Three Essays in Theoretical Economics

Inaugural-Dissertation

zur Erlangung des Grades

Doctor oeconomiae publicae (Dr. oec. publ.)

an der Ludwig-Maximilians-Universität München

Volkswirtschaftliche Fakultät

2021

vorgelegt von

Lion Heinrich Henrich

Referent: Prof. Dr. Klaus M. Schmidt

Korreferent: Prof. Dr. Gerhard Illing

Datum der mündlichen Prüfung: 05. Juli 2021

Promotionsabschlussberatung: 14. Juli 2021

Acknowledgements

First and foremost, I am deeply indebted to my mentor and first supervisor Klaus Schmidt. During my time at his chair, he was always willing to offer helpful advice and valuable as well as thoughtful comments which helped me tremendously to complete this work. Furthermore, his continuous efforts to make our chair such a welcoming place to work at and such a sublime place to conduct research are highly appreciated. In particular, I am grateful for the annual organization of highly educating and pleasurable research seminars in Obergurgl and Ohlstadt from which I benefited a lot.

Second, I owe a great debt of gratitude to my second supervisor Gerhard Illing who volunteered on rather short notice to support me in my efforts to write a paper based on macroeconomic theory, even though he knew that it was not my main area of expertise. His accessibility in combination with his competent advice was invaluable for the project's completion.

Third, I thank Florian Englmaier for his willingness to join my thesis committee despite his already overflowing schedule. Further, I thank him very much for his steadfast willingness to support me in my research efforts and his continuous encouragement over the years.

Additionally, I want to express my gratitude to Matthias Lang whose advice and support concerning my single-author project helped me tremendously to improve it. Besides, his trust to appoint me as a teaching assistant for the Virtual Journal Club he organized last term earned me one of the best teaching experiences of my life. Thank you very much.

I am also extremely grateful to Friedrich Breyer, who was my professor and mentor during my time at the University of Konstanz. He was the first person to employ me as a research assistant and who encouraged me to pursue a Ph.D. in economics. His way of analytical thinking, his discipline, and his engagement in the public economic debate have been constant sources of inspiration and impetus to me, throughout my entire academic career and until this day.

Many thanks go to my coauthors, Annemarie Gronau (chapter 2) and Markus Epp (chapter 3). Aside from our working relationship, I am proud to call both of them good friends. Annemarie helped me to find and complete our, i.e. my first, project. She always treated me as an equal as well as an appreciated partner, and further, always offered an open mind in combination with helpful words when I had personal issues.

My history with Markus Epp goes back to the first semester term in our Bachelor's program in Konstanz. Since then, his friendship has been invaluable to me. Without his tenacity, his kindness, his ambition, his intelligence, and, above all, his care as a friend, I never would be where I am right now. I cannot thank you enough.

Special thanks also go to Sven Resnjanskij who has been a reliable and competent advisor as well as a personal friend throughout the years. Further, I am indebted to Karen Bernhardt-Walther who during the last six months was a constant source of valuable writing tips and encouragement.

My research projects strongly benefited from comments made by participants of workshops and seminars. In particular, I am grateful to Peter Schwardmann, Yves Le Yaouanq, Johannes Maier, Taisuke Imai, Marco Schwarz, Lukas Buchheim, Peter Zorn, and Franziska Hünnekes for their insightful comments on my research.

Furthermore, I am deeply indebted to my colleagues at the Seminar for Economic Theory, Else Christensen, Dominik Fischer, Benjamin Häusinger, Hung-Ni Chen, Hoa Ho, Christoph Schwaiger, and Friederike Reichel. Thank you for all of your support and generally, an unforgettable time I was allowed to spend with you. I genuinely enjoyed all events we shared together.

Aside from our chair, I had the great fortune of meeting sublime young researchers at LMU, across all Ph.D. cohorts. In particular, many thanks go to Dominik Grothe, Michael Kaiser, Alexander Schwemmer, Timm Opitz, Leonie Bielefeld, and Clarissa Mang. I especially want to thank Hirofumi Okoshi for being a wonderful person and colleague, as well as a personal friend. Thank you for believing in me even at times I just wanted to give up.

Many thanks also go to Sabine Wilhelm-Kauf for her reliable, competent, and overall invaluable help with all kinds of administrative and organizational issues.

