Board's Choice of Successor. *Journal of Corporate Finance*, 29:495–510.

Campbell, T. C., Gallmeyer, M., Johnson, S. A., Rutherford, J., and Stanley, B. W. (2011). CEO Optimism and Forced Turnover. *Journal of Financial Economics*, 101(3):695–712.

Cinelli, C., Forney, A., and Pearl, J. (2022). A Crash Course in Good and Bad Controls. *Sociological methods & research*, page 004912412210995.

Crossland, C. and Hambrick, D. C. (2011). Differences in Managerial Discretion Across Countries: How Nation-Level Institutions Affect the Degree to Which CEOs Matter. *Strategic Management Journal*, 32(8):797–819.

Cunningham, S. (2021). *Causal Inference: The Mixtape*. Yale University Press, New Haven.

Deshmukh, S., Goel, A. M., and Howe, K. M. (2013). CEO Overconfidence and Dividend Policy. *Journal of Financial Intermediation*, 22(3):440–463.

Galasso, A. and Simcoe, T. S. (2011). CEO Overconfidence and Innovation. *Management Science*, 57(8):1469–1484.

Gopinath, G., Kalemli-Özcan, Ş., Karabarbounis, L., and Villegas-Sanchez, C. (2017). Capital Allocation and Productivity in South Europe. *Quarterly Journal of Economics*, 132(4):1915–1967.

Hackbarth, D. (2008). Managerial Traits and Capital Structure Decisions. *Journal of Financial and Quantitative Analysis*, 43(4):843–881.

Hall, B. (1990). The Manufacturing Sector Master File: 1959-1987. NBER Working Paper 3366, National Bureau of Economic Research, Cambridge, MA.

Hall, B. J. and Liebman, J. B. (1998). Are CEOs Really Paid Like Bureaucrats? *Quarterly Journal of Economics*, 113(3):653–691.

Hall, B. J. and Murphy, K. J. (2002). Stock Options for Undiversified Executives. *Journal of Accounting and Economics*, 33(1):3–42.

Hambrick, D. C. (2007). Upper Echelons Theory: An Update. *Academy of Management Review*, 32(2):334–343.

Hirshleifer, D., Low, A., and Teoh, S. H. (2012). Are Overconfident CEOs Better Innovators? *Journal of Finance*, 67(4):1457–1498.

Holmstrom, B. (1999). Managerial Incentive Problems: A Dynamic Perspective. *The Review of Economic Studies*, 66(1):169–182.

Hribar, P. and Yang, H. (2016). CEO Overconfidence and Management Forecasting. *Contemporary Accounting Research*, 33(1):204–227.

Humphery-Jenner, M., Lisic, L. L., Nanda, V., and Silveri, S. D. (2016). Executive Overconfidence and Compensation Structure. *Journal of Financial Economics*, 119(3):533–558.

Hünermund, P. and Bareinboim, E. (2023). Causal Inference and Data Fusion in Econometrics. *The Econometrics Journal*.

Kim, J.-B., Wang, Z., and Zhang, L. (2016). CEO Overconfidence and Stock Price Crash Risk. *Contemporary Accounting Research*, 33(4):1720–1749.

Koszegi, B. (2014). Behavioral Contract Theory. *Journal of Economic Literature*, 52(4):1075–1118.

Leng, J., Ozkan, A., Ozkan, N., and Trzeciakiewicz, A. (2021). CEO Overconfidence and the Probability of Corporate Failure: Evidence from the United Kingdom. *The European Journal of Finance*, 27(12):1–25.

Malmendier, U., Pezone, V., and Zheng, H. (2022). Managerial Duties and Managerial Biases. *Management Science*, 69(6):3157–3758.

Malmendier, U. and Tate, G. (2005). CEO Overconfidence and Corporate Investment. *Journal of Finance*, 60(6):2661–2700.

Malmendier, U. and Tate, G. (2008). Who Makes Acquisitions? CEO Overconfidence and the Market's Reaction. *Journal of Financial Economics*, 89(1):20–43.

Malmendier, U., Tate, G., and Yan, J. O. (2011). Overconfidence and Early-Life Experiences: The Effect of Managerial Traits on Corporate Financial Policies. *Journal of Finance*, 66(5):1687–1733.

Marshall, A., McCann, L., and McColgan, P. (2016). The Choice of Debt Source by UK Firms. *Journal of Business Finance & Accounting*, 43(5-6):729–764.

Modigliani, F. and Miller, M. H. (1958). The Cost of Capital, Corporation Finance and the Theory of Investment. *American Economic Review*, 48(3):261–297.

Moore, D. A. and Healy, P. J. (2008). The Trouble with Overconfidence. *Psychological Review*, 115(2):502–517.

Nilsson, A., Bonander, C., Strömberg, U., and Björk, J. (2021). A Directed Acyclic Graph for Interactions. *International Journal of Epidemiology*, 50(2):613–619.

Otto, C. A. (2014). CEO Optimism and Incentive Compensation. *Journal of Financial Economics*, 114(2):366–404.

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. Cambridge University Press, Cambridge, 2nd edition.

Schrand, C. M. and Zechman, S. L. (2012). Executive Overconfidence and the Slippery Slope to Financial Misreporting. *Journal of Accounting and Economics*, 53(1-2):311–329.

Thomas, S., Tahoun, A., and Walker, M. (2010). IFRS Adoption in Europe and Investment-Cash Flow Sensitivity: Outsider Versus Insider Economies. *The International Journal of Accounting*, 45(2):143–168.

Zavertiaeva, M. A., López-Iturriaga, F. J., and Kuminova, E. V. (2018). Better Innovators or More Innovators? Managerial Overconfidence and Corporate R&D. *Managerial and Decision Economics*, 39(4):447–461.

# 3 Paying for Optimism: Stock Options for Rank and File Employees

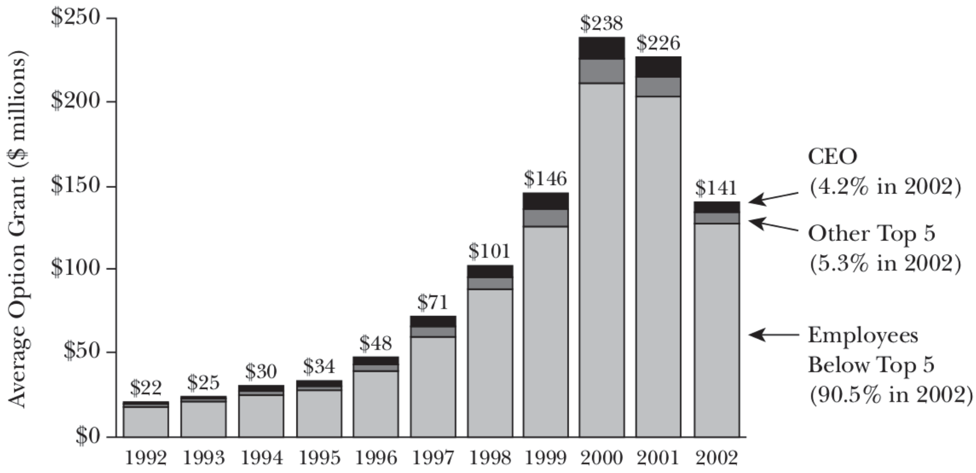
Simon Schulten

## 3.1 Introduction

Firms often pay stock options or share profits with rank and file employees. For example, Hall and Murphy (2003) find that roughly 90% of stock options granted by S&P 500 firms are paid out to employees who are not in the top management level (see Figure 11). At first glance, these contract features are in line with the large literature on CEO compensation (for a recent survey see Edmans and Gabaix, 2011). The idea underlying this literature is to align incentives between the CEO and shareholders by incentivizing CEO effort provision. It is, however, implausible that a rank and file employee has a unilateral impact on the stock price (or profits). More plausibly, one may assume that rank and file employees influence the stock price (profits) collectively. Such a collective impact, however, amounts to a public goods provision problem, which does not explain the use of stock options or unconditional profit sharing. Thus, traditional moral hazard models of effort provision are unable to explain these contract features for rank and file employees.

To explain stock options for rank and file employees, Section 3.2 introduces a model where the agent (the employee, he) has anticipatory utility in addition to material utility. In the model, the agent's action is to accumulate company specific skill, but, in contrast to the CEO setting, the agent's action does not increase the probability of the good state of the world. This is by design as each rank and file employee is aware that she has, by herself, minuscule impact on firm-wide outcomes. Further, I assume the principal (the firm, she) can observe the agent's

Figure 11: Grant-Date Values of Employee Stock Options in the S&P 500, reproduced from Hall and Murphy (2003)



company-specific skill level and can condition a bonus payment on skill level. I suppose the principal can only commit to paying the bonus if doing so is not overly costly. While this assumption is reduced form, such a skill bonus may be linked to a promotion, and promotable positions are linked to overall firm performance. Another reason for imperfect commitment might be overly short-termist management incentives, which leads managers to sacrifice bonus payments in the short-term despite profitability in the longer term (for models of short-termism see e.g. Bolton et al., 2006; Stein, 1988).

Company-specific skills are, by definition, not transferable to another firm. Consequently, accumulating company-specific skills does not increase the agent's outside option, while general skills do. Assuming a non-zero chance to sign with another company, the agent will prefer to invest in general rather than company-specific skills (holding incentives fixed). From the perspective of the principal, company-specific skills are more expensive to incentivize as the risk of skill-devaluing rehiring at another company needs to be compensated. Considering risk aversion makes the risk of skill-devaluing even more expensive, though I restrict attention to risk neutrality for ease of exposition. Furthermore, if general and company-specific skills have the same diminishing returns, the firm will optimally implement a lower company-specific skill level. A same sized decrease in the costs of incentivizing skills, then, boosts profits more for company-specific skills than for general skills. This is one reason why company-specific skills are regarded as valuable, but hard to incentivize.

In addition to the base salary and the skill dependent bonus payment the firm may grant the agent stock options that are independent of skill and have a higher value in the good state of the world. In the rational benchmark the sum of base salary and stock options is pinned down by the agent's participation constraint. As principal and agent share the probability of the good state as their common prior, there is no reason to grant options.[75] This mirrors the puzzle that stock options are used in practice despite the agent's inability to increase the stock price or the probability of the good state of the world.

I assume that the principal faces an agent who derives anticipatory utility in addition to material utility and chooses her beliefs optimally (Brunnermeier and Parker, 2005). Thus, the agent faces a tension between increasing her belief in the future and distorting her skill decisions. A higher belief in the good state (optimism) increases anticipatory utility. At the same time it causes costly over-investment in the skill from the agent's perspective, since she believes she is getting the bonus more frequently.

This tradeoff between anticipatory utility is mediated by the convex cost of the skill. A higher bonus implements a higher skill level, which makes over-investment in the skill more costly as it now occurs at a higher point of the convex cost function. From the principle's perspective this renders unconditional stock options useful as they enable him to reinforce the agent's tendency

[75]Assuming any level of risk aversion would pin down zero stock options in the rational benchmark.

to become optimistic without increasing the skill level directly. The same mechanism holds true for any variable pay component that is unconditional on skill and correlates with the bonus, such as profit participation or bonuses that condition on company wide targets rather than individual skill. Examples for such variable pay components beside stock options include car manufacturers sharing profits with assembly line workers[76] and long term incentive plans often combining bonuses with individual targets and company-, division- or team-level targets.

I derive the optimal contract the principal offers when facing a risk neutral optimal expectations agent.[77] Stock options play a role in the optimal contract if the agent's weight on anticipated utility is in an intermediate range. If, on the one hand, the agent puts too little weight on anticipatory utility it becomes too expensive to induce additional optimism through stock options. If, on the other hand, the agent puts too much weight on anticipatory utility she already believes in the good state with certainty, negating any scope to further induce optimism through stock options. In the intermediate range, however, the principal uses stock options to induce the agent to become fully optimistic. The agent's optimism creates a virtual rent, that the principal extracts through a combination of increased skill and lower base salary.

My main contribution is to the literature on stock options for rank and file employees, or broad-based stock options. Hall and Murphy (2003) review the literature on stock options in compensation practices. Managerial incentives to maximize firm value is one of the main arguments for using stock options in compensation packages, but other reasons include attraction of motivated employees and employee retention (Oyer and Schaefer, 2005). Hochberg and Lindsey (2010) find a positive causal effect of stock options on firm performance using an IV approach. Interestingly, their effect is stronger for firms that grant stock options to non-executive employees. Furthermore, Chang et al. (2015) find that stock options for rank and file employees correlate with innovation as measured by patents and patent citations. I provide a behavioral micro-foundation in line with these effects.[78]

A competing explanation why firms pay stock options even though their incentive effect is muted for rank and file employees is a model with stochastic outside options (Oyer, 2004). In this model, employee turnover is costly and salaries are sticky, so deterministic contracts are not easily adjusted to the stochastic outside option. Stock options, however, allow the principal to inject stochasticity into the agent's payment, which turns out to be optimal, even though the agent is risk averse and the stock options do not have an incentive effect. Another explanation highlights tax benefits of stock options (Babenko and Tserlukevich, 2009), but it is unclear how these benefits compare to the drawback of diluting shareholder equity. I add a behavioral

[76]For example, a yearly bonus depends on overall profits at many car companies. See `https://www.auto-motor-und-sport.de/verkehr/bonus-zahlungen-der-autobauer-2022/`, last accessed August 2nd 2023.

[77]I restrict attention to linear contracts which is innocuous under risk neutrality. Under some assumptions linear contracts are also optimal under risk aversion (see Holmstrom and Milgrom, 1987; Edmans and Gabaix, 2011).

[78]Innovation is often firm specific and is thus fueled by company specific skills.

explanation for stock options to the literature: they reinforce the agent's tendency to optimism.

The paper also contributes to a literature on contracting with behavioral agents (see Kőszegi, 2014; Spiegler, 2011, for reviews). Most of the existing literature considers settings where the overconfidence (optimism) bias is exogenously fixed (Sandroni and Squintani, 2007; Landier and Thesmar, 2008; De la Rosa, 2011; Spinnewijn, 2013). If the optimism bias is exogenous, it cannot explain the use of stock options, as the optimism-reinforcement channel is missing.

Contract theory with anticipatory utility is relatively understudied. Most closely related are the companion papers Immordino et al. (2011, 2015) who study a moral hazard setting where the agent chooses his effort after privately receiving an informative signal about the project's profitability in the good state. In their model, effort increases the likelihood of the good state, which is in contrast to my model. Bridet and Schwardmann (2023) also study contracting with endogenous optimism, albeit in an adverse selection framework. Empirical evidence of belief-based utility in contracting settings remains, to the best of my knowledge, scarce.

The theoretical literature on belief formation with anticipatory utility outside the contract theory setting is more developed (see e.g. Akerlof and Dickens, 1982; Caplin and Leahy, 2001; Brunnermeier and Parker, 2005; Brunnermeier et al., 2007; Bénabou, 2013; Schwardmann, 2019). There is also empirical support for utility derived from beliefs in the economic literature (see, e.g. Zimmermann, 2020; Kocher et al., 2014; Falk and Zimmermann, 2016; Coutts, 2019). For a survey of the psychology literature on believe based utility see Kunda (1990).

The rest of the paper is organized as follows. Section 3.2 introduces the model setup. Section 3.3 discusses the agent's problem, while Section 3.4 solves for the optimal contract. Section 3.5 concludes.

## 3.2 Model Setup

A risk neutral principal (she, the firm) seeks to hire a risk neutral agent (he, the employee) to perform a project. Profits increase in firm specific skills, $s$, that the agent can accumulate at a cost of $c(s) = \frac{1}{2}s^2$. The principal offers a contract, which consists of a bonus, $b$, to incentivize skill accumulation, stock options, $n$, to manage optimism without directly impacting the skill decision of the agent, and a wage, $w$, that shifts the contract without impacting skill and optimism.[79] Throughout, I assume that each component of the contract is weakly positive. Negative variable pay (bonus or options) sets a perverse incentive to sabotage, which is implausible in practice. I also rule out negative wages, since employees usually do not pay the principle. Since the wage is a simple shift of the contract, I normalize any (potentially relevant) minimum wage to zero.

The good and bad state of the world realize with (objective) probability $q$ and $1-q$, respectively. In contrast to standard moral hazard settings, the objective probability of the good state, $q$, is unaffected by the agent's skill. The principal chooses the contract $(w, b, n)$ to

[79] To a risk neutral rational agent wage, $w$, and stock options, $n$, are perfect substitutes.

maximize her profits, $\pi = s - w - q(bs + n)$ subject to incentive and participation constraints.

The principal cannot fully commit to the bonus in every state of the world. If the bonus is linked to a promotion, this amounts to assuming promotable positions are harder to come by if the firm is not doing well. Options are more profitable in the good state, but also have an *option value* in the bad state, say $v_h > v_l > 0$. Any standalone value $v_l$ may be relegated to the fixed wage and without loss of generality I normalize $v_h - v_l = 1$. The model may be extended to the case where bonus and stock options are not perfectly correlated by introducing separate probabilities $q_b$ and $q_n$, but this is suppressed here for clarity.

The agent chooses skill, $s$, to maximize (expected) material utility,

$$u = w - \frac{1}{2}s^2 + p(bs + n), \tag{13}$$

where $p$ is the agent's subjective probability of the good state of the world. Following the optimal expectations framework (Brunnermeier and Parker, 2005; Spiegler, 2008, 2019), the agent chooses the belief, $p$, to maximize (expected) felicity,

$$f = w - \frac{1}{2}s^2 + \alpha p(bs + n) + (1 - \alpha)q(bs + n), \tag{14}$$

where $\alpha \in (0, 1)$, $(1 - \alpha)$ is the weight on anticipated (material) utility.[80] The agent optimizes his beliefs, $p$, realizing the he takes those beliefs as given when choosing skill to maximize material utility $u$. In the rational benchmark, the agent cannot distort her beliefs and thus $p = q$. The agent has the option not to work at the principal's firm and instead earn a wage on a competitive labour market. This gives him a fixed outside option of $\hat{u}$.

### 3.2.1 Discussion of Model Setup

The model is designed to emphasize that an individual employee does not have an effect on aggregate firm outcomes like profits or the stock price. Further, stock options do not play a direct role in incentivizing skill accumulation. This is fundamentally different to standard moral hazard models where effort increases the likelihood of the good state. Immordino et al. (2011, 2015) incorporate anticipated utility, but also make the probability of the good state a function of effort.

In contrast, I model a complementarity between stock options and bonus, through the agent's belief choice. By increasing the payoff in the good state without incentivizing the costly skill directly, the principal can reinforce the optimistic tendency of the agent. Optimism, in turn, makes it cheaper to incentivize the skill, since the agent believes he is getting the bonus more often. Failure of the firm to fully commit to the bonus is instrumental here, but not without

[80] In the bulk of the paper I restrict attention to $\alpha \in \left(0, \frac{1-q}{2-q}\right)$, as for larger weights anticipation dominates material utility and the agent is becomes vacuously optimistic.

support. If the firm is doing well, more opportunities for promotion open up, which is beneficial to agents with company-specific skills. Fundamentally, it assumes good times for the firm are also good for its workers, which is well established in the literature on rent sharing (see e.g. Kline et al., 2019; Saez et al., 2019; Lamadon et al., 2022). Another reason for imperfect commitment might be overly short-termist management incentives, which leads managers to sacrifice bonus payments in the short-term despite profitability in the longer term (for models of short-termism see e.g. Bolton et al., 2006; Stein, 1988).

My model assumes the contract is linear in the bonus, which is an innocuous assumption when the agent is risk neutral and rational. The restriction to linear contracts is widespread in the literature for tractability. Indeed, under some conditions linear contracts are also optimal (see Holmstrom and Milgrom, 1987; Edmans and Gabaix, 2011). In practice, bonus contracts are often conditional on meeting a target, where partially meeting the target earns a fraction of the contract in a (near) linear fashion and overachieving may be rewarded similarly. Pertaining to the promotion interpretation, note that promotions may be (linearly) increasing in skill as well.

### 3.3 The Agent's Problem

The optimal expectations agent chooses his belief $p$ to maximize felicity $f$ as given in Equation 14. He receives utility from exaggerating the probability of the good state, since he anticipates receiving the bonus and options more often. At the same time, the agent realises that such optimism leads him to acquire more company-specific skills, which is costly. The agent trades off the increase in anticipated utility from choosing optimistic beliefs with the decrease in material utility from over-accumulation of the skill.

Accordingly, I solve the agent's problem using backwards induction. First, the agent maximizes his material utility $u$ in Equation 13 by choosing the optimal skill level taking his *subjective* belief as given. In doing so he balances the marginal benefit from skill, namely receiving the bonus $b$ with probability $p$, with the marginal skill cost $c'(s) = s$. Thus, he sets skill level $s = pb$. This skill decision is analogous to the rational benchmark of $s^{rat} = qb$, except the optimal expectations agent has subjective expectations rather than the objectively correct probability $q$ that a rational agent and the realistic principal apply.

It remains to solve for the agents belief $p$. The agent chooses his subjective beliefs optimally. In particular, he chooses his belief $p$ to maximize felicity $f$ in Equation 14. He receives the wage $w$ and has to pay skill costs $\frac{1}{2}s^2$ for sure, so expected material and anticipated utility only differ in the variable pay components. On the one hand, he anticipates to get the bonus and stock options, $bs+n$, with probability $p$, to which he attaches a weight of $\alpha$. On the other hand, he will get material utility of $bs + n$ with probability $q$, to which he attaches the converse weight $1 - \alpha$. He knows he will choose skill $s = pb$, so he maximizes $w - \frac{1}{2}p^2b^2 + \alpha p(pb^2 + n) + (1 - \alpha)q(pb^2 + n)$.

Thus, the agent balances the marginal increase in anticipatory utility from choosing a higher belief, with the marginal cost of over-skilling in material utility.

I require the agent's beliefs to be a well defined probability, $p \in [0,1]$. There is an interior solution for low enough weight, $\alpha$, on anticipated utility, but as $\alpha$ grows the corner solution $p = 1$ is reached:

$$p = \begin{cases} \frac{\alpha}{1-2\alpha}\frac{n}{b^2} + \frac{1-\alpha}{1-2\alpha}q & \text{if } \alpha \leq \frac{(1-q)b^2}{(2-q)b^2+n} \\ 1 & \text{else} \end{cases} \tag{15}$$

The optimal expectations of the agent, $p$, are not only a function the exogenous parameters $\alpha$ and $q$, but also of the endogenous variables $b$ and $n$. This highlights the principal's scope to influence the agent's beliefs. By paying more stock options the principal can reinforce the agents tendency to optimism. This works simply by increasing the agent's payoff contingency in the good state. On the other hand, the agent's beliefs decrease in the bonus. The agent realises that a higher bonus means a higher skill level, all else equal. This changes the trade-off between anticipated utility and material utility, since the costly over-accumulation of skill now occurs at a higher level of the (quadratic) cost function. The results from the agent's problem are summarized in Lemma 3.1.

**Lemma 3.1.** *In the rational benchmark, agent and principle share a common prior $q$ and the agent chooses skill level $s = qb$.*
*The optimal expectations agent chooses skill level $s = pb$, where $p = \frac{\alpha}{1-2\alpha}\frac{n}{b^2} + \frac{1-\alpha}{1-2\alpha}q$ if $\alpha \leq \frac{(1-q)b^2}{(2-q)b^2+n}$ and $p = 1$ otherwise.*

The proof is in Appendix C.1. The lemma highlights the (potentially) extreme optimism of the agent. If $\alpha > \frac{1-q}{2-q}$, the agent will believe $p = 1$ even when he receives no stock options at all. In the next section I consider the optimal contract assuming $\alpha \leq \frac{1-q}{2-q}$ to discount this extreme and somewhat vacuous case.

## 3.4 The Principal's Problem

The principal uses a linear contract $(w, b, n)$, which consists of a wage $w$, a bonus $b$, and stock options $n$. The bonus $b$ incentivizes skill, though this is potentially quite costly because the agent requires compensation not only for the direct cost of accumulating the skill, but also for the risk of not receiving the bonus in the bad state of the world.[81] The principal can use stock options $n$ to manage the agent's optimism without *directly* affecting his skill decision. Of course, inducing more optimism will indirectly increase the agent's skill through the belief channel. The wage $w$ can shift the contract without impacting skill or optimism.

[81] Whether the principal has to cover the agent's skill cost hinges on whether the participation constraint is binding. I consider both cases. I restrict attention, however, to the risk neutral case. In the risk averse case the risk of not receiving the bonus is even more costly and optimism, consequently, more profitable for principal.

In the rational benchmark, the principal's program is

$$\begin{aligned} \max_{w,b,n} \quad & s - w - q(bs+n) \\ \text{s.t.} \quad & s = qb \quad \text{(IC)} \\ & w - \frac{1}{2}s^2 + q(bs+n) \geq \hat{u} \quad \text{(PC)} \\ & w, b, n \geq 0, \end{aligned} \tag{16}$$

where I assume each wage component is at least weakly positive. First consider the case where the outside option is small, $\hat{u} < \frac{1}{8}$, so that the participation constraint (PC) is slack.

If this is the case we can ignore the PC and the program simplifies to

$$\begin{aligned} \max_{w,b,n} \quad & qb - w - q^2b^2 - qn \\ & w, b, n \geq 0. \end{aligned} \tag{17}$$

It is easy to see that the wage will be zero at the optimum, $w = 0$, since it decreases profits and does not play any other role if the PC is slack. An analogous argument holds for stock options $n$ in the rational benchmark. The first order condition of the reduced principal's problem then is $\frac{\partial \pi}{\partial b} = q - 2q^2 b \stackrel{!}{=} 0$, which yields the rational benchmark bonus $b^r = \frac{1}{2q}$. This implements a skill of $s^r = qb = \frac{1}{2}$ and a profit of $\pi^r = \frac{1}{4}$. The resulting contract is $(w^r, b^r, n^r) = (0, \frac{1}{2q}, 0)$, mirroring the puzzle that stock options should not be used to remunerate (unbiased) rank and file employees.

Next, consider the case where the principal faces an optimal expectations agent. The principal's program facing the optimal expectations agent is

$$\begin{aligned} \max_{w,b,n} \quad & s - w - q(bs+n) \\ & s = pb \quad \text{(IC)} \\ & p = \frac{\alpha}{1-2\alpha}\frac{n}{b^2} + \frac{1-\alpha}{1-2\alpha}q \quad \text{(OE)} \\ & w - \frac{1}{2}s^2 + \alpha p(bs+n) + (1-\alpha)q(bs+n) \geq \hat{u} \quad \text{(PC)} \\ & w, b, p, n \geq 0, \quad p \leq 1, \end{aligned} \tag{18}$$

where the incentive constraint (IC), $s = pb$ is complicated by the fact that the beliefs $p$ is now a function of the bonus and stock options. This highlights that the principal determines the agent's beliefs through her choice of the ratio of options to bonus $\frac{n}{b^2}$. In particular, she may induce full optimism by choosing to pay $n = \frac{b^2}{\alpha}\left[1 - 2\alpha - (1-\alpha)q\right]$ stock options, which implements $p = 1$. The other corner, $p = 0$, is not feasible, as $p > q > 0$ for $n, b \geq 0$.[82]

[82] Implementing unbiased beliefs, $p = q$, requires negative stock options, which mean the principal hands out put options instead of the call options that we typically observe. In a put option the agent would lose instead of earn money in the good state of the world, which yields a perverse incentive for the agent to sabotage the firm.

Further, the optimal expectations agent evaluates the contracts felicity $f$, when deciding whether to accept the contract. Since the belief $p$ is optimally chosen his PC is less strict for any given contract. The PC is slack if $\hat{u} < \frac{1-2q}{8q^2(q^2-q)}$ for $q < 0.5$ and $\hat{u} < \frac{2q-1}{8q^2}$ for $q \leq 0.5$. It is, again, easy to see that the principle does not pay a salary at the optimum, $w = 0$, as profits are decreasing in $w$ and the PC is slack. As skill is the only source of profit in this model and $s = pb$, it follows directly that $b = 0$ cannot be optimal. This yields the much simpler problem,

$$\begin{aligned} \max_{b,p} \quad & pb - \frac{1-\alpha}{\alpha} q b^2 (p-q) \\ \text{s.t. } & b > 0, \quad n \geq 0, \quad p \leq 1. \end{aligned} \tag{19}$$

Recall that I restrict attention to the interesting case, where $\alpha < \frac{1-q}{2-q}$. Otherwise, the agent's weight on anticipatory utility is so large that she is fully optimistic regardless of the contract's features. The optimal contract is twofold. If the weight is large enough, $\alpha \in [\frac{q}{1+q}, \frac{1-q}{2-q}]$ for $q < 0.5$, the principal uses stock options to implement full optimism, $p = 1$. Otherwise, the weight on anticipated utility is so low that stock options are not profitable.

**Proposition 3.1.** *Suppose the participation constraint is slack, then in the optimal contract stock options are used to implement $p = 1$ if $\alpha \in [\frac{q}{1+q}, \frac{1-q}{2-q}]$ for $q < 0.5$. Otherwise, no stock options are used.*

The proof and the full contract is in the Appendix C.2.

The agent's tendency to become optimistic is beneficial for the principal, since it reduces the cost of incentivizing company-specific skill accumulation. Recall that those costs accrue through two channels. First, the agent incurs direct costs of skill accumulation $c(s)$. Second, the agent understands she is only rewarded for skill in the good state. Any increase in the agent's *subjective* probability therefore decreases the bonus needed to implement the targeted skill level. Once the weight on anticipated utility is large enough the principal uses her power over the agent's beliefs to full extent and implements $p = 1$. The agent's belief around this cutoff is not continuous, rather it jumps discontinuously. This is to the principal's benefit and so profits are also discontinuous around the cutoff.

The optimal skill level is greater than in the rational benchmark. Skill, in this model, is the only source of production and thus, one might suspect a higher skill level signals a welfare increase. If one takes the stance that anticipated utility is permanent and should be used to evaluate welfare, welfare is increasing in skill. If, however, one takes the position that anticipatory utility is transitory, one should restrict attention to material utility when evaluating welfare. In this case, the over-accumulation of skill due to the agent's (contract-induced) optimism leads to lower welfare.

### 3.4.1 Binding Participation Constraint

The case where the participation constraint is binding is significantly more complicated and directly solving for the optimal contract is intractable. It can, however, still be shown that the optimal contract exhibits $w = 0$ as before and that stock options play a role if the weight on anticipated utility is large enough.

Suppose the optimal contract $(w, b, n)$ would feature $w > 0$. Consider a decrease in the wage, $w' = w - q\epsilon$, and an increase in stock options, $n' = n + \epsilon$, that keeps the (expected) payments by the principle fixed. At fixed beliefs, the agent would accepts this new contract, since it creates slack in the participation constraint,

$$
\begin{aligned}
&w' - \frac{1}{2}s^2 + \alpha p(bs + n') + (1-\alpha)q(bs + n') \\
&= w - q\epsilon + q(n+\epsilon) + \underbrace{\alpha(p-q)(n+\epsilon)}_{>0} - \frac{1}{2}s^2 - pqb^2 - \alpha(p-q)pb^2 > \hat{u}.
\end{aligned}
$$

The agent then optimally revises her beliefs given the new contract, which gives him (weakly) higher felicity, so that he will still accept the contract at the new beliefs, $p' > p$. This yields higher profits for the principle, as she now implements a higher skill for the same bonus without increasing total payments. Thus, the wage must be zero at the optimum, $w = 0$.

Consider now the case where the participation constraint in Equation 18 is binding. It turns out that the optimal contract specifies no salary and positive stock options if $\alpha > \frac{1}{3}$.

**Proposition 3.2.** *If the participation constraint is binding the optimal contract does not specify a wage* $w = 0$. *The optimal contract features positive stock options,* $n > 0$ *if* $\alpha > \frac{1}{3}$.

The proof is in Appendix 3.2.

## 3.5 Conclusion

This paper provides a micro-foundation for paying rank and file employees in stock options, so-called broad based stock options. The agent chooses her beliefs optimally and weighs anticipatory utility against losses in material utility due to over-investment in skill. This tendency to become optimistic is mediated by the convex cost of skill: at a higher skill level over-accumulation of skill is more costly. Thus, a higher bonus conditional on skill limits the agent's optimism. By granting stock options, which are unconditional of skill, the principal can reinforce the agent's tendency to optimism without moving up the skill cost function. Reinforcing the agent's optimism increases skill accumulation indirectly as the agent believes she will receive the bonus more often than she actually does.

The model is in line with causal evidence that firms that give out stock options lead to higher firm performance (Hochberg and Lindsey, 2010). If one considers proprietary innovation, it is also in line with empirical evidence that firms using broad based stock options innovate more (Chang et al., 2015). While it is interesting to explore this further, future work may depart from using observational data and test the mechanism more directly, for example in a lab or field experiment. While there is a theoretical literature on anticipatory utility, empirical evidence is lagging behind, especially in contract theory settings (Zimmermann, 2020; Kocher et al., 2014; Falk and Zimmermann, 2016). Contract theory settings are especially interesting as the principle manipulates the agent's beliefs through the contract. It is an open question whether the agent foresees that this will be costly for her and attenuates the belief manipulation or whether she does not understand and distorts beliefs as in the theory.

## References


Akerlof, G. A. and Dickens, W. T. (1982). The economic consequences of cognitive dissonance. *The American economic review*, 72(3):307–319.

Babenko, I. and Tserlukevich, Y. (2009). Analyzing the tax benefits from employee stock options. *The Journal of Finance*, 64(4):1797–1825.

Bénabou, R. (2013). Groupthink: Collective delusions in organizations and markets. *Review of economic studies*, 80(2):429–462.

Bolton, P., Scheinkman, J., and Xiong, W. (2006). Executive Compensation and Short-Termist Behaviour in Speculative Markets. *The Review of Economic Studies*, 73(3):577–610.

Bridet, L. and Schwardmann, P. (2023). Selling dreams: Endogenous optimism in lending markets. *unpublished manuscript*.

Brunnermeier, M. K., Gollier, C., and Parker, J. A. (2007). Optimal beliefs, asset prices, and the preference for skewed returns. *American Economic Review*, 97(2):159–165.

Brunnermeier, M. K. and Parker, J. A. (2005). Optimal expectations. *American Economic Review*, 95(4):1092–1118.

Caplin, A. and Leahy, J. (2001). Psychological expected utility theory and anticipatory feelings. *The Quarterly Journal of Economics*, 116(1):55–79.

Chang, X., Fu, K., Low, A., and Zhang, W. (2015). Non-executive employee stock options and corporate innovation. *Journal of financial economics*, 115(1):168–188.

Coutts, A. (2019). Testing models of belief bias: An experiment. *Games and Economic Behavior*, 113:549–565.

De la Rosa, L. E. (2011). Overconfidence and moral hazard. *Games and Economic Behavior*, 73(2):429–451.

Edmans, A. and Gabaix, X. (2011). Tractability in incentive contracting. *The Review of Financial Studies*, 24(9):2865–2894.

Falk, A. and Zimmermann, F. (2016). Beliefs and utility: Experimental evidence on preferences for information. *unpublished manuscript*.

Hall, B. J. and Murphy, K. J. (2003). The trouble with stock options. *Journal of economic perspectives*, 17(3):49–70.

Hochberg, Y. V. and Lindsey, L. (2010). Incentives, Targeting, and Firm Performance: An Analysis of Non-executive Stock Options. *The Review of Financial Studies*, 23(11):4148–4186.

Holmstrom, B. and Milgrom, P. (1987). Aggregation and linearity in the provision of intertemporal incentives. *Econometrica: Journal of the Econometric Society*, pages 303–328.

Immordino, G., Menichini, A. M. C., and Romano, M. G. (2011). A simple impossibility result in behavioral contract theory. *Economics Letters*, 113(3):307–309.

Immordino, G., Menichini, A. M. C., and Romano, M. G. (2015). Contracts with wishful thinkers. *Journal of Economics & Management Strategy*, 24(4):863–886.

Kline, P., Petkova, N., Williams, H., and Zidar, O. (2019). Who Profits from Patents? Rent-Sharing at Innovative Firms*. *The Quarterly Journal of Economics*, 134(3):1343–1404.