Besides, I am happy to have been part of the GRK 1928 graduate program and want to thank the program coordinator Carsten Eckel for all of his efforts. Special thanks go to GRK 1928

and the Deutsche Forschungsgemeinschaft (DFG) for their generous financial support, particularly for the GRK Munich Young Economists' Meeting 2019 I co-organized.
Last, I can only express my most profound and deepest love and gratitude towards my family, most notably, my parents Wolfram Henrich and Petra Habrock-Henrich as well as my brother Tristan Henrich. Your unrelenting support, your love, your care, and the feeling of safety and security you give me have been the cornerstones of my life. Words cannot express how grateful I am for all that. My greatest appreciation and love goes to my girlfriend, Silvia. You stayed with me for all those years, even though we used to live hundreds of kilometers apart. Your support and love throughout these years permitted me to keep going. Thank you. I am looking forward to our future life together.

Lion Henrich

Munich, March 2021

To Silvia.

Always.

Contents

List of Figures

List of Tables

Preface

> “In my view, economic theory is "just" an arena for the *investigation of concepts* we use in thinking about real-life economic situations. [...] Through the investigation of these concepts, we try to better understand reality, and the models provide a language that enables us to think about economic interactions in a systematic way. ”
>
> A. Rubinstein (2012), *Introduction*

This dissertation is special in the sense that it contains three distinct topics from seemingly entirely different areas of economics: Decision theory, Agency theory, and Macroeconomics. Thus, at first glance, it consists of three projects concerning scenarios without any common features. Especially in times of increased specialization with regard to the own field of expertise, such a mixture of projects containing not only topics from Microeconomics in form of individual decision making and Game theory but also aggregate behavior of the overall economy might appear surprising.

From a high-level overview, however, there exist at least two connected unifying elements. First, all projects represent applications of economic theory to real-life questions and puzzles in economics. Using economic theories based on specific assumptions regarding human and/or other economic entities' behavior permits economists to explore complex problems within a simplified representation by focusing on the essentials. It provides an elegant tool and a commonly comprehensible language to analyze real-life economic issues. In essence, the use of economic models to study economic interactions presents a mode of thought that is not restricted to one specific area of economics only but that is broadly applicable. To say it with the words of John Maynard Keynes (1938): "Economics is a science of thinking in terms of

models joined to the art of choosing models which are relevant to the contemporary world."
Second, strongly related to the first point, there is the author's general and broad interest in basically all economic topics. In fact, the first point enabled and fueled the second. Early on during my Bachelor's program, I recognized the beauty and elegance of economic models to express and exemplify real-life phenomena. It was an astonishing experience to discover a new technique that enabled me to analytically think about individual decision making, strategic interaction among economic agents, as well as the behavior of an economy as a whole. The fascination of learning a method applicable to setups in Microeconomics as well as Macroeconomics has remained to this day. For me, economic models have always been more than a summary of numbers or variables: they are applied logic and exhibit meaning that can be expressed in form of mathematical equations and graphs. When applied sophisticatedly, they represent a valuable analytical tool able to shed light on various puzzling economic phenomena in different areas and to provide answers to problems we face in reality. In this dissertation, I aim to accomplish exactly that.
All chapters included in this work are based on self-contained papers and can be read independently of each other. They are complemented by separated appendices and a joint bibliography I present at the end of this dissertation. In what follows, I provide an overview of the main points developed in each chapter, respectively.

In Chapter 1, I analyze the consequences arising if individuals overstate the representativeness of finite random sample data, i.e., the consequences of individuals being ignorant about the existence of random errors associated with randomly drawn samples. Decades of research in psychology as well as in economics have shown that human beings tend to be "semi-sophisticated" statisticians at best and tend to be prone to apparently "irrational" behavior. Figuring out in what way exactly individuals fail to behave as sophisticated statisticians might thus lead to explanations for observed deviations from the behavior postulated by the theory of strict rationality. The focus of my approach lies on probabilistic assessments of an uncertain future event, preceding actual decision making. I show how beliefs based on random error neglect are systematically flawed and stand in contrast to those made by rational agents, or as I call them in this chapter: sophisticated statisticians.
In contrast to the vast majority of literature, I do not assume perception biases or the simple

use of heuristics as an explanation for rather particular behavioral fallacies. In my approach, all individuals are fully capable of using the best-possible unbiased estimation strategies in each context. Further, they do not specify the underlying model incorrectly. For example, they do not assume a linear regression model while the true connection of two variables is quadratic. However, they do not behave like fully rational agents because they fail to account for the existence of random errors when dealing with limited sample data. In other words, they believe their sample estimates – mean and variance in the context of a static random variable or intercept and slope parameter in the context of regression analysis – to be constants, not random variables.

The rather simple assumption of individuals naively extrapolating from sample properties to distribution properties has great explanatory power concerning statistical fallacies. It can account for various well-established empirical phenomena, such as framing effects, overprecision, overinflated skepticism about other people's opinions, misperception of differences in groups, and unjustified belief in the persistence of trends.