Kocher, M. G., Krawczyk, M., and van Winden, F. (2014). 'let me dream on!'anticipatory emotions and preference for timing in lotteries. *Journal of Economic Behavior & Organization*, 98:29–40.

Kőszegi, B. (2014). Behavioral contract theory. *Journal of Economic Literature*, 52(4):1075–1118.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological bulletin*, 108(3):480.

Lamadon, T., Mogstad, M., and Setzler, B. (2022). Imperfect competition, compensating differentials, and rent sharing in the us labor market. *American Economic Review*, 112(1):169–212.

Landier, A. and Thesmar, D. (2008). Financial contracting with optimistic entrepreneurs. *The Review of Financial Studies*, 22(1):117–150.

Oyer, P. (2004). Why do firms use incentives that have no incentive effects? *The Journal of Finance*, 59(4):1619–1650.

Oyer, P. and Schaefer, S. (2005). Why do some firms give stock options to all employees?: An empirical examination of alternative theories. *Journal of financial Economics*, 76(1):99–133.

Saez, E., Schoefer, B., and Seim, D. (2019). Payroll taxes, firm behavior, and rent sharing: Evidence from a young workers' tax cut in sweden. *American Economic Review*, 109(5):1717–63.

Sandroni, A. and Squintani, F. (2007). Overconfidence, insurance, and paternalism. *American Economic Review*, 97(5):1994–2004.

Schwardmann, P. (2019). Motivated health risk denial and preventative health care investments. *Journal of health economics*, 65:78–92.

Spiegler, R. (2008). On two points of view regarding revealed preference and behavioral economics. *The foundations of positive and normative economics: A handbook*, pages 95–115.

Spiegler, R. (2011). *Bounded rationality and industrial organization*. Oxford University Press.

Spiegler, R. (2019). Behavioral economics and the atheoretical style. *American Economic Journal: Microeconomics*, 11(2):173–194.

Spinnewijn, J. (2013). Insurance and perceptions: How to screen optimists and pessimists. *The Economic Journal*, 123(569):606–633.

Stein, J. C. (1988). Takeover threats and managerial myopia. *Journal of political economy*, 96(1):61–80.

Zimmermann, F. (2020). The dynamics of motivated beliefs. *American Economic Review*, 110(2):337–363.

# A Appendix: Managing Bidder Learning in Retail Auction

## A.1 Seller's Problem

In this Section we present a simplified model used to derive Observations 1.1 and 1.2 in Section 1.5. We denote the number of overbidders by $o_t$ and the number of non-overbidders by $s_t$. We assume that non-overbidders buy at the fixed price whenever they do not buy in the auction. To simplify, we assume that all bidders have the same latent bid $\beta > p$. We assume overbidders never buy at the fixed price. While this may sound restrictive, the seller will always supply each overbidder through the auction at a auction price weakly higher than the fixed price. All bidders have unit demand.[83]

**Definition A.1.** Seller's Problem

A profit-maximizing firm solves the following problem:

$$\max_{\{q_t\}_{t=0}^{\infty}} \sum_t \delta^t \pi(q_t, o_t, s_t),$$

where

$$\pi(q_t, o_t, s_t) = \begin{cases} \beta \cdot q_t + p \cdot s_t & if\ q_t \leq o_t \\ p(o_t + s_t) & if\ o_t < q_t \leq o_t + s_t \\ 0 & if\ o_t + s_t < q_t \end{cases}$$

subject to:

$$o_{t+1} = \begin{cases} o_t - (\epsilon + \iota) q_t & if\ q_t \leq o_t \\ o_t & if\ q_t > o_t \end{cases}$$

$$s_{t+1} = \begin{cases} s_t + \iota q_t & if\ q_t \leq o_t \\ s_t & if\ q_t > o_t \end{cases}.$$

## A.2 Proof of Observation 1.1

*Proof.* We guess two policy functions (always choose $q_t = o_t + s_t$) and always choose $q_t = s_t$). Since the union of these conditions covers the parameter space the desired result follows.

Because $o_0, s_0$ are both larger than one and $\epsilon, \iota$ are both smaller than one we guarantee that $o_t > 0 \wedge s_t > 0\ \forall t$.

In the remainder of this proof we drop the time index to simplify our notation.

We can simplify the strategy space because some actions are dominated and some are outcome-equivalent. All actions with $q > o + s$ are dominated because profits are zero and we

[83] Most of these assumptions are for illustrative purposes and are dropped in the empirical model in Section 1.6.1

can get positive profits with $q = o + s$. Profits are constant over $o < q \leq o + s$. Thus we can eliminate this interval from the action space if we include its upper boundary $o + s$.

Having simplified the strategy space in this way we state the Bellman equation.

$$V(o,s) = max_{q \in Q} \begin{cases} q \cdot \beta + sp + \delta V(o - (\epsilon + \iota) \cdot q, s + \iota \cdot q) & if\ q \leq o \\ (s+o)p + \delta V(o,s) & if\ q = o + s \end{cases},$$

where $Q = [0, o] \cup \{o + s\}$.

We guess and verify the policy $q = o + s$. The result of this policy is is that the firm sells $o + s$ unity each period at a price of $p$. This leads to the following value function

$$V(o,s) = \Sigma_{k=0}^{\infty} \delta^k (o_t + s_t)p = \frac{(o_t + s_t)p}{1 - \delta}.$$

We derive conditions under which this value function solves the Bellman equation

$$\frac{(o+s)p}{1-\delta} = max_{q \in Q} \begin{cases} q \cdot \beta + sp + \delta \frac{(o+s-\epsilon q)p}{1-\delta} & if\ q \leq o \\ (s+o)p + \delta \frac{(o+s)p}{1-\delta} & if\ q = o + s \end{cases},$$

where $Q = [0, o] \cup \{o + s\}$.

We need to check two cases, either the left arm ($q \leq o$) of the right-hand side of the Bellman equation rises or falls in $q$. It (weakly) rises if

$$\beta \geq \frac{\delta}{1-\delta} \epsilon p.$$

In this case profits are either maximized at $q = o + s$ or at $q = o$. They are maximized at $q = o + s$ and our guess is true if

$$o \cdot \beta + sp + \delta \frac{(s + (1-\epsilon)o)p}{1-\delta} \leq (o+s)p + \delta \frac{(s+o)p}{1-\delta} \tag{20}$$

$$\leftrightarrow \frac{\beta - p}{p} \leq \frac{\delta}{1-\delta} \epsilon. \tag{21}$$

If

$$\frac{\beta}{p} < \frac{\delta}{1-\delta} \epsilon \tag{22}$$

the left arm ($q < o$) of the right-hand side of the Bellman falls in $q$.

In this case profits are either maximized at $q = 0$ or at $q = o + s$. They are maximized at $q = o + s$ if

$$sp + \delta\frac{(s+o)p}{1-\delta} < (o+s)p + \delta\frac{(s+o)p}{1-\delta}$$
$$\leftrightarrow 0 < op,$$

which is true. Since condition 21 is strictly stronger than 22, we can verify our guess of no overbidding if condition 21 holds.

We guess that the seller wants all auctions to end in an overpay. Then the seller derives a profit of $p \cdot s$ from the initial non-overbidders in perpetuity. They derive a profit of $\beta$ per overbidder in each period from a steadily declining stock of overbidders. This results in $o_t(1-\epsilon-\iota)^k\beta$ in each future period $k$. In each future period a fraction $i$ of the current overbidders is transformed into non-overbidders $o_t(1-\epsilon-\iota)^{p-1}\iota p$. Consequently, in period $k$ there are $\Sigma_{p=1}^{k} o_t(1-\epsilon-\iota)^{p-1}\iota p$ that were generated through intensive margin learning. The discounted sum of these period profits yields the value function under the conjecture that the seller ends all auctions in an overpay

$$\begin{aligned}
V(o,s) &= \Sigma_{k=0}^{\infty}\delta^k(o(1-\epsilon-\iota)^k\beta + sp) + \Sigma_{k=1}^{\infty}\delta^k\Sigma_{p=1}^{k} o(1-\epsilon-\iota)^{p-1}\iota p \\
&= o\beta\Sigma_{k=0}^{\infty}\delta^k(1-\epsilon-\iota)^k + sp\Sigma_{k=0}^{\infty}\delta^k + \frac{o\iota p}{1-\epsilon-\iota}\Sigma_{k=1}^{\infty}\delta^k\Sigma_{p=0}^{k}(1-\epsilon-\iota)^p - 1 \\
&= \frac{o\beta}{1-\delta(1-\epsilon-\iota)} + \frac{sp}{1-\delta} + \frac{o\iota p}{1-\epsilon-\iota}\Sigma_{k=1}^{\infty}\delta^k\left(\frac{1-(1-\epsilon-\iota)^{k+1}}{\epsilon+\iota} - 1\right) \\
&= \frac{o\beta}{1-\delta(1-\epsilon-\iota)} + \frac{sp}{1-\delta} \\
&+ \frac{o_t\iota p}{(1-\epsilon-\iota)(\epsilon+\iota)}\Sigma_{k=1}^{\infty}\delta^k(1-\epsilon-\iota) - (1-\epsilon-\iota)\delta^k(1-\epsilon-\iota)^k \\
&= \frac{o\beta}{1-\delta(1-\epsilon-\iota)} + \frac{sp}{1-\delta} + \frac{o_t\iota p}{\epsilon+\iota}\Sigma_{k=1}^{\infty}\delta^k - \delta^k(1-\epsilon-\iota)^k \\
&= \frac{o\beta}{1-\delta(1-\epsilon-\iota)} + \frac{sp\delta}{1-\delta} + \frac{o\iota p}{\epsilon+\iota}\left[\frac{\delta}{1-\delta} - \frac{\delta(1-\epsilon-\iota)}{1-\delta(1-\epsilon-\iota)}\right].
\end{aligned}$$

We look for conditions under which this conjecture for the value function solves the seller's Bellman equation (equation 20). If

$$\beta < \delta(\epsilon+\iota)\left(\frac{\beta}{1-\delta(1-\epsilon-\iota)} + \frac{\iota p}{\epsilon+\iota}\left[\frac{\delta}{1-\delta} - \frac{\delta(1-\epsilon-\iota)}{1-\delta(1-\epsilon-\iota)}\right]\right) - \frac{\delta p}{1-\delta} \tag{23}$$

there is no overpaying because the left arm of profits fall in $q$. Then we have to compare $q = 0$ with $q = o + s$. Since the latter leads to higher period profits and both lead to the same future profits the firm prefers $q = o + s$, which refutes our conjecture.

If condition 23 does not hold the left-arm of the values function rises in $q$ and the seller ends every auction in overpaying if he prefers that to $q = o$. This is the case if

$$\beta - p \geq \delta(\epsilon+\iota)\Bigg( \beta(1-\delta(1-\epsilon-\iota))^{-1}$$
$$+ \iota p(\epsilon+\iota)^{-1}\left[\frac{\delta}{1-\delta} - \frac{\delta(1-\epsilon-\iota)}{1-\delta(1-\epsilon-\iota)}\right]\Bigg) - \iota\delta p(1-\delta)^{-1}$$
$$\leftrightarrow$$
$$\beta - p \geq \delta\left(\iota+\epsilon\right)\beta\left(1-\delta\left(1-\epsilon-\iota\right)\right)^{-1} + \iota p\frac{\delta^2}{1-\delta} - \iota p\frac{\delta}{1-\delta}$$
$$- \iota p\frac{\delta^2\left(1-\epsilon-\iota\right)}{1-\delta\left(1-\epsilon-\iota\right)}$$
$$\leftrightarrow$$
$$\beta - p \geq \delta\left(\iota+\epsilon\right)\beta\left(1-\delta\left(1-\epsilon-\iota\right)\right)^{-1} + \iota p\frac{\delta^2-\delta}{1-\delta} - \iota p\frac{\delta^2\left(1-\epsilon-\iota\right)}{1-\delta\left(1-\epsilon-\iota\right)}$$
$$\leftrightarrow$$
$$\beta - p \geq \delta\left(\iota+\epsilon\right)\beta\left(1-\delta\left(1-\epsilon-\iota\right)\right)^{-1} + \iota p\frac{\delta^2-\delta}{\left(1-\delta\right)\left(\ 1-\delta\left(1-\epsilon-\iota\right)\right)},$$

where the last step follows if since$\frac{1}{d} > 1-\epsilon-\iota$, which is always true since $\epsilon, \iota$ and $d$ are all between zero and one. Having simplified the condition so far we can collect terms and solve for a condition on $\epsilon$

$$\begin{aligned}
\left(1-\delta\left(1-\epsilon-\iota\right)\right)\left(\beta-p\right) &\geq \delta\left(\iota+\epsilon\right)\beta + \iota p\frac{\delta^2-\delta}{1-\delta} \\
\leftrightarrow \left(1-d\right)\beta - \frac{1-2\delta+\delta^2+\delta e-\delta^2\epsilon}{1-d}p &\geq 0 \\
\leftrightarrow \left(1-\delta\right)\beta - \left(1-\delta\right)p - \delta\epsilon p &\geq 0 \\
\leftrightarrow \frac{\beta-p}{p}\frac{1-\delta}{\delta} &\geq \epsilon.
\end{aligned}$$

This condition covers all cases in which the other strategy is not optimal. Consequently, the seller either sets $q = o$ or $o < q \leq o + s$.

□

## A.3 Proof of Proposition 1.2

*Proof.* The potential outcome for the the untreated (people that did not overpay) is the probability that an overbidder submits a strict non-overbid,

$$E[\textit{non-overbid}_t^{t+1}(0)|u_i] = \mathbb{P}(p_{t+1} > \beta_{i,t+h} > b(q_{t+h})|u_i).$$

If we exogenously assign a bidder to the treated status they either stay an overbidder, become a non-overbidder or leave. In the cases in which they become a non-overbidder they also change their latent bid distribution. This leads to a change in probabilities which we denote by switching from $P$ to $P'$. We calculate the potential outcome of a bidder treated in $t$ and observed in $t+1$ as

$$\begin{aligned} E[\textit{non-overbid}_t^{t+1}(1)|u_i] =& E[(1-\epsilon_i-\iota_i)|u_i]\mathbb{P}(\beta_{i,t+1} < p_{t+1} \wedge \beta_{i,t+1} > b(q_{t+1})|u_i) \\ &+E[\iota_i|u_i]P'(\beta_{i,t+1} < p_{t+1} \wedge \beta_{i,t+1} > b(q_{t+1})|u_i). \end{aligned}$$

Adding an intelligent zero and taking the difference of potential outcomes yields the following expression for the treatment effects

$$\begin{aligned} E[TE^{t+1}_{\textit{non-overbid}}] &= E[-\epsilon_i|u_i]\mathbb{P}(p_{t+1} > \beta_{i,t+1} > b(q_{t+1})|u_i) \\ +\iota_i(P'(p_{t+1} &> \beta_{i,t+h} > b(q_{t+1})|u_i) - \mathbb{P}(p_{t+1} > \beta_{i,t+1} > b(q_{t+1})|u_i)) \end{aligned}$$

□

## A.4 Proof of Proposition 1.3

*Proof.* We are interested in the treatment effect of overpaying in period $t$, but we can only estimate it from pooling observations over a time period. Previously, we have written down the treatment effect from overpaying in $t$ on the next auction, i.e. in auction $t+1$. If we add up these treatment effects, we do not account for the possibility that there could be double treatments. To rule out that effects come from double treatments we first write down the probability of a type change from treatments in periods $t+1$ to $t+k-1$. Using this probability we can calculate the effect from treatment in t excluding subsequent treatments. In Lemma 1.1 we sum over these treatment effects to get an expression for the treatment effect from pooled observations, that we can then use to recover learning rates.

Let $p_{l,k}$ denote the probability that an overbidder changes his type to dropout because of a treatment in periods $t+1$ to $t+k-1$. Notice that a change in type is irreversible in our model, so we simply need to sum over the probability to change type at a specific point in time, but not before that point in time. The probability of a type change in period $t+m$ is the probability of treatment, $\textit{overpaid}_{i,t+m}$ times the probability of changing type due to treatment, $\epsilon_i$. The converse probability is the probability of neither changing to dropout nor to non-overbidder before $t+m$. Thus, we get the following expression for $p_{l,k}$ (the argument for $p_{s,k}$ is analogous).

$$E[p_{l,k}|\epsilon_i, \iota_i, u_i] = E\left[\Sigma_{m=1}^{k-1} overpaid_{i,t+m}\epsilon_i(1-(\epsilon_i+\iota_i)overpaid_{i,t+m-1})^m|\epsilon_i, \iota_i, u_i\right]$$

$$E[p_{s,k}|\epsilon_i, \iota_i, u_i] = E\left[\Sigma_{m=1}^{k-1} overpaid_{i,t+m}\iota_i(1-(\epsilon_i+\iota_i)overpaid_{i,t+m-1})^m|\epsilon_i, \iota_i, u_i\right].$$

Next we use these probabilities to characterize the potential outcomes in period $t+k$ for a bidder who we assign exogenously to be either untreated ($E[\textit{non-overbid}_t^{t+k}(0)|u_i]$) or treated ($E[\textit{non-overbid}_t^{t+k}(1)|u_i]$) in period $t$. Note that the term $(1-E[p_{l,k}|\epsilon_i, \iota_i, u_i])$ captures that the type change did not occur in periods $t+1$ to $t+k-1$.

$$E[\textit{non-overbid}_t^{t+k}(0)|u_i] = E[(E[p_{l,k}|\epsilon_i, \iota_i, u_i])\mathbb{1}(\beta_{i,t+k} < p_{t+k} \wedge \beta_{i,t+k} > b(q_{t+k})|u_i)]$$

$$= E[E[p_{l,k}|\epsilon_i, \iota_i, u_i]|u_i] \cdot \mathbb{P}(\beta_{i,t+k} < p_{t+k} \wedge \beta_{i,t+k} > b(q_{t+k})|u_i)$$

$$E[\textit{non-overbid}_t^{t+k}(1)|u_i] = E[(1-\epsilon_i)(E[p_{l,k}|\epsilon_i, \iota_i, u_i])\mathbb{1}(\beta_{i,t+k} < p_{t+k} \wedge \beta_{i,t+k} > b(q_{t+k})|u_i)]$$

$$= E[(1-\epsilon_i)(E[p_{l,k}|\epsilon_i, \iota_i, u_i])|u_i] \cdot \mathbb{P}(\beta_{i,t+k} < p_{t+k} \wedge \beta_{i,t+k} > b(q_{t+k})|u_i).$$

The last step in each follows because conditional on $u_i$, $\epsilon_i$ and $\beta_{i,t+k}$ are independent. If we take the difference of potential outcomes we get the treatment effect.

$$E[\textit{TE-overbid}_t^{t+k}|u_i] = E[-\epsilon_i \cdot E[p_{l,k}|\epsilon_i, \iota_i, u_i]|u_i]\mathbb{P}(\beta_{i,t+k} < p_{t+k} \wedge \beta_{i,t+k} > b(q_{t+k})|u_i).$$

The calculation for the treatment effect on observed overbids is analogous.

□

## A.5 Proof of Lemma 1.1

*Proof.* Proof of Proposition 1 We take the expression for the potential outcome of the untreated and the treatment effect from the proof of Proposition 1.3 and divide one by the other.

$$\frac{-\Sigma_{m=0}^{k} E[TE_{non\text{-}overbid}^{t+m}]}{\Sigma_{m=0}^{k} E[\textit{non-overbid}_t^{t+m}(0)|u_i]}$$
$$= \frac{\Sigma_{m=0}^{k} E[\epsilon_i(1-E[p_{l,m}|\epsilon_i, \iota_i, u_i])|u_i]\mathbb{P}(p_{t+m} > \beta_{i,t+m} > b(q_{t+m})|u_i)}{\Sigma_{m=0}^{k} E[(1-E[p_{l,m}|\epsilon_i, \iota_i, u_i])|u_i]\mathbb{P}(p_{t+m} > \beta_{i,t+m} > b(q_{t+m})|u_i)}$$
$$= \frac{\Sigma_{m=0}^{k} E[\epsilon_i(1-E[p_{l,m}|\epsilon_i, \iota_i, u_i])|u_i]}{\Sigma_{m=0}^{k} E[(1-E[p_{l,m}|\epsilon_i, \iota_i, u_i])|u_i]}.$$

The proof for the treatment effect on strict overbids is analogous. □

## A.6 Replication Files

The replication files are available from the authors or at `https://www.dropbox.com/scl/fo/m3g3d24uuglpvr91zkh4d/h?rlkey=dyennumcdltsoala319wwt8tl&dl=0`

## A.7 Control Variables

We calculate two sets of bidder history variables. First, we calculate histories for the bidders in our control and treatment group. Given that we restrict attention to the first overbid of each bidder our bidder history variables only capture behavior and experience in the auctions of this seller before that first overbid. Hence our variables do not control for a previous overbid as there was none by construction. When a bidder overbids on the first bid we do not observe a history before that, because there was none. In this case we substitute the average from treatment and control bidders in the same auction. This substitute may not be available if all control and treatment bidders were new bidders. In this case we keep the NA and exclude these observations in regression that include bidders histories.

The bidder history variables roughly fall into two categories. First, there are variables that measure the average previous behavior. For example, the average difference to the high bid measures whether a bidder usually bids early in the auction and the share of bids by phone measures whether a bidder usually bids by telephone or online. Second, some variables refer more to the experience that the bidder had in the previous auctions. For example, the time in the market measures how many hours have past since the first observed bid for that bidder in our sample and total savings measures how much money the bidder has saved compared to the fixed price.

We calculate the same set of bidder history variables also for the other bidders in the auction, even if they are not in the treatment or control group. Referring back to Section 1.6.5, this controls for the other bidders individual characteristics $u_{-}i$, that were left out of the DAG for simplicity.

Table 14 gives summary statistics for all history variables that we calculate. It is evident that there are differences between the treatment and control groups, which reassures us that it is helpful to control for this set of proxy variables.

Table 14: Average value of bidder history variables, and fixed price ($p_t$) and number of products ($q_t$) at the first overbid, split by overpaid.

| | 0 | | 1 | |
|---|---|---|---|---|
| | Mean | Std. Dev. | Mean | Std. Dev. |
| fixed price | 29.63 | 67.20 | 30.78 | 85.86 |
| quantity | 282.88 | 275.17 | 272.61 | 257.81 |
| new bidder | 0.39 | 0.49 | 0.35 | 0.48 |
| own number of bids | 6.74 | 6.95 | 7.25 | 7.63 |
| own average savings, logged | 3.50 | 1.07 | 3.63 | 1.09 |
| own average bid, logged | 3.56 | 0.62 | 3.60 | 0.66 |
| own time in market (hours) | 1479.14 | 2491.12 | 1756.69 | 2690.80 |
| own share of bids by phone | 0.82 | 0.32 | 0.79 | 0.34 |
| own average difference to the high bid | 11.44 | 18.44 | 12.19 | 24.16 |
| others average number of bids | 44.56 | 39.27 | 47.42 | 39.18 |
| others logged total savings | 4.99 | 1.33 | 5.13 | 1.30 |
| others logged average bid | 3.35 | 0.47 | 3.43 | 0.49 |
| others time in market (hours) | 2502.92 | 2471.15 | 2701.00 | 2404.57 |
| others fraction of new bidders | 0.13 | 0.14 | 0.10 | 0.11 |
| others share of bids by phone | 0.58 | 0.15 | 0.60 | 0.15 |
| others average difference to the high bid | 7.73 | 3.75 | 8.21 | 4.63 |

| Category | Share Overpaid | n |
|---|---|---|
| Heimwerken & Garten | 0.32 | 6246 |
| Mode & Accessoires | 0.28 | 13269 |
| Beauty & Wellness | 0.26 | 15257 |
| Uhren | 0.25 | 8059 |
| Schmuck | 0.22 | 7950 |
| Haushalt | 0.20 | 16786 |
| Möbel & Heimtextilien | 0.16 | 6981 |
| Freizeit & Sammeln | 0.08 | 157 |

Table 15: Average probability of a bidder to be treated (overpay) at their first overbid by show category.

| Weekday | Share Overpaid | n |
|---|---|---|
| Sunday | 0.21 | 13909 |
| Monday | 0.17 | 9706 |
| Tuesday | 0.27 | 10209 |
| Wednesday | 0.28 | 9329 |
| Thursday | 0.24 | 9856 |
| Friday | 0.25 | 9581 |
| Saturday | 0.26 | 12115 |

Table 16: Average probability of a bidder to be treated (overpay) at their first overbid by day of the week.

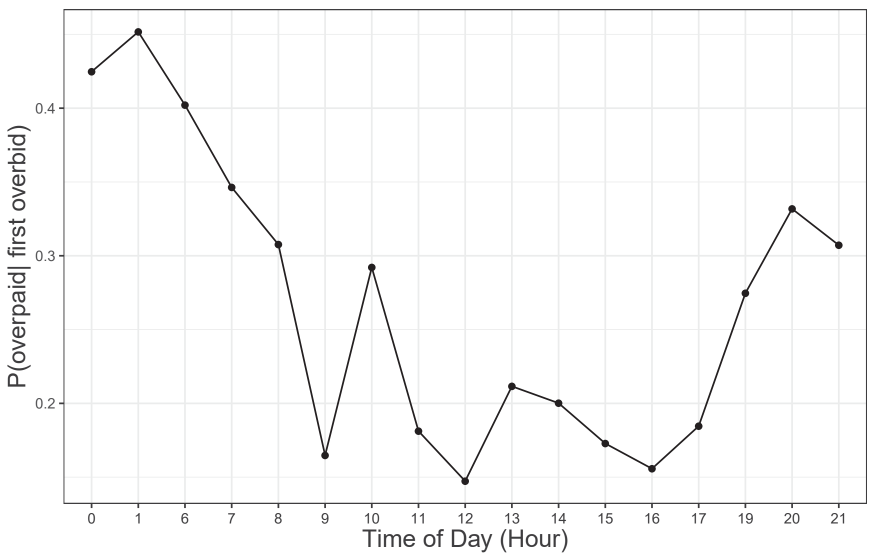



Figure 12: Average probability of a bidder to be treated (overpay) at their first overbid by time of day. Averages are by hour. The time between 18:00 and 19:00 is missing from hour data because of a coding error.

## A.8 Structural Break

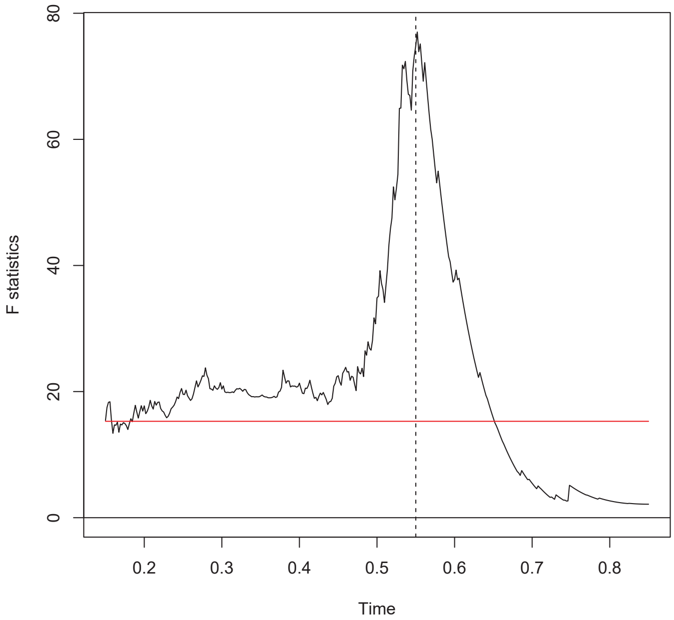



Figure 13: Time series of F statistics for a single shift hypothesis, fitted at every day in our sample. The red line gives the critical value at the 1 percent significance level. We accept the most probable break-point at the dashed line, 16th of May 2018.

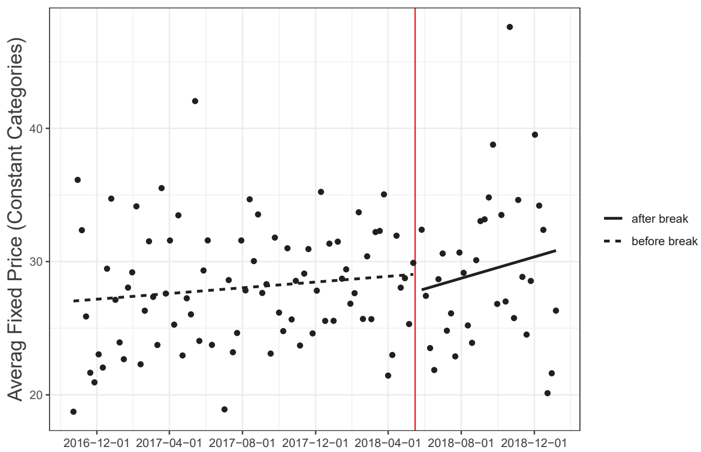
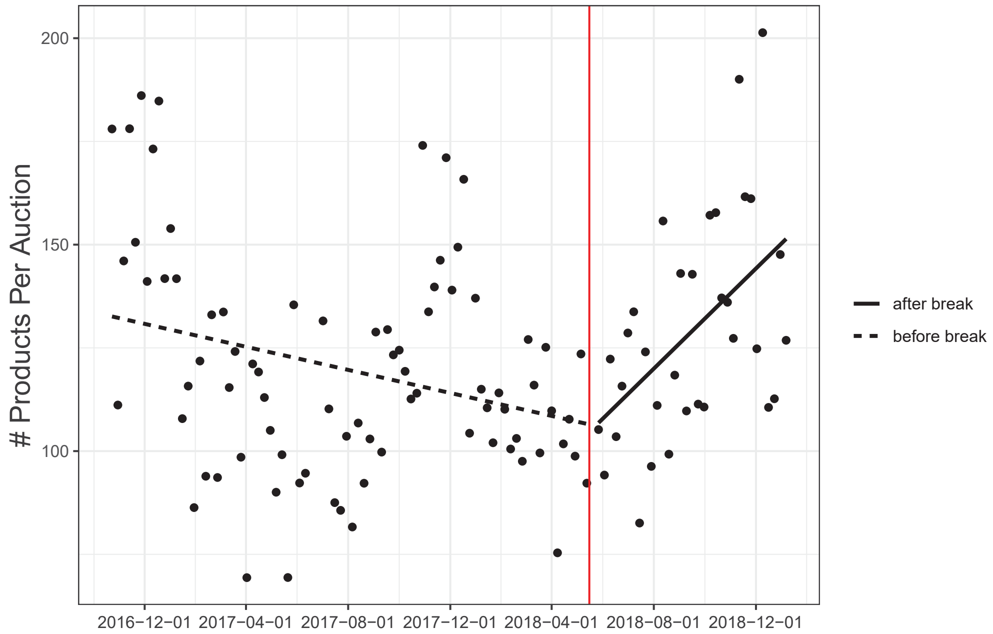
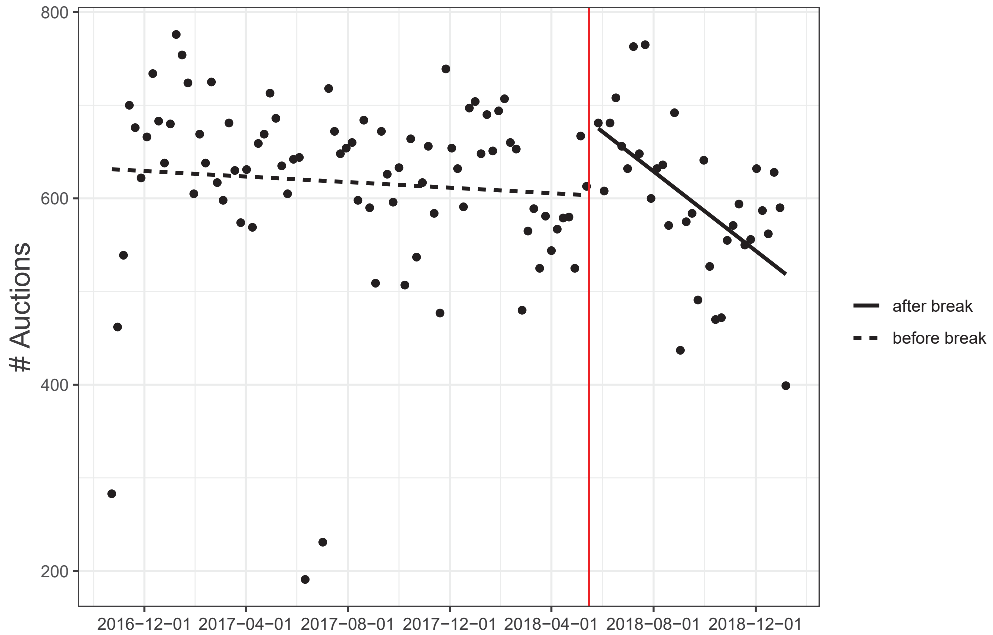



Figure 14: Changes in the fixed price of the auctioned products. We use weekly averages and fit a linear trend at both sides of the structural break.




(a) Number of products per auction.




(b) Number of auctions each week.

Figure 15: Changes in number of auctions and number of products per auction to both sides of the structural break. We use weekly averages and fit a linear trend at both sides of the structural break.

## A.9 Regression Using Period 90-180 Days After Treatment

We redo the regression results for the period of 90-180 days after treatment.

Table 17: Overpaying Reduces #Overbids and #Non-Overbids (90-180)

| | # Overbids | # Overbids | # Non-Overbids | # Non-Overbids |
|---|---|---|---|---|
| Overpaid | −0.173*** | −0.204*** | −0.213 | −0.422** |
| | (0.035) | (0.041) | (0.150) | (0.171) |
| Num.Obs. | 124 136 | 77 029 | 124 136 | 77 029 |
| R2 | 0.060 | 0.121 | 0.095 | 0.204 |
| Cf. Mean | 1.755 | 1.995 | 8.886 | 10.405 |
| Bidder History | No | Yes | No | Yes |
| Window | 90-180 | 90-180 | 90-180 | 90-180 |

* p < 0.1, ** p < 0.05, *** p < 0.01

Table 18: Overpaying Reduces Future Revenue (90-180)

| | Revenue | Revenue |
|---|---|---|
| Overpaid | −0.598 | −7.314 |
| | (4.865) | (6.005) |
| Num.Obs. | 124 136 | 77 029 |
| R2 | 0.051 | 0.126 |
| Cf. Mean | 193.165 | 224.008 |
| Bidder History | No | Yes |
| Window | 90-180 | 90-180 |

* p < 0.1, ** p < 0.05, *** p < 0.01

## A.10 Poisson Regression

We redo the regressions using Poisson regression instead of OLS as we are using count data.

Table 19: Overpaying Reduces #Overbids and #Non-Overbids (Poisson)

| | # Overbids | # Overbids | # Non-Overbids | # Non-Overbids |
|---|---|---|---|---|
| Overpaid | −0.133*** | −0.121*** | −0.034* | −0.040** |
| | (0.024) | (0.022) | (0.020) | (0.019) |
| Num.Obs. | 115 295 | 71 261 | 115 295 | 71 261 |
| R2 | | | | |
| R2 Pseudo | 0.132 | 0.206 | 0.127 | 0.257 |
| Cf. Mean | 1.431 | 1.616 | 5.795 | 6.846 |
| Bidder History | No | Yes | No | Yes |
| Window | 0-90 | 0-90 | 0-90 | 0-90 |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 20: Overpaying Reduces #Overbids and #Non-Overbids (90-180, Poisson)

| | # Overbids | # Overbids | # Non-Overbids | # Non-Overbids |
|---|---|---|---|---|
| Overpaid | −0.103*** | −0.105*** | −0.022 | −0.034** |
| | (0.021) | (0.022) | (0.018) | (0.016) |
| Num.Obs. | 124 136 | 77 029 | 124 136 | 77 029 |
| R2 | | | | |
| R2 Pseudo | 0.101 | 0.205 | 0.162 | 0.345 |
| Cf. Mean | 1.752 | 1.991 | 8.878 | 10.383 |
| Bidder History | No | Yes | No | Yes |
| Window | 90-180 | 90-180 | 90-180 | 90-180 |

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

## A.11 First Five Overbids

We redo the empirical exercise separately using the first five overbids of each bidder. Learning may turn an initial overbidder into a non-overbidder or dropout already after the first overbid if it was overpaid. Accordingly, we observe fewer second overbids than first overbids as is depicted in Figure 16. Learning after the first overbid is likely to be skewed to well-learning overbidders and, hence, we expect to back-out smaller learning rates using the second overbid for each bidder compared to the first overbid. Figure 17 reports the backed-out extensive margin learning rates for the first five overbids. The pattern is the same across methods (ols and poisson regression) and time periods of aggregation (90 days after overbid and 90-180 days after overbid). Subsequent overbids are associated with smaller epsilon values, bar some outliers.