Further, I show how this theoretical set-up can be incorporated into existing economic models. One example is the bidding behavior in auctions. This theory can explain why overbidding is a more prominent feature in First-Price Sealed-Bid auctions than underbidding. Additionally, it illustrates why people who, according to the standard theory, bid too little, deviate stronger from the rational bidding behavior than those who bid too much.

In addition to its explanatory power and its broad applicability, this theory has the advantage of specifying the degree by which individuals neglecting random errors deviate from rational assessments: it lies entirely in the size of the samples these individuals consider. If the individual only considers enough realizations, his probabilistic assessment will be approximately close to the one of a sophisticated statistician. This might offer a way to easily improve peoples' predictions and the resulting actions: even though the actual neglect may not be cured, increasing the officially required number of outcomes an individual has to consider before forming an opinion can significantly improve the resulting behavior.

In Chapter 2, which is joint work with Annemarie Gronau (LMU), we show how a principal should optimally apply the concept of Autonomy Support as a non-monetary incentive to an agent in order to increase the likelihood of successful innovative activity. Autonomy Support refers to actions and management methods of, for example, a supervisor to encourage independent choices and the initiative of employees. Further, it is supposed to provide meaning and on-the-job training, and, most importantly, renounce to put any form of pressure on the agent to take certain actions (Stone et al., 2009).

Within firms, workers frequently exhibit hands-on experience with products and production processes. As a result, they may have implicit and unique knowledge to improve them. For example, they might be suited best to fully understand customers' wishes and finding ways to (creatively) meet them. Hence, encouraging their creativity can lead to meaningful improvements for the production and the design of a new product and consequentially, its ability to compete with other products. A vast amount of literature has shown, however, that companies can only make limited use of monetary incentives to encourage these innovations; such explicit incentives are only valid tools to incentivize target-oriented solutions to specific problems but not for unconstrained innovations needing proper creativity. Thus, relying solely on monetary incentives deprives a company of the full innovative potential of its workforce. In practice, the workers under consideration are not specifically employed for completing creative tasks. It is the environment at work that influences whether they feel free to think about new ideas and process improvements, and further, whether they feel safe enough to share these ideas with their superiors. In other words, their environment influences whether they dare to challenge the status quo by being creative.

In contrast to the majority of literature in organizational economics, our focus lies not on one specific management practice, but on the general behavior of a leader that fosters and invites novel ideas. We argue that leadership behavior that is successful in instilling such innovative activity is Autonomy Supportive.

To illustrate our point, we develop a theoretical model in which a representative agent enters a firm with an initial level of Autonomy Support, e.g., from prior education or previous employment. We assume, however, that the level of Autonomy Support does not remain constant over time: a single act of encouragement does not plausibly motivate innovation indefinitely. Thus, we assume that the level of Autonomy Support an agent possesses depreciates over time at a given rate. As a result, the investments in Autonomy Support must be made repeatedly to remain effective, capturing the necessity of ongoing leadership behavior in a working relationship. We account for this repeated setup by modeling the interaction of a principal with

an agent, not in a static, but a dynamic two-period interaction setup.
Our results generate several different investment patterns: First of all, we find that the principal invests just enough to achieve a level that provides optimal innovative effort by the agent after observing his initial endowment of Autonomy Support. This captures the intuition that while some individuals need extensive investments in the form of encouragement and skill training, for others, shorter (but nevertheless regular) meetings might suffice. In other words, leadership behavior must adapt to the specific requirements of an employee. Second, we find that for extremely low and/or high initial levels of Autonomy Support, the principal does not provide a positive amount. The reason is that either, the agent's requirements present investments that are too high and thus, too costly for the principal, or that an agent simply does not need additional support to facilitate his innovative activity. Third, we show that the specific investment dynamics are subject to the principal's discounting of future periods' payoffs and the rate by which the agent's Autonomy Support depreciates: the principal in our model begins the relationship with comparably high amounts of investments if she does not discount future periods too strongly and her support has a longer-lasting effect.

In Chapter 3, which is joint work with Markus Epp (University of Freiburg), we analyze whether central bankers should be rewarded for keeping interest rates up and whether monetary policy should "keep its powder dry" in anticipation of deteriorating economic conditions. These requests are in contrast to the optimal monetary policy framework in typical models of macroeconomic stabilization, where monetary policy cannot be stored and where interest rates are the instrument and not a goal of stabilization. Given the central bank has a dual mandate with a certain degree of freedom in putting weights on price stability and economic activity, our results are as follows.
First, letting a non-committing central banks' objective function also host explicit preferences for keeping interest rates above their natural level, the associated precautionary interest rate-setting (PIRS) creates a *dry powder paradox*: incentivizing higher nominal interest rates only leads to on average lower nominal interest rates in equilibrium. The reason is that such an incentive causes a deflationary bias, lowering average inflation and output gaps, which in turn influence rational expectations strongly enough to push nominal interest rates down.
This deflationary bias, however, helps to address the time inconsistency problem, which arises

when the central bank cannot credibly commit and the steady-state output level is distorted. A positive incentive to keep nominal interest rates above zero counters the inflationary bias in this case and leads to a combination of inflation and output gap that is welfare optimal for society. We show that the optimal PIRS-weight inducing this allocation has closed-form expressions.