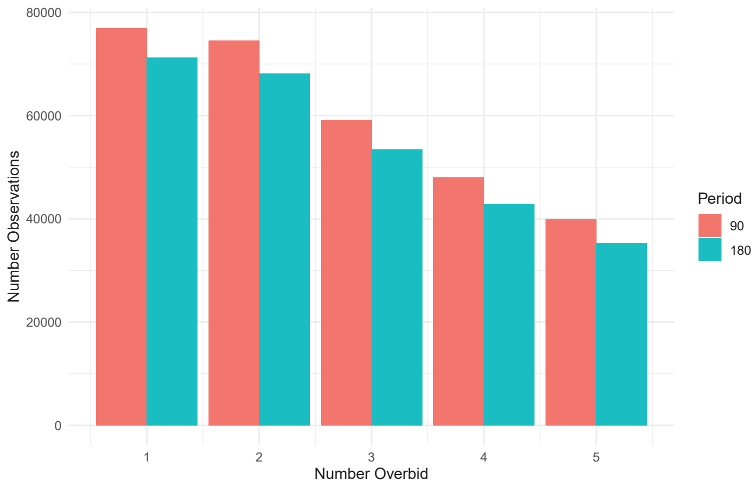



Figure 16: Number of Observations First Five Overbids

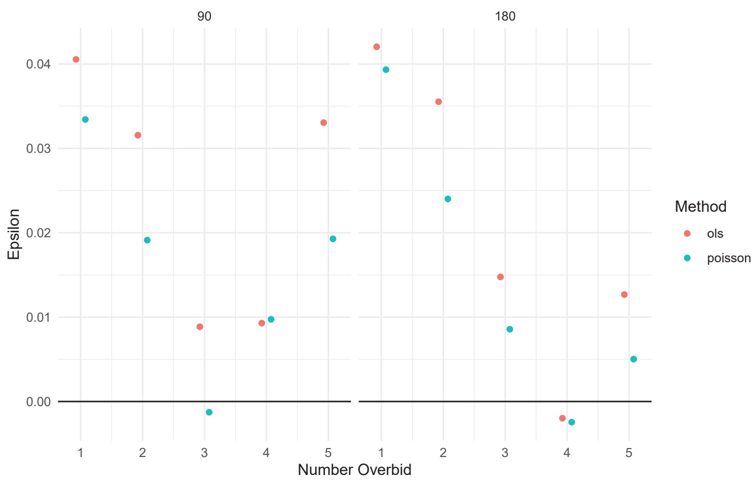



Figure 17: Epsilon First Five Overbids

# B Appendix: Managerial Overconfidence in Europe

## B.1 Replication Files

The replication files are available from the authors or at `https://www.dropbox.com/scl/fo/m3g3d24uuglpvr91zkh4d/h?rlkey=dyennumcdltsoala319wwt8tl&dl=0`

## B.2 Data Preparation

### B.2.1 BoardEx Data

We use data on director compensation packages from the "BoardEx Core Reports - Compensation" of Europe and the UK. The option-based overconfidence classification approaches require information on vesting dates, expiry dates, and strike prices of option grants. The BoardEx data provides this information. We assume that options vest at the report date of the information whenever the vesting date is unavailable and the option is listed as "exercisable". We exclude option grants with missing information on the vesting or expiry dates. Options granted as part of a long-term incentive plan are listed separately in the BoardEx data but not treated differently in our data set.

The data contains detailed information on compensation packages, including salary, bonus, pension, share grants, option grants, and other cash payments (e.g., relocation costs and fringe benefits). The raw data is provided in separate files for the "main actors" of the firm and for further senior management disclosed earners and also split between Europe and the UK. We combine this information in the first step. The data set contains multiple entries for individuals within a company in a year due to updates in the company reports, when individuals receive separate remuneration for different functions within the firm, or when individuals are present in multiple raw data sets. We treat these cases in the following way: First, we select the maximum of each compensation component among the last available information for each person's position within a company in a year. We choose the maximum because we believe that omissions of information are the most likely explanation for differences.[84] We then aggregate information over different positions of a person in a company within a year.

The fraction of shares held by the CEO and the number of options vested in the first six months of the year relative to the total number of shares outstanding are used as control variables in the analyses of Sections 2.4 and 2.5. BoardEx reports the total value of shares held by an individual. We use the maximum when there are multiple observations for a person within a company and a year. We divide this information by the market capitalization provided in the Amadeus data to obtain the fraction of shares held by the individual. The number of vested options is based on aggregating individual option grants divided by the shares outstanding

[84]Typically, there are only minor differences due to individuals being reported in the EU, and the UK data set and exchange rate fluctuations.

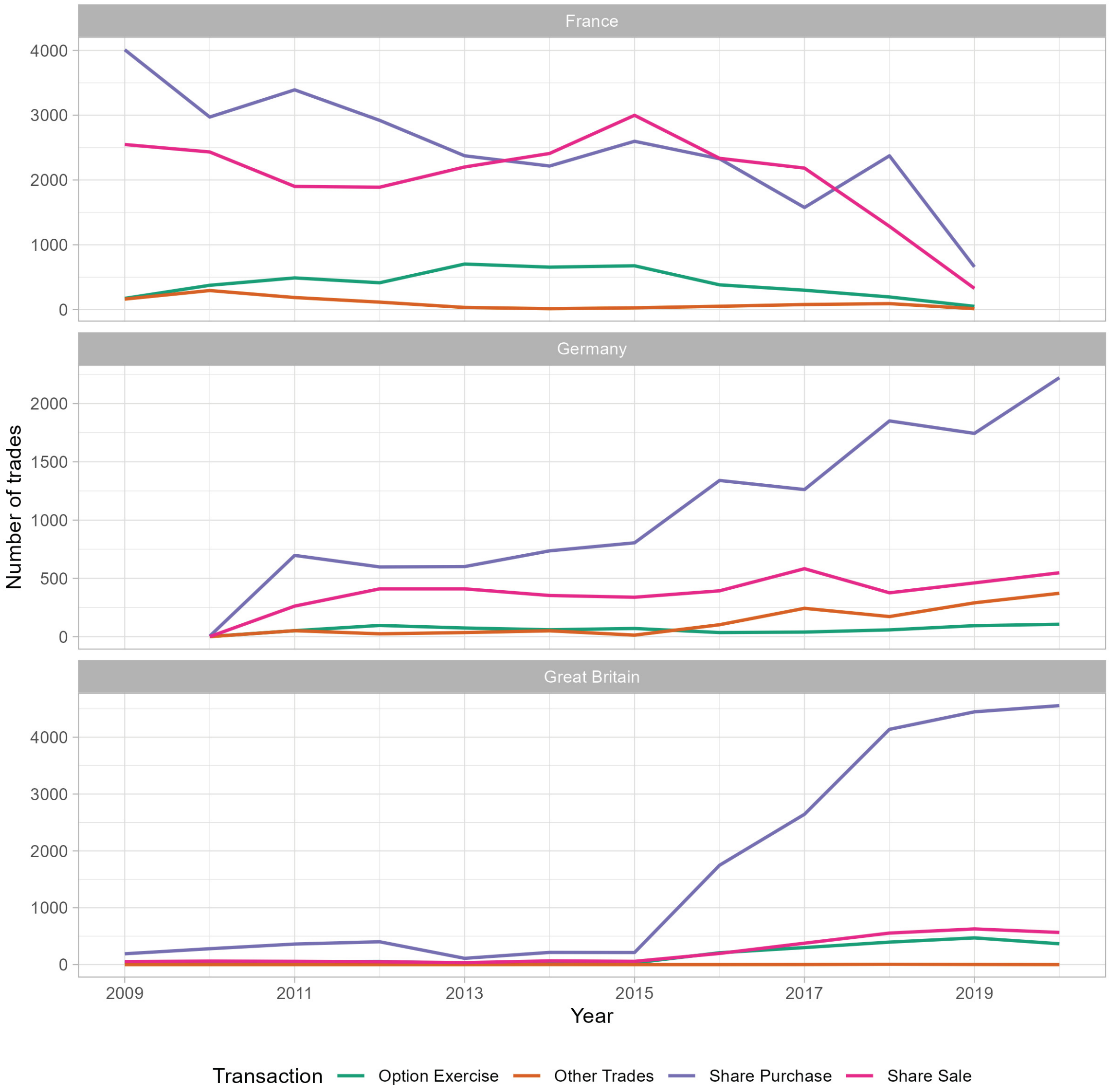
provided in the Amadeus data set. We assume that immediately vested options are exercisable from the beginning of the year they were granted.

### B.2.2 Amadeus Data

The Amadeus data preparation follows a three-step procedure. First, we drop values with obvious reporting mistakes. Specifically, we drop values of total assets, total assets per employee, tangible fixed assets, intangible fixed assets, fixed assets, other fixed assets, current assets, operating revenue, and sales that are below 1000 Euro, and values of Tobin's Q, cost of goods sold, and long-term debt below zero. Further, we drop values of the number of employees that are below one or above 2000000. In the second step, we verify the internal consistency of the balance sheet data. We calculate the ratio between aggregated information and the sum of their individual components, and drop values, when they differ by more than two percent. Specifically, we verify the consistency of fixed assets, total assets, profits after tax, total assets per employee, and Tobin's Q. Finally, we fill gaps in the data of three years and less. We linearly interpolate the variables cash flow, cash and cash equivalents, depreciation and amortization, EBIT, long-term debt, market capitalization, current liabilities, other current liabilities, non-current liabilities, other non-current liabilities, number of employees, operating revenue, profits after tax, profits before tax, sales, shareholder funds, tangible fixed assets, taxation, Tobin's Q, and total assets. Moreover, we fill gaps in the variables shares outstanding and nominal values downwards and replace missing values with zeros in the variables extraordinary profits and losses.

## B.3 Descriptive Statistics

Figure 18: Number of Directors' Dealings per Year, by Nature of Transaction and Country.




*Source:* Scraped data about directors' dealings, own illustration.
*Notes:* This figure shows the number of directors' dealings over time split by type of financial instrument and reported separately for each country.

## B.4 Correlation by Country

Table 21: Correlation Matrix - France.

| | Optimism | Overconfidence | Net-Buyer (GSZ) | Net-Buyer (MT) | Longholder | Holder67 |
|---|---|---|---|---|---|---|
| Optimism | 1 | | | | | |
| Overconfidence | -0.117 | 1 | | | | |
| Net-Buyer (GSZ) | 0.078 | 0.104 | 1 | | | |
| Net-Buyer (MT) | | | 0.375 | 1 | | |
| Longholder | | 0.268 | 0.153 | | 1 | |
| Holder67 | | | -0.394 | | -0.121 | 1 |

Note: Correlation coefficients are rounded to the third decimal.

Table 22: Correlation Matrix - Germany.

| | Optimism | Overconfidence | Net-Buyer (GSZ) | Net-Buyer (MT) | Longholder | Holder67 |
|---|---|---|---|---|---|---|
| Optimism | 1 | | | | | |
| Overconfidence | -0.101 | 1 | | | | |
| Net-Buyer (GSZ) | -0.115 | 0.078 | 1 | | | |
| Net-Buyer (MT) | | | -0.165 | 1 | | |
| Longholder | 0.196 | -0.069 | -0.061 | 0.225 | 1 | |
| Holder67 | 0.002 | 0.028 | -0.094 | 0.095 | -0.129 | 1 |

Note: Correlation coefficients are rounded to the third decimal.

Table 23: Correlation Matrix - UK.

| | Optimism | Overconfidence | Net-Buyer (GSZ) | Net-Buyer (MT) | Longholder | Holder67 |
|---|---|---|---|---|---|---|
| Optimism | 1 | | | | | |
| Overconfidence | -0.101 | 1 | | | | |
| Net-Buyer (GSZ) | -0.115 | 0.078 | 1 | | | |
| Net-Buyer (MT) | | | -0.165 | 1 | | |
| Longholder | 0.196 | -0.069 | -0.061 | 0.225 | 1 | |
| Holder67 | 0.002 | 0.028 | -0.094 | 0.095 | -0.129 | 1 |

Note: Correlation coefficients are rounded to the third decimal.

## B.5 Investment Cash Flow Sensitivity and Overconfidence

Table 24: Investment Cash Flow Sensitivity and Overconfidence - Holder67

| Dep. Var. | Investment | | | | |
|---|---|---|---|---|---|
| Model: | (1) | (2) | (3) | (4) | (5) |
| *Variables* | | | | | |
| Holder67 | | | -0.013* | -0.014* | -0.014** |
| | | | (0.007) | (0.007) | (0.007) |
| Cash Fl. × Holder67 | | | 0.059 | 0.074* | 0.076* |
| | | | (0.043) | (0.042) | (0.044) |
| Cash Fl. | 0.115** | -0.256 | 0.070 | -0.312* | -0.279 |
| | (0.049) | (0.190) | (0.058) | (0.187) | (0.191) |
| Lag Q | 0.007** | 0.006 | 0.008** | 0.006 | 0.006* |
| | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) |
| Stocks | | 0.080 | | 0.083 | 0.089 |
| | | (0.114) | | (0.115) | (0.115) |
| Options | | -0.284 | | -0.183 | -0.210 |
| | | (0.213) | | (0.208) | (0.219) |
| Size | | -0.022*** | | -0.022*** | -0.022*** |
| | | (0.007) | | (0.007) | (0.008) |
| Eff. Board Size | | 0.012 | | 0.011 | 0.011 |
| | | (0.015) | | (0.015) | (0.015) |
| Cash Fl. × Lag Q | | -0.003 | | -0.005 | -0.004 |
| | | (0.011) | | (0.011) | (0.011) |
| Cash Fl. × Stocks | | -0.886 | | -0.872 | -0.969 |
| | | (0.674) | | (0.669) | (0.692) |
| Cash Fl. × Options | | 0.654 | | -0.089 | 0.048 |
| | | (1.314) | | (1.280) | (1.320) |
| Cash Fl. × Size | | 0.029* | | 0.029* | 0.027 |
| | | (0.016) | | (0.016) | (0.019) |
| Cash Fl. × Eff. Board Size | | 0.097 | | 0.109* | 0.110* |
| | | (0.065) | | (0.064) | (0.066) |
| *Fixed-effects* | | | | | |
| Firm (342) | Yes | Yes | Yes | Yes | Yes |
| Year (9) | Yes | Yes | Yes | Yes | Yes |
| Year × Cash Fl. | Yes | Yes | Yes | Yes | Yes |
| Industry × Cash Fl. | No | No | No | No | Yes |
| *Fit statistics* | | | | | |
| Observations | 1,592 | 1,592 | 1,592 | 1,592 | 1,592 |
| Adjusted $R^2$ | 0.434 | 0.448 | 0.435 | 0.450 | 0.448 |

*Notes:* The dependent variable in all columns is investments, defined as the change in tangible fixed assets plus depreciation and amortization and normalized by lagged total assets. Cash flow is profit after tax plus depreciation and amortization, normalized by lagged total assets. Tobin's Q is lagged market capitalization by total assets. Stock ownership is the share of company stocks held by the CEO at the beginning of the year, normalized by market capitalization. Vested options are the number of options held by the CEO that are exercisable within the first six months of the year, normalized by the number of outstanding shares. Size is logged and lagged total assets. Efficient board size is a dummy that takes the value one when the board size is between four and twelve. Industry fixed effects are based on 12 Fama–French industry groups. *Signif. Codes:* ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered on the firm-level.

Table 25: Investment Cash Flow Sensitivity and Overconfidence - Longholder

| Dep. Var. | Investment | | | | |
|---|---|---|---|---|---|
| Model: | (1) | (2) | (3) | (4) | (5) |
| *Variables* | | | | | |
| Longholder | | | -0.007 | -0.012 | -0.011 |
| | | | (0.009) | (0.007) | (0.007) |
| Cash Fl. × Longholder | | | 0.008 | 0.041 | 0.045 |
| | | | (0.045) | (0.041) | (0.042) |
| Cash Fl. | 0.114** | -0.431** | 0.106* | -0.485*** | -0.462** |
| | (0.047) | (0.177) | (0.054) | (0.184) | (0.197) |
| Lag Q | 0.008** | 0.006 | 0.008** | 0.006 | 0.005 |
| | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) |
| Stocks | | 0.067 | | 0.058 | 0.056 |
| | | (0.107) | | (0.108) | (0.108) |
| Options | | -0.389* | | -0.381* | -0.406* |
| | | (0.216) | | (0.216) | (0.227) |
| Size | | -0.023*** | | -0.024*** | -0.023*** |
| | | (0.007) | | (0.007) | (0.007) |
| Eff. Board Size | | 0.004 | | 0.004 | 0.003 |
| | | (0.009) | | (0.009) | (0.009) |
| Cash Fl. × Lag Q | | 0.003 | | 0.004 | 0.004 |
| | | (0.011) | | (0.011) | (0.011) |
| Cash Fl. × Stocks | | -0.807 | | -0.794 | -0.812 |
| | | (0.641) | | (0.656) | (0.663) |
| Cash Fl. × Options | | -0.245 | | -0.227 | -0.153 |
| | | (1.074) | | (1.080) | (1.133) |
| Cash Fl. × Size | | 0.047*** | | 0.048*** | 0.049*** |
| | | (0.015) | | (0.015) | (0.018) |
| Cash Fl. × Eff. Board Size | | 0.082 | | 0.087 | 0.086 |
| | | (0.057) | | (0.058) | (0.060) |
| *Fixed-effects* | | | | | |
| Firm (399) | Yes | Yes | Yes | Yes | Yes |
| Year (9) | Yes | Yes | Yes | Yes | Yes |
| Year × Cash Fl. | Yes | Yes | Yes | Yes | Yes |
| Industry × Cash Fl. | No | No | No | No | Yes |
| *Fit statistics* | | | | | |
| Observations | 1,843 | 1,843 | 1,843 | 1,843 | 1,843 |
| Adjusted $R^2$ | 0.444 | 0.468 | 0.444 | 0.468 | 0.469 |

*Notes:* The dependent variable in all columns is investments, defined as the change in tangible fixed assets plus depreciation and amortization and normalized by lagged total assets. Cash flow is profit after tax plus depreciation and amortization, normalized by lagged total assets. Tobin's Q is lagged market capitalization by total assets. Stock ownership is the share of company stocks held by the CEO at the beginning of the year, normalized by market capitalization. Vested options are the number of options held by the CEO that are exercisable within the first six months of the year, normalized by the number of outstanding shares. Size is logged and lagged total assets. Efficient board size is a dummy that takes the value one when the board size is between four and twelve. Industry fixed effects are based on 12 Fama–French industry groups. *Signif. Codes:* ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered on the firm-level.

Table 26: Investment Cash Flow Sensitivity and Overconfidence - Net-Buyer

| Dep. Var. | Investment | | | | |
|---|---|---|---|---|---|
| Model: | (1) | (2) | (3) | (4) | (5) |
| *Variables* | | | | | |
| Cash Fl. | 0.095*** | -0.206 | 0.049 | -0.291* | -0.262 |
| | (0.031) | (0.167) | (0.036) | (0.165) | (0.164) |
| Lag Q | 0.005** | 0.004 | 0.006** | 0.005* | 0.005* |
| | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) |
| Stocks | | 0.076 | | 0.067 | 0.018 |
| | | (0.085) | | (0.081) | (0.077) |
| Options | | 0.230 | | 0.212 | 0.216 |
| | | (0.206) | | (0.207) | (0.212) |
| Size | | -0.012* | | -0.012* | -0.011* |
| | | (0.006) | | (0.006) | (0.007) |
| Eff. Board Size | | 0.005 | | 0.005 | 0.005 |
| | | (0.007) | | (0.006) | (0.007) |
| Cash Fl. × Lag Q | | -0.006 | | -0.007 | -0.012 |
| | | (0.012) | | (0.012) | (0.012) |
| Cash Fl. × Stocks | | -0.497 | | -0.434 | -0.089 |
| | | (0.508) | | (0.479) | (0.447) |
| Cash Fl. × Options | | 0.368 | | 0.578 | 0.962 |
| | | (1.065) | | (1.013) | (1.012) |
| Cash Fl. × Size | | 0.027** | | 0.030** | 0.031** |
| | | (0.013) | | (0.013) | (0.013) |
| Cash Fl. × Eff. Board Size | | 0.041 | | 0.044 | 0.027 |
| | | (0.052) | | (0.048) | (0.051) |
| Net Buyer | | | -0.009 | -0.011* | -0.014** |
| | | | (0.006) | (0.006) | (0.006) |
| Cash Fl. × Net Buyer | | | 0.091** | 0.102** | 0.138*** |
| | | | (0.042) | (0.041) | (0.040) |
| *Fixed-effects* | | | | | |
| Firm (351) | Yes | Yes | Yes | Yes | Yes |
| Year (9) | Yes | Yes | Yes | Yes | Yes |
| Year × Cash Fl. | Yes | Yes | Yes | Yes | Yes |
| Industry × Cash Fl. | No | No | No | No | Yes |
| *Fit statistics* | | | | | |
| Observations | 1,747 | 1,747 | 1,747 | 1,747 | 1,747 |
| Adjusted $R^2$ | 0.479 | 0.486 | 0.481 | 0.488 | 0.494 |

*Notes:* The dependent variable in all columns is investments, defined as the change in tangible fixed assets plus depreciation and amortization and normalized by lagged total assets. Cash flow is profit after tax plus depreciation and amortization, normalized by lagged total assets. Tobin's Q is lagged market capitalization by total assets. Stock ownership is the share of company stocks held by the CEO at the beginning of the year, normalized by market capitalization. Vested options are the number of options held by the CEO that are exercisable within the first six months of the year, normalized by the number of outstanding shares. Size is logged and lagged total assets. Efficient board size is a dummy that takes the value one when the board size is between four and twelve. Industry fixed effects are based on 12 Fama–French industry groups. *Signif. Codes:* ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered on the firm-level.

Table 27: Investment Cash Flow Sensitivity and Overconfidence - Overconfidence + Optimism

| Dep. Var. | | | Investment | | |
|---|---|---|---|---|---|
| Model: | (1) | (2) | (3) | (4) | (5) |
| *Variables* | | | | | |
| Overconfidence | | | 0.031 | 0.043 | 0.044 |
| | | | (0.038) | (0.038) | (0.037) |
| Cash Fl. × Overconfidence | | | -0.115 | -0.040 | -0.086 |
| | | | (0.112) | (0.123) | (0.156) |
| Optimism | | | 0.006 | 0.005 | 0.002 |
| | | | (0.027) | (0.026) | (0.026) |
| Cash Fl. × Optimism | | | -0.009 | -0.007 | 0.065 |
| | | | (0.111) | (0.105) | (0.151) |
| Cash Fl. | 0.108* | -0.548* | 0.110* | -0.590** | -0.684* |
| | (0.055) | (0.283) | (0.058) | (0.287) | (0.354) |
| Lag Q | 0.005 | 0.006 | 0.004 | 0.005 | 0.004 |
| | (0.004) | (0.005) | (0.004) | (0.005) | (0.005) |
| Stocks | | -0.039 | | -0.159 | -0.199 |
| | | (0.242) | | (0.209) | (0.208) |
| Options | | -0.272 | | -0.225 | -0.182 |
| | | (0.393) | | (0.401) | (0.410) |
| Size | | -0.007 | | -0.013 | -0.015 |
| | | (0.017) | | (0.015) | (0.015) |
| Eff. Board Size | | 0.009 | | 0.009 | 0.007 |
| | | (0.007) | | (0.008) | (0.008) |
| Cash Fl. × Lag Q | | 0.034** | | 0.036** | 0.022 |
| | | (0.017) | | (0.017) | (0.018) |
| Cash Fl. × Stocks | | -0.836 | | -0.502 | -0.001 |
| | | (0.562) | | (0.636) | (0.781) |
| Cash Fl. × Options | | 1.797 | | 1.522 | 2.030 |
| | | (1.750) | | (2.126) | (2.011) |
| Cash Fl. × Size | | 0.063*** | | 0.064*** | 0.083*** |
| | | (0.023) | | (0.024) | (0.030) |
| Cash Fl. × Eff. Board Size | | -0.069 | | -0.056 | -0.042 |
| | | (0.061) | | (0.067) | (0.064) |
| *Fixed-effects* | | | | | |
| Firm (94) | Yes | Yes | Yes | Yes | Yes |
| Year (9) | Yes | Yes | Yes | Yes | Yes |
| Year × Cash Fl. | Yes | Yes | Yes | Yes | Yes |
| Industry × Cash Fl. | No | No | No | No | Yes |
| *Fit statistics* | | | | | |
| Observations | 447 | 447 | 447 | 447 | 447 |
| Adjusted $R^2$ | 0.481 | 0.520 | 0.481 | 0.521 | 0.543 |

*Notes:* The dependent variable in all columns is investments, defined as the change in tangible fixed assets plus depreciation and amortization and normalized by lagged total assets. Cash flow is profit after tax plus depreciation and amortization, normalized by lagged total assets. Tobin's Q is lagged market capitalization by total assets. Stock ownership is the share of company stocks held by the CEO at the beginning of the year, normalized by market capitalization. Vested options are the number of options held by the CEO that are exercisable within the first six months of the year, normalized by the number of outstanding shares. Size is logged and lagged total assets. Efficient board size is a dummy that takes the value one when the board size is between four and twelve. Industry fixed effects are based on 12 Fama–French industry groups. *Signif. Codes:* ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered on the firm-level.

Table 28: Investment Cash Flow Sensitivity and Overconfidence - Holder67

| Dep. Var. | | | Investment | | |
|---|---|---|---|---|---|
| Model: | (1) | (2) | (3) | (4) | (5) |
| *Variables* | | | | | |
| Holder67 | | | -0.012* | -0.013* | -0.014* |
| | | | (0.007) | (0.007) | (0.007) |
| Cash Fl. × Holder67 | | | 0.059 | 0.075* | 0.077* |
| | | | (0.043) | (0.042) | (0.044) |
| Cash Fl. | 0.113** | -0.243 | 0.067 | -0.300 | -0.269 |
| | (0.049) | (0.190) | (0.059) | (0.187) | (0.191) |
| Lag Q | 0.007** | 0.006 | 0.008** | 0.006 | 0.006 |
| | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) |
| Stocks | | 0.078 | | 0.081 | 0.087 |
| | | (0.114) | | (0.115) | (0.115) |
| Options | | -0.270 | | -0.174 | -0.199 |
| | | (0.214) | | (0.208) | (0.219) |
| Size | | -0.022*** | | -0.022*** | -0.023*** |
| | | (0.007) | | (0.007) | (0.008) |
| Eff. Board Size | | 0.011 | | 0.011 | 0.010 |
| | | (0.015) | | (0.015) | (0.015) |
| Cash Fl. × Lag Q | | -0.003 | | -0.005 | -0.003 |
| | | (0.011) | | (0.011) | (0.011) |
| Cash Fl. × Stocks | | -0.880 | | -0.868 | -0.951 |
| | | (0.675) | | (0.670) | (0.694) |
| Cash Fl. × Options | | 0.592 | | -0.156 | -0.014 |
| | | (1.316) | | (1.282) | (1.321) |
| Cash Fl. × Size | | 0.026 | | 0.026 | 0.025 |
| | | (0.017) | | (0.016) | (0.019) |
| Cash Fl. × Eff. Board Size | | 0.107 | | 0.119* | 0.121* |
| | | (0.067) | | (0.065) | (0.068) |
| *Fixed-effects* | | | | | |
| Firm (338) | Yes | Yes | Yes | Yes | Yes |
| Year (9) | Yes | Yes | Yes | Yes | Yes |
| Year × Cash Fl. | Yes | Yes | Yes | Yes | Yes |
| Industry × Cash Fl. | No | No | No | No | Yes |
| *Fit statistics* | | | | | |
| Observations | 1,576 | 1,576 | 1,576 | 1,576 | 1,576 |
| Adjusted $R^2$ | 0.433 | 0.447 | 0.435 | 0.449 | 0.447 |

*Notes:* The dependent variable in all columns is investments, defined as the change in tangible fixed assets plus depreciation and amortization and normalized by lagged total assets. Cash flow is profit after tax plus depreciation and amortization, normalized by lagged total assets. Tobin's Q is lagged market capitalization by total assets. Stock ownership is the share of company stocks held by the CEO at the beginning of the year, normalized by market capitalization. Vested options are the number of options held by the CEO that are exercisable within the first six months of the year, normalized by the number of outstanding shares. Size is logged and lagged total assets. Efficient board size is a dummy that takes the value one when the board size is between four and twelve. Industry fixed effects are based on 12 Fama–French industry groups. *Signif. Codes:* ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered on the firm-level.

## B.6 Overconfidence and Innovation

Table 29: Innovation and Overconfidence - Holder67

| Dep. Var. | Total Citations (Adjusted) | |
|---|---|---|
| Model: | (1) | (2) |
| *Variables* | | |
| Holder67 | 0.265*** | 0.157 |
| | (0.098) | (0.109) |
| ln(Lag Sales+1) | 0.863*** | -0.142 |
| | (0.266) | (0.106) |
| ln(Lag C/L+1) | 0.283 | 0.517*** |
| | (0.204) | (0.128) |
| Stocks | -24.908 | -13.162*** |
| | (20.040) | (3.406) |
| Options | 0.976 | -0.789 |
| | (0.662) | (0.920) |
| ln(Lag RD Stock+1) | | 0.913*** |
| | | (0.249) |
| CEO Controls | Yes | Yes |
| *Fixed-effects* | | |
| Country (2) | Yes | Yes |
| Industry (18) | Yes | Yes |
| Year (9) | Yes | Yes |
| Firm effects (BGV) | Yes | Yes |
| *Fit statistics* | | |
| Observations | 671 | 671 |
| Pseudo $R^2$ | 0.923 | 0.941 |

*Notes:* The dependent variable in all columns is citation-weighted patent counts, normalized by the mean number of citations across firms in this year. We add the value one before taking logs. Sales is logged and lagged operating revenue. The capital labor ratio is computed as the log of the difference between lagged total assets and lagged number of employees. Stock ownership is the share of company stocks held by the CEO at the beginning of the year, normalized by market capitalization. Vested options are the number of options held by the CEO that are exercisable within the first six months of the year, normalized by the number of outstanding shares. R&D Stock are constructed following Hall (1990). For details, see Section 1.3. We use a set of CEO controls, such as age, tenure, and their respective quadratic terms. Industry fixed effects are based on SIC two-digit industry codes. BGV (Blundell et al., 1999) firm effects are average citation-weighted patent counts over the ten years preceding our sample as a control variable. *Signif. Codes:* ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered on the industry-level.

Table 30: Innovation and Overconfidence - Longholder

| Dep. Var. | Total Citations (Adjusted) | |
|---|---|---|
| Model: | (1) | (2) |
| *Variables* | | |
| Longholder | 2.666*** | 1.769*** |
| | (0.302) | (0.266) |
| ln(Lag Sales+1) | 0.680*** | -0.188 |
| | (0.247) | (0.172) |
| ln(Lag C/L+1) | 1.086*** | 0.467*** |
| | (0.062) | (0.144) |
| Stocks | -8.849 | -0.342 |
| | (17.612) | (7.778) |
| Options | -26.490 | -5.614 |
| | (51.684) | (9.545) |
| ln(Lag RD Stock+1) | | 1.078*** |
| | | (0.286) |
| CEO Controls | Yes | Yes |
| *Fixed-effects* | | |
| Country (2) | Yes | Yes |
| Industry (20) | Yes | Yes |
| Year (9) | Yes | Yes |
| Firm effects (BGV) | Yes | Yes |
| *Fit statistics* | | |
| Observations | 849 | 849 |
| Pseudo $R^2$ | 0.889 | 0.924 |

*Notes:* The dependent variable in all columns is citation-weighted patent counts, normalized by the mean number of citations across firms in this year. We add the value one before taking logs. Sales is logged and lagged operating revenue. The capital labor ratio is computed as the log of the difference between lagged total assets and lagged number of employees. Stock ownership is the share of company stocks held by the CEO at the beginning of the year, normalized by market capitalization. Vested options are the number of options held by the CEO that are exercisable within the first six months of the year, normalized by the number of outstanding shares. R&D Stock are constructed following Hall (1990). For details, see Section 1.3. We use a set of CEO controls, such as age, tenure, and their respective quadratic terms. Industry fixed effects are based on SIC two-digit industry codes. BGV (Blundell et al., 1999) firm effects are average citation-weighted patent counts over the ten years preceding our sample as a control variable. *Signif. Codes:* ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered on the industry-level.

Table 31: Innovation and Overconfidence - Net-Buyer

| Dep. Var. | Total Citations (Adjusted) | |
|---|---|---|
| Model: | (1) | (2) |
| *Variables* | | |
| Net Buyer | 2.905*** | 1.468*** |
| | (0.514) | (0.248) |
| ln(Lag Sales+1) | 0.779*** | -0.521*** |
| | (0.291) | (0.187) |
| ln(Lag C/L+1) | 0.991** | -0.087 |
| | (0.403) | (0.321) |
| Stocks | 3.454 | 3.372 |
| | (15.023) | (7.017) |
| Options | 1.312 | -0.713 |
| | (0.870) | (0.938) |
| ln(Lag RD Stock+1) | | 1.645*** |
| | | (0.276) |
| CEO Controls | Yes | Yes |
| *Fixed-effects* | | |
| Country (2) | Yes | Yes |
| Industry (22) | Yes | Yes |
| Year (9) | Yes | Yes |
| Firm effects (BGV) | Yes | Yes |
| *Fit statistics* | | |
| Observations | 856 | 856 |
| Pseudo $R^2$ | 0.868 | 0.918 |

*Notes:* The dependent variable in all columns is citation-weighted patent counts, normalized by the mean number of citations across firms in this year. We add the value one before taking logs. Sales is logged and lagged operating revenue. The capital labor ratio is computed as the log of the difference between lagged total assets and lagged number of employees. Stock ownership is the share of company stocks held by the CEO at the beginning of the year, normalized by market capitalization. Vested options are the number of options held by the CEO that are exercisable within the first six months of the year, normalized by the number of outstanding shares. R&D Stock are constructed following Hall (1990). For details, see Section 1.3. We use a set of CEO controls, such as age, tenure, and their respective quadratic terms. Industry fixed effects are based on SIC two-digit industry codes. BGV (Blundell et al., 1999) firm effects are average citation-weighted patent counts over the ten years preceding our sample as a control variable. *Signif. Codes:* ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered on the industry-level.

Table 32: Innovation and Overconfidence - Overconfidence and Optimism

| Dep. Var. | Total Citations (Adjusted) | |
|---|---|---|
| Model: | (1) | (2) |
| *Variables* | | |
| Overconfidence | -1.446*** | -0.309 |
| | (0.321) | (2.968) |
| ln(Lag Sales+1) | -0.189 | -0.179 |
| | (0.136) | (0.245) |
| ln(Lag C/L+1) | -0.250 | -0.743 |
| | (0.158) | (0.642) |
| Stocks | -70.064 | -43.391 |
| | (88.535) | (123.203) |
| Options | -7.686 | 28.699 |
| | (91.645) | (205.086) |
| ln(Lag RD Stock+1) | | 2.895 |
| | | (3.348) |
| CEO Controls | Yes | Yes |
| *Fixed-effects* | | |
| Country (2) | Yes | Yes |
| Industry (9) | Yes | Yes |
| Year (9) | Yes | Yes |
| Firm effects (BGV) | Yes | Yes |
| *Fit statistics* | | |
| Observations | 103 | 103 |
| Pseudo $R^2$ | 0.952 | 0.954 |

*Notes:* The dependent variable in all columns is citation-weighted patent counts, normalized by the mean number of citations across firms in this year. We add the value one before taking logs. Sales is logged and lagged operating revenue. The capital labor ratio is computed as the log of the difference between lagged total assets and lagged number of employees. Stock ownership is the share of company stocks held by the CEO at the beginning of the year, normalized by market capitalization. Vested options are the number of options held by the CEO that are exercisable within the first six months of the year, normalized by the number of outstanding shares. R&D Stock are constructed following Hall (1990). For details, see Section 1.3. We use a set of CEO controls, such as age, tenure, and their respective quadratic terms. Industry fixed effects are based on SIC two-digit industry codes. BGV (Blundell et al., 1999) firm effects are average citation-weighted patent counts over the ten years preceding our sample as a control variable. *Signif. Codes:* ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered on the industry-level.

Table 33: Innovation and Overconfidence - Holder67

| Dep. Var. | Total Citations (Adjusted) | |
|---|---|---|
| Model: | (1) | (2) |
| *Variables* | | |
| Holder67 | -0.326 | -0.326 |
| | (0.499) | (0.492) |
| ln(Lag Sales+1) | -0.039 | -0.091 |
| | (0.067) | (0.106) |
| ln(Lag C/L+1) | 0.810** | 0.775* |
| | (0.353) | (0.401) |
| Stocks | -18.284** | -17.008*** |
| | (8.949) | (6.403) |
| Options | -0.283 | -0.193 |
| | (0.879) | (0.732) |
| ln(Lag RD Stock+1) | | 0.149 |
| | | (0.218) |
| CEO Controls | Yes | Yes |
| *Fixed-effects* | | |
| Country (1) | Yes | Yes |
| Industry (18) | Yes | Yes |
| Year (9) | Yes | Yes |
| Firm effects (BGV) | Yes | Yes |
| *Fit statistics* | | |
| Observations | 593 | 593 |
| Pseudo $R^2$ | 0.824 | 0.825 |

*Notes:* The dependent variable in all columns is citation-weighted patent counts, normalized by the mean number of citations across firms in this year. We add the value one before taking logs. Sales is logged and lagged operating revenue. The capital labor ratio is computed as the log of the difference between lagged total assets and lagged number of employees. Stock ownership is the share of company stocks held by the CEO at the beginning of the year, normalized by market capitalization. Vested options are the number of options held by the CEO that are exercisable within the first six months of the year, normalized by the number of outstanding shares. R&D Stock are constructed following Hall (1990). For details, see Section 1.3. We use a set of CEO controls, such as age, tenure, and their respective quadratic terms. Industry fixed effects are based on SIC two-digit industry codes. BGV (Blundell et al., 1999) firm effects are average citation-weighted patent counts over the ten years preceding our sample as a control variable. *Signif. Codes:* ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered on the industry-level.

## B.7 Compensation and Overconfidence

Table 34: Compensation and Overconfidence - Holder67

| Dep. Var.<br>Model: | ln(Options+1)<br>(1) | ln(Stocks+1)<br>(2) | ln(Bonus+1)<br>(3) | ln(Salary+1)<br>(4) | ln(Total+1)<br>(5) |
|---|---|---|---|---|---|
| *Variables* | | | | | |
| Holder67 | 0.333 | 0.206 | -0.153 | 0.140** | 0.133** |
| | (0.225) | (0.166) | (0.150) | (0.057) | (0.062) |
| ln(Mkt Cap) | 0.431*** | 0.179 | 0.759*** | 0.127** | 0.313*** |
| | (0.139) | (0.136) | (0.129) | (0.055) | (0.057) |
| SD Return | 0.023** | -0.001 | -0.007 | -0.011*** | -0.007** |
| | (0.012) | (0.012) | (0.010) | (0.003) | (0.003) |
| Leverage | -0.030 | 0.017 | -0.118 | -0.407*** | -0.383*** |
| | (0.134) | (0.122) | (0.176) | (0.094) | (0.105) |
| MtB Ratio | 0.005 | -0.017 | -0.005 | -0.053** | -0.062** |
| | (0.065) | (0.067) | (0.054) | (0.024) | (0.030) |
| Cash Fl. | -0.545** | -0.006 | 0.360 | 0.017 | -0.124 |
| | (0.266) | (0.138) | (0.238) | (0.132) | (0.100) |
| RD Expenses | 0.121 | -0.536 | 1.540** | 0.303 | 0.265 |
| | (1.090) | (0.689) | (0.679) | (0.332) | (0.399) |
| RD Missing | 0.157 | 0.012 | 0.673*** | 0.057 | 0.074 |
| | (0.272) | (0.164) | (0.248) | (0.088) | (0.071) |
| Ex. Board Size | -0.075 | -0.085 | -0.099 | -0.025 | -0.045 |
| | (0.093) | (0.093) | (0.069) | (0.026) | (0.028) |
| EBIT | 0.099 | -0.127 | 0.051 | -0.078 | 0.042 |
| | (0.316) | (0.191) | (0.245) | (0.196) | (0.180) |
| Stock Return | 0.043 | 0.021 | 0.231** | 0.017 | 0.083 |
| | (0.113) | (0.061) | (0.109) | (0.036) | (0.051) |
| *Fixed-effects* | | | | | |
| Firm (383) | Yes | Yes | Yes | Yes | Yes |
| Year (9) | Yes | Yes | Yes | Yes | Yes |
| Age (47) | Yes | Yes | Yes | Yes | Yes |
| Tenure (35) | Yes | Yes | Yes | Yes | Yes |
| Gender (2) | Yes | Yes | Yes | Yes | Yes |
| *Fit statistics* | | | | | |
| Observations | 2,274 | 2,274 | 2,274 | 2,274 | 2,274 |
| Adjusted $R^2$ | 0.331 | 0.733 | 0.606 | 0.627 | 0.737 |

*Notes:* The dependent variables are logged compensation components, indicated in the column headers. We add the value one before taking logs. Market capitalization is calculated as the product of the number of outstanding shares and the end-of-year market price. Standard deviation of stock returns is calculated based on monthly returns over the previous five years and is divided by 100. Leverage is long-term debt divided by total assets, Market-to-book ratio is the ratio of market capitalization and total assets. Cash flow is profit after tax plus depreciation and amortization, normalized by lagged total assets. R&D expenses are normalized by total assets. R&D Missing is a dummy indicating whether R&D expenses are missing. Executive board size is the number of executive directors on the board. EBIT is normalized by total assets. *Signif. Codes:* ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered on the firm-level.

Table 35: Compensation and Overconfidence - Longholder

| Dep. Var.<br>Model: | ln(Options+1)<br>(1) | ln(Stocks+1)<br>(2) | ln(Bonus+1)<br>(3) | ln(Salary+1)<br>(4) | ln(Total+1)<br>(5) |
|---|---|---|---|---|---|
| *Variables* | | | | | |
| Longholder | 0.514 | -0.543* | 0.056 | -0.151* | -0.036 |
| | (0.345) | (0.313) | (0.275) | (0.082) | (0.075) |
| ln(Mkt Cap) | 0.460*** | 0.143 | 0.737*** | 0.110** | 0.310*** |
| | (0.130) | (0.118) | (0.121) | (0.051) | (0.052) |
| SD Return | 0.023** | 0.003 | -0.005 | -0.009*** | -0.007*** |
| | (0.011) | (0.011) | (0.009) | (0.003) | (0.003) |
| Leverage | -0.022 | -0.041 | -0.185 | -0.401*** | -0.379*** |
| | (0.134) | (0.135) | (0.190) | (0.099) | (0.105) |
| MtB Ratio | 0.005 | -0.005 | 0.007 | -0.051** | -0.064** |
| | (0.066) | (0.059) | (0.054) | (0.023) | (0.029) |
| Cash Fl. | -0.525* | 0.088 | 0.427* | 0.051 | -0.063 |
| | (0.267) | (0.134) | (0.249) | (0.144) | (0.128) |
| RD Expenses | 0.115 | -0.427 | 1.615** | 0.245 | 0.233 |
| | (1.101) | (0.702) | (0.690) | (0.335) | (0.389) |
| RD Missing | 0.111 | 0.044 | 0.457* | -0.022 | 0.040 |
| | (0.262) | (0.151) | (0.243) | (0.098) | (0.069) |
| Ex. Board Size | -0.082 | -0.074 | -0.106* | -0.026 | -0.037 |
| | (0.091) | (0.081) | (0.064) | (0.025) | (0.026) |
| EBIT | 0.086 | -0.099 | 0.072 | -0.085 | 0.019 |
| | (0.333) | (0.192) | (0.245) | (0.196) | (0.179) |
| Stock Return | 0.038 | 0.060 | 0.262** | 0.018 | 0.090* |
| | (0.106) | (0.061) | (0.110) | (0.034) | (0.049) |
| *Fixed-effects* | | | | | |
| Firm (445) | Yes | Yes | Yes | Yes | Yes |
| Year (9) | Yes | Yes | Yes | Yes | Yes |
| Age (47) | Yes | Yes | Yes | Yes | Yes |
| Tenure (35) | Yes | Yes | Yes | Yes | Yes |
| Gender (2) | Yes | Yes | Yes | Yes | Yes |
| *Fit statistics* | | | | | |
| Observations | 2,660 | 2,660 | 2,660 | 2,660 | 2,660 |
| Adjusted $R^2$ | 0.312 | 0.733 | 0.605 | 0.590 | 0.756 |

*Notes:* The dependent variables are logged compensation components, indicated in the column headers. We add the value one before taking logs. Market capitalization is calculated as the product of the number of outstanding shares and the end-of-year market price. Standard deviation of stock returns is calculated based on monthly returns over the previous five years and is divided by 100. Leverage is long-term debt divided by total assets, Market-to-book ratio is the ratio of market capitalization and total assets. Cash flow is profit after tax plus depreciation and amortization, normalized by lagged total assets. R&D expenses are normalized by total assets. R&D Missing is a dummy indicating whether R&D expenses are missing. Executive board size is the number of executive directors on the board. EBIT is normalized by total assets. *Signif. Codes:* ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered on the firm-level.

Table 36: Compensation and Overconfidence - Net-Buyer

| Dep. Var.<br>Model: | ln(Options+1)<br>(1) | ln(Stocks+1)<br>(2) | ln(Bonus+1)<br>(3) | ln(Salary+1)<br>(4) | ln(Total+1)<br>(5) |
|---|---|---|---|---|---|
| *Variables* | | | | | |
| Net Buyer | 0.011 | -0.174 | 0.014 | 0.068 | 0.148 |
| | (0.191) | (0.289) | (0.270) | (0.111) | (0.107) |
| ln(Mkt Cap) | 0.339*** | 0.173* | 0.670*** | 0.170*** | 0.321*** |
| | (0.091) | (0.092) | (0.109) | (0.047) | (0.045) |
| SD Return | 0.017** | 0.012 | 0.013 | 0.009** | 0.014*** |
| | (0.008) | (0.009) | (0.009) | (0.004) | (0.004) |
| Leverage | -0.105 | 0.034 | -0.115 | -0.191 | -0.151 |
| | (0.068) | (0.080) | (0.098) | (0.162) | (0.191) |
| MtB Ratio | 0.009 | 0.014 | -0.008 | -0.020 | -0.013 |
| | (0.032) | (0.037) | (0.035) | (0.016) | (0.018) |
| Cash Fl. | 0.029 | 0.015 | 0.054 | 0.060** | 0.054*** |
| | (0.030) | (0.029) | (0.033) | (0.027) | (0.020) |
| RD Expenses | -1.660 | -0.651 | 1.017 | 0.008 | -0.130 |
| | (1.127) | (1.170) | (0.840) | (0.387) | (0.462) |
| RD Missing | 0.060 | -0.084 | 0.465** | 0.107 | 0.137 |
| | (0.203) | (0.207) | (0.200) | (0.087) | (0.085) |
| Ex. Board Size | -0.060 | -0.125* | -0.087 | -0.008 | -0.026 |
| | (0.063) | (0.068) | (0.059) | (0.023) | (0.024) |
| EBIT | -0.046 | -0.010 | 0.281 | -0.395 | -0.162 |
| | (0.278) | (0.198) | (0.290) | (0.251) | (0.189) |
| Stock Return | 0.048 | 0.078 | 0.328*** | 0.035 | 0.086** |
| | (0.082) | (0.072) | (0.112) | (0.035) | (0.039) |
| *Fixed-effects* | | | | | |
| Firm (525) | Yes | Yes | Yes | Yes | Yes |
| Year (9) | Yes | Yes | Yes | Yes | Yes |
| Age (51) | Yes | Yes | Yes | Yes | Yes |
| Tenure (39) | Yes | Yes | Yes | Yes | Yes |
| Gender (2) | Yes | Yes | Yes | Yes | Yes |
| *Fit statistics* | | | | | |
| Observations | 3,338 | 3,338 | 3,338 | 3,338 | 3,338 |
| Adjusted $R^2$ | 0.400 | 0.737 | 0.618 | 0.642 | 0.757 |

*Notes:* The dependent variables are logged compensation components, indicated in the column headers. We add the value one before taking logs. Market capitalization is calculated as the product of the number of outstanding shares and the end-of-year market price. Standard deviation of stock returns is calculated based on monthly returns over the previous five years and is divided by 100. Leverage is long-term debt divided by total assets, Market-to-book ratio is the ratio of market capitalization and total assets. Cash flow is profit after tax plus depreciation and amortization, normalized by lagged total assets. R&D expenses are normalized by total assets. R&D Missing is a dummy indicating whether R&D expenses are missing. Executive board size is the number of executive directors on the board. EBIT is normalized by total assets. *Signif. Codes:* ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered on the firm-level.

Table 37: Compensation and Overconfidence - Longholder

| Dep. Var.<br>Model: | ln(Options+1)<br>(1) | ln(Stocks+1)<br>(2) | ln(Bonus+1)<br>(3) | ln(Salary+1)<br>(4) | ln(Total+1)<br>(5) |
|---|---|---|---|---|---|
| *Variables* | | | | | |
| Longholder | 0.269 | -0.733** | 0.059 | -0.118 | -0.049 |
| | (0.345) | (0.315) | (0.295) | (0.088) | (0.081) |
| ln(Mkt Cap) | 0.462*** | 0.174 | 0.725*** | 0.105** | 0.309*** |
| | (0.135) | (0.120) | (0.123) | (0.052) | (0.054) |
| SD Return | 0.025** | 0.000 | -0.006 | -0.009*** | -0.006** |
| | (0.011) | (0.011) | (0.009) | (0.003) | (0.003) |
| Leverage | -0.024 | -0.007 | -0.204 | -0.412*** | -0.380*** |
| | (0.134) | (0.126) | (0.192) | (0.097) | (0.104) |
| MtB Ratio | -0.004 | -0.031 | 0.006 | -0.051** | -0.067** |
| | (0.065) | (0.059) | (0.054) | (0.023) | (0.029) |
| Cash Fl. | -0.452 | 0.034 | 0.431* | 0.045 | -0.069 |
| | (0.276) | (0.132) | (0.259) | (0.150) | (0.130) |
| RD Expenses | 0.023 | -0.736 | 1.447** | 0.251 | 0.219 |
| | (1.102) | (0.721) | (0.691) | (0.326) | (0.391) |
| RD Missing | 0.011 | -0.051 | 0.426* | 0.005 | 0.042 |
| | (0.278) | (0.140) | (0.242) | (0.104) | (0.072) |
| Ex. Board Size | -0.113 | -0.112 | -0.109 | -0.029 | -0.045 |
| | (0.094) | (0.084) | (0.068) | (0.026) | (0.028) |
| EBIT | 0.028 | -0.165 | 0.023 | -0.081 | 0.012 |
| | (0.335) | (0.187) | (0.249) | (0.192) | (0.179) |
| Stock Return | 0.053 | 0.069 | 0.300*** | 0.023 | 0.098** |
| | (0.108) | (0.058) | (0.114) | (0.033) | (0.049) |
| *Fixed-effects* | | | | | |
| Firm (418) | Yes | Yes | Yes | Yes | Yes |
| Year (9) | Yes | Yes | Yes | Yes | Yes |
| Age (47) | Yes | Yes | Yes | Yes | Yes |
| Tenure (35) | Yes | Yes | Yes | Yes | Yes |
| Gender (2) | Yes | Yes | Yes | Yes | Yes |
| *Fit statistics* | | | | | |
| Observations | 2,451 | 2,451 | 2,451 | 2,451 | 2,451 |
| Adjusted $R^2$ | 0.282 | 0.759 | 0.600 | 0.632 | 0.724 |

*Notes:* The dependent variables are logged compensation components, indicated in the column headers. We add the value one before taking logs. Market capitalization is calculated as the product of the number of outstanding shares and the end-of-year market price. Standard deviation of stock returns is calculated based on monthly returns over the previous five years and is divided by 100. Leverage is long-term debt divided by total assets, Market-to-book ratio is the ratio of market capitalization and total assets. Cash flow is profit after tax plus depreciation and amortization, normalized by lagged total assets. R&D expenses are normalized by total assets. R&D Missing is a dummy indicating whether R&D expenses are missing. Executive board size is the number of executive directors on the board. EBIT is normalized by total assets. *Signif. Codes:* ***: 0.01, **: 0.05, *: 0.1. Standard errors are clustered on the firm-level.

# References

Blundell, R., Griffith, R., and van Reenen, J. (1999). Market Share, Market Value and Innovation in a Panel of British Manufacturing Firms. *The Review of Economic Studies*, 66(3):529–554.

Hall, B. (1990). The Manufacturing Sector Master File: 1959-1987. NBER Working Paper 3366, National Bureau of Economic Research, Cambridge, MA.

# C Appendix: Paying for Optimism

## C.1 Proof of Lemma 3.1

The optimal expectations agent chooses the skill level to maximize material utility. The maximization problem and first order condition are in Equation 24. The second derivative with respect to skill is negative, $\frac{\partial^2 u}{\partial s^2} = -1 < 0$, so this is indeed a maximum.

$$\begin{aligned} \max_s \quad & w - \frac{1}{2}s^2 + p(bs + n) \\ \text{FOC:} \quad & \frac{\partial u}{\partial s} = -s + pb \overset{!}{=} 0 \\ & s^* = pb \end{aligned} \tag{24}$$

This result is analogous to the rational benchmark, where agent and principle share a common prior and the agent attains skill $s = qb$. Taking into account her optimal skill decision, $s = pb$, the optimal expectations agent chooses her belief $p$ to maximize felicity, given in Equation 14. The maximization problem and first order condition are in Equation 25.

$$\begin{aligned} \max_p \quad & w - \frac{1}{2}p^2b^2 + \alpha p(pb^2 + n) + (1-\alpha)q(pb^2 + n) \\ \text{FOC:} \quad & \frac{\partial f}{\partial p} = -pb^2 + \alpha 2pb^2 + \alpha n + (1-\alpha)qb^2 \overset{!}{=} 0 \\ \Leftrightarrow \quad & p^* = \frac{\alpha}{1-2\alpha}\frac{n}{b^2} + \frac{1-\alpha}{1-2\alpha} \cdot q \end{aligned} \tag{25}$$

The second derivative of the agent's felicity is $\frac{\partial f}{\partial p} = -b^2 + \alpha 2b^2$, which is negative if $\alpha < \frac{1}{2}$ and positive for $\alpha > 0.5$. Thus, $p^*$ is a maximum for $\alpha \in [0, 0.5]$. Note that $p^*$ has a discontinuity at $\alpha = 0.5$, where the left limit is $+\infty$ and the right limit is $-\infty$. I require the belief to be a well defined probability, and thus $p \in [0,1]$ and $p^* \notin [0,1]$ for $\alpha > \frac{(1-q)b^2}{(2-q)b^2+n}$. From the second derivative we can immediately conclude that $p = 1$ is optimal if $\frac{(1-q)b^2}{(2-q)b^2+n} < \alpha < 0.5$. Note that the first order solution is a minimum for $\alpha > 0.5$ and so it remains to verify that $f(p = 1) > f(p = 0)$ for $\alpha > 0.5$ to show the desired result. In fact, $f(p = 1) - f(p = 0) = b^2(\alpha + (1-\alpha)q - \frac{1}{2} + \alpha n > 0$ as $\alpha > 0.5$, $b > 0$, and $n \geq 0$.

## C.2 Proof of Proposition 3.1

Consider the principal's profit maximization problem in Equation 18, where the incentive constraint is complicated by the endogenous beliefs in the optimal expectations constraint (OE). The bonus will be strictly positive at the optimum, $b > 0$, as it is the only source of profit. Since the agent's belief $p$ is increasing in stock options $n$ and the bonus is positive $b > 0$, we can conclude that the belief can't be zero either as $p \geq \frac{1-\alpha}{1-2\alpha}q > 0$. Note that the corner solution

$p = 1$ will be obtained for $\alpha \geq \frac{(1-q)b^2}{(2-q)b^2+n}$ as discussed in Appendix C.1.

If the outside option is small enough the participation constraint is slack and can thus be ignored. Note that if the PC is non-binding the wage $w$ only decreases profits and plays no other role. Thus, $w = 0$ in the optimum. Plugging in the remaining constraints into the profit function and rearranging yields a simplified principal's problem:

$$\begin{aligned} \max_{w,b,n} \quad & pb - \frac{1-\alpha}{\alpha} qb^2(p-q) \\ & n \geq 0, \quad p \leq 1, \end{aligned} \tag{26}$$

The corresponding Lagrangian is $\mathcal{L} = pb - \frac{1-\alpha}{\alpha} qb^2(p-q) + \mu(p-1)$ and its first derivatives are

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial b} &= p - \frac{1-\alpha}{\alpha} 2qb(p-q) \\ \frac{\partial \mathcal{L}}{\partial p} &= b - \frac{1-\alpha}{\alpha} qb^2 + \mu. \end{aligned} \tag{27}$$

Keeping track of possible corner solutions, $n = 0$ or $p = 1$, the Kuhn Tucker conditions are

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial b} &\overset{!}{=} 0 \\ \frac{\partial \mathcal{L}}{\partial p} &\overset{!}{=} 0 \\ \mu(p-1) = 0 \implies \mu &= 0 \text{ or } p = 1 \\ \mu, n &\geq 0 \\ 0 < p &\leq 1. \end{aligned} \tag{28}$$

This yields four solutions candidates,

1. $b > 0, n = 0, p = \frac{1-\alpha}{1-2\alpha} q < 1$

2. $b > 0, n > 0, \frac{1-\alpha}{1-2\alpha} q < p < 1$

3. $b > 0, n > 0, p = 1$

4. $b > 0, n = 0, p = 1,$

where cases three and four are extreme in the sense that the agent's optimism is maximized, $p = 1$. Indeed, case four is somewhat vacuous, since it is the only case where the agent puts so much weight on anticipatory utility that his optimism is maximized without any reinforcement through stock options, $\alpha > \frac{1-q}{2-q}$. For cases one, two, and three the weight is low enough that the principal could implement an interior belief, if she wanted to. It remains to solve these cases and compare the resulting profits to show which case is optimal.

**Case 1.** In the first case, stock options play no role, $n = 0$ and so we immediately conclude $p = \frac{1-\alpha}{1-2\alpha} q < 1$. Substituting into $\frac{\partial \mathcal{L}}{\partial b} \overset{!}{=} 0$ yields $b = \frac{1}{2q}$, which is the same bonus as in the rational benchmark. It follows that skill $s = pb = \frac{1}{2}\frac{1-\alpha}{1-2\alpha}$, which is bigger than in the rational benchmark, as $p > q$. The resulting profits are $\pi = \frac{1}{4}\frac{1-\alpha}{1-2\alpha}$.

**Case 2.** In the second case, stock options are positive so that an interior belief obtains. Thus, $\mu = 0$ and I can solve for the bonus from $\frac{\partial \mathcal{L}}{\partial p} \overset{!}{=} 0 \Leftrightarrow b = \frac{\alpha}{1-\alpha}\frac{1}{q}$. Substituting into $\frac{\partial \mathcal{L}}{\partial b} \overset{!}{=} 0$ yields $p = 2q$. Recall that stock options can be retrieved from the optimal expectations constraint (OE), $p = p = \frac{\alpha}{1-2\alpha}\frac{n}{b^2} + \frac{1-\alpha}{1-2\alpha}q \Leftrightarrow n = \frac{b^2}{\alpha}[(1-2\alpha)p - (1-\alpha)q]$. This yields $n = \frac{\alpha}{(1-\alpha)^2}\frac{1}{q}(1-3\alpha)$, which is positive for $\alpha < \frac{1}{3}$. The corresponding profit is $\pi = \frac{\alpha}{1-\alpha}$.

**Case 3.** In the third case, we can calculate the bonus from $\frac{\partial \mathcal{L}}{\partial b} \overset{!}{=} 0$, which yields $b = \frac{\alpha}{1-\alpha}\frac{1}{2q}\frac{1}{1-q}$. Given the bonus we can solve for the stock options that implement $p = 1$. This gives $n = \frac{b^2}{\alpha}[(1-2\alpha)p - (1-\alpha)q] = \frac{\alpha}{(1-\alpha)^2}\frac{1}{4q^2(1-q)^2}[1 - 2\alpha - (1-\alpha)q]$, which is positive if and only if $1 - 2\alpha - (1-\alpha)q > 0 \Leftrightarrow \alpha < \frac{1-q}{2-q}$. The corresponding profits are $\pi = \frac{1}{4}\frac{\alpha}{1-\alpha}\frac{1}{q}\frac{1}{1-q}$.

**Case 4.** In the fourth case, the agent puts so much weight on anticipatory utility, $\alpha > \frac{1-q}{2-q}$, that he is fully optimistic, $p = 1$, even though he receives no stock options at all, $n = 0$. The principle chooses the same bonus as in Case 3, $b = \frac{\alpha}{1-\alpha}\frac{1}{2q}\frac{1}{1-q}$, which yields profits of $\pi = \frac{\alpha}{(1-\alpha)^2}\frac{2(1-\alpha)(1-q)-\alpha}{4q(1-q)^2}$.

It turns out that Case 1 (weakly) dominates Case 2, $\pi_1 \geq \pi_2 \Leftrightarrow (1-3\alpha)^2 \geq 0$, which holds for the relevant range $\alpha < \frac{1-q}{2-q}$. It remains to compare profits of Cases 1 and 3. It turns out that case three dominates case one if and only if $\frac{\alpha - 2\alpha^2}{(1-\alpha)^2} \geq q(1-q)$, which is a lower bound on $\alpha$. Recall that both cases apply only for $\alpha < \frac{1-q}{2-q}$. This defines a proper interval in which Case 3 dominates Case 1, such that stock options play a role in the contract if and only if $\alpha \in [\frac{q}{1+q}, \frac{1-q}{2-q}]$ if $q \leq 0.5$. For $q > 0.5$, only a single point, $\alpha = \frac{1-q}{2-q}$, satisfies $\frac{\alpha-2\alpha^2}{(1-\alpha)^2} \geq q(1-q)$ and thus the interval is degenerate.

Accordingly, the optimal contract is

$$
\begin{aligned}
w &= 0. \\
b &= \begin{cases} \frac{1}{2q} & \text{if } \alpha < \frac{q}{1+q} \\ \frac{\alpha}{(1-\alpha)}\frac{1}{2q}\frac{1}{(1-q)} & \text{if } \alpha > \frac{q}{1+q} \text{ for } q \leq 0.5 \text{ and } \alpha \leq \frac{1-q}{2-q} \\ \frac{\alpha}{(1-\alpha)}\frac{1}{2q}\frac{1}{(1-q)} & \text{if } \alpha > \frac{1-q}{2-q} \end{cases} \\
n &= \begin{cases} 0 & \text{if } \alpha < \alpha < \frac{q}{1-q} \\ \frac{\alpha}{(1-\alpha)^2}\frac{1-2\alpha-(1-\alpha)q}{4q^2(1-q)^2} & \text{if } \alpha > \frac{q}{1+q} \text{ for } q \leq 0.5 \text{ and } \alpha \leq \frac{1-q}{2-q} \\ 0 & \text{if } \alpha > \frac{1-q}{2-q} \end{cases}
\end{aligned}
\tag{29}
$$

## C.3 Proof of Proposition 3.2

Consider the principal's problem in Equation 18 when the participation constraint is binding. To show that the salary must be zero, suppose that it is positive, $w > 0$, at the optimal contract

$(w, b, n)$. Holding total payments constant, the principal can decrease the salary, $w' = w - q\epsilon$ and increase stock options $n' = n + \epsilon$. Fixing the old belief $p$, the agent would accept this new contract, $(w', b, n')$, as the change in payment creates slack in the participation constraint,

$$\begin{aligned} & w' - \frac{1}{2}s^2 + \alpha p(bs + n') + (1-\alpha)q(bs + n') \\ & = w - q\epsilon + q(n+\epsilon) + \alpha(p-q)(n+\epsilon) - \frac{1}{2}s^2 + pqb^2 + \alpha(p-q)pb^2 \\ & = w + qn + \underbrace{\alpha(p-q)(n+\epsilon)}_{>0} - \frac{1}{2}s^2 + pqb^2 + \alpha(p-q)pb^2 > \hat{u}. \end{aligned} \tag{30}$$

At the new contract the agent will re-optimize his beliefs to $p' > p$, which gives him (weakly) higher felicity, preserving the slack in the participation constraint. The principal benefits as the higher belief $p' > q$ yields a higher skill for the same bonus, thus increasing profits. It follows that the wage, $w$, must be zero at the optimum.

To show that stock options must play a role in the optimal contract (at least for some $\alpha$ region), assume that the optimal contract is such that $(w, b, n) = (0, b, 0)$. Then, the principal can increase stock options, $n$, while holding the skill, $s$, fixed through an appropriate decrease in bonus, $b$. Per the implicit function theorem I can calculate the appropriate decrease in the bonus, $\frac{db}{dn}$, when holding the skill fixed, $G(b, n) = s - \frac{\alpha}{1-2\alpha}\frac{n}{b} - \frac{1-\alpha}{1-2\alpha}qb$. Then,

$$\begin{aligned} \frac{db}{dn} &= \frac{\frac{-\partial G}{\partial n}}{\frac{\partial G}{\partial b}} \\ &= \frac{\frac{\alpha}{1-2\alpha}\frac{1}{b}}{\frac{\alpha}{1-2\alpha}\frac{n}{b^2} - \frac{1-\alpha}{1-2\alpha}q} \\ &= \frac{\alpha b}{\alpha n - (1-\alpha)qb^2}. \end{aligned} \tag{31}$$

At $n = 0$ this reduces to $\frac{\alpha}{1-\alpha}\frac{1}{q}$. Note that this swap of bonus to stock options, which keeps skill constant, creates slackness in the participation constraint. The participation constraint can be written as $s^2(\alpha - \frac{1}{2}) + (1-\alpha)qbs + \alpha pn + (1-\alpha)qn = \hat{u}$. The change in the participation constraint is,

$$(1-\alpha)q\frac{db}{dn}s + \alpha p + (1-\alpha)q = (1-\alpha)q > 0.$$

But profits change with this swap of bonus to options,

$$\begin{aligned} -q\frac{db}{dn}s - q &\overset{!}{>} 0 \\ \Leftrightarrow p &> \frac{1-\alpha}{\alpha}q \\ \Leftrightarrow \frac{\alpha}{1-2\alpha}q &> \frac{1-\alpha}{\alpha}q \\ \Leftrightarrow \alpha &> \frac{1}{3}. \end{aligned} \tag{32}$$

Thus, if $\alpha > \frac{1}{3}$, the principle will use stock options to induce (at least some) optimism in the agent even with a binding participation constraint.

# D Appendix: Statements and Curriculum Vitae

## Affidavit

Ich, Herr Simon Christian Schulten, versichere an Eides statt, dass die vorliegende Dissertation von mir selbstständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist.

English version:
I, Mr. Simon Christian Schulten, solemnly declare that the present dissertation has been independently and without impermissible external assistance created by myself, in accordance with the "Principles for Safeguarding Good Academic Practice at Heinrich Heine University Düsseldorf."

---

**Simon Schulten**
University of Düsseldorf

# Coauthorship Declaration

**Managing Bidder Learning in Retail Auctions**

We, the coauthors Paul Schäfer and Simon Schulten, have equally and substantially contributed to every facet of this study. Our collaboration has proven fruitful, with both of us actively participating in conceptualization, research, data analysis, and writing of this paper. This joint effort highlights the balanced nature of our coauthorship.

Paul Schäfer

**Paul Schäfer**
University of Bonn

**Simon Schulten**
University of Düsseldorf

# Coauthorship Declaration

**CEO Overconfidence in Europe**

We, the coauthors David Zeimentz, Simon Schulten and Dennis Gottschlich, jointly contributed to this work. While Zeimentz and Schulten were particularly involved in the project's conceptualization, data acquisition, coding, and writing efforts, Gottschlich significantly added to the paper by providing manual data work, a literature review, and formatting. He further provided valuable support in terms of conceptualization and coding tasks.

**David Zeimentz**
University of Düsseldorf

**Simon Schulten**
University of Düsseldorf

**Dennis Gottschlich**
University of Düsseldorf

# Simon Schulten

*www.Simon-Schulten.de*
*Simon.Schulten@hhu.de*

## Research Interests

**Applied Microeconomics, Behavioral Economics, Industrial Organisation**.

## References

**Prof. Dr. Paul Heidhues**, *DICE, University of Düsseldorf*.
heidhues@dice.hhu.de

**Prof. Dr. Joel Stiebale**, *DICE, University of Düsseldorf*.
stiebale@dice.hhu.de

**Prof. Dr. Florian Heiß**, *University of Düsseldorf*.
florian.heiss@hhu.de

## Education

2018 – 2023 **PhD in Economics**, *DICE, University of Düsseldorf*.

2015 – 2018 **Master of Science in Economics**, *University of Mannheim*.

2011 – 2015 **Bachelor of Science in Economics**, *University of Mannheim*.

Fall 2013 **Exchange semester**, *Xiamen University*.

## Research

Job Market Paper **Managing Bidder Learning in Retail Auctions**.
with Paul Schäfer

**Managerial Overconfidence in Europe**.
with David Zeimentz, Dennis Gottschlich

**Paying for Optimism: A Model of Stock Options for Rank and File Employees**.
Single authored

## Seminars, Conferences and Workshops

2023 Behavioral IO & Marketing Symposium (University of Michigan).

2022 EARIE (University of Vienna), CORE Brown Bag Seminar (UCLouvain).

2021 Causal Data Science Meeting (CBS and Maastricht Universtiy).

2019 Paris Summer School (PSE), CISS (Montenegro).

2017 Datafest Germany (University of Mannheim, team member).

2016 Datafest Germany (LMU Munich, participant).

2015 Datafest Germany (University of Mannheim, participant).

## Employment

2018 – 2023 **Doctoral Researcher**, *DICE, University of Düsseldorf*.

Fall 2017 **Research Assistant**, *University of Mannheim*.

Spring 2017 **Intern Market Design**, *Centre for European Economic Research (ZEW)*.

2016 **Teaching Assistant**, *University of Mannheim*.

## Teaching

Fall '19 – '22 & Spring '23 **Methods in Institutional Economics**, *DICE, University of Düsseldorf*.
I give a lecture with integrated exercise sessions.

2019 – 2022 **Thesis and Term Paper Supervision**, *DICE, University of Düsseldorf*.
I supervised multiple term papers and co-supervised two Bachelors theses.

Fall 2016 **Econ 101**, *University of Mannheim*.
I gave two exercise sessions.

Spring 2016 **Micro A**, *University of Mannheim*.
I gave two exercise sessions.

## Personal and technical skills

Languages **German** (native)**, English** (fluent)

Programming **R** (advanced)**, Python** (intermediate)**, Otree** (basic)

Coursera **Specialization: Algorithms** — Courses 1 and 2.

**Simon Schulten**
August 2023