Although the benefits of PIRS arising from the time inconsistency problem vanish under commitment, the effects imposed on average inflation by an explicit weight on keeping interest rates above their natural level may still provide a rationale for using it. Our finding illustrates the Fisherian nature of the long-run biases in the New Keynesian model: higher nominal interest rates drive inflation in the long-run. Hence, when the objective actually is to create some leeway away from the zero lower bound, such a reward might help to accomplish that goal. Nevertheless, not-considering a potentially binding effective lower bound, there is no welfare-based argument for a positive weight on keeping interest rates at higher levels. The reason is that the ability to commit enables the central bank to overcome the time inconsistency bias in the long-run without any additionally needed tools.

Further, we show how PIRS might prove to be a valuable strategy for a central bank to avoid the scenario of "fiscal dominance", in which interest expenses of a government cannot be covered by tax collection or rollovers quickly enough, such that a central bank would have to sacrifice its policy instrument to avoid government-insolvency.

Chapter 1

Random Error Neglect:

A theory of random samples and flawed conclusions

1.1 Introduction

Dealing with uncertainty is relevant for many economic settings. Predicting the next realization of a random process and quantifying the risk associated with this forecast is essential for investment decisions, purchases, procurement, gambling or even getting married. In the past, people had to rely on the auguries of higher authorities, like the Oracle of Delphi, to gain some insight into a "God-given" future. But ever since the first developments of probability theory in the 17th century by Blaise Pascal and Pierre de Fermat, more and more analytical approaches were devised.

In contrast to today's highly sophisticated algorithms using machine learning tools handling huge amounts of data, individuals usually have only limited information based on, for example, their own experiences or those made by acquaintances. Alternatively, they might face binding resource or cognitive constraints when collecting data. Hence, it is of great importance to understand how single individuals form beliefs about an uncertain future based on limited data. In general, statistical (or econometric) theory offers ways about how to do so in a sophisticated manner. In particular, it prescribes how to deal with the innate risk-taking associated with sampling, or to be more precise, the potential perils arising when forming assessments based on a limited number of data points. When using the average of a sample to guess about a future realization, one always has to account for how reliable the calculated sample average is. When forming beliefs based on sample data, the question is how representative the sample under consideration truly is.

If individuals are well aware of these issues and thus, form beliefs the same way a sophisticated statistician would, their assessments concerning an uncertain future will not be systematically flawed. However, decades of research in psychology as well as in economics have shown that human beings tend to be "semi-sophisticated" statisticians at best (see Kahneman (2011) for an overview). As a result, their probabilistic assessments, as well as the resulting behavior, tend to show significant flaws. Figuring out in what way exactly individuals fail to behave like sophisticated statisticians might thus lead to explanations for observed "misbehaving" (Thaler and Ganser, 2015).

In this chapter, I address the question about the consequences arising if individuals overstate the representativeness of a finite random sample, i.e. individuals neglecting the existence of random errors associated with randomly drawn samples. I argue that, depending on the size of the random sample, the resulting probability assessments of a *Random Error Neglecting Agent* (henceforth: RENA) are severely flawed. Random error neglect implies that an agent displays unjustifiably high trust in his estimates of the population parameters derived from his sample data. This can be exemplified best by the famous joke statisticians frequently use about a group of blind men who want to assess the shape of an elephant by all touching different parts of it: all men believe in a different shape and properties of the animal, dependent on whether a man touches just the trunk or just the abdominal parts because each of them fully trusts his own limited experience and ignores the possibility that there is more to be known.

This chapter shows how the assumption of individuals neglecting the random error of samples causes deviations from the rational, i.e. statistically sophisticated, assessment. Further, I delineate the extent of these deviations. I demonstrate that within this rather simple specification lies great explanatory power: the theory can serve as an endogenous explanation for well-established empirical findings like framing effects, overconfidence of individuals in form of overprecision, excessive distrust with respect to other people's opinions, the disbelief in the regression-to-the-mean phenomenon, unjustified beliefs in the persistence of trends as well as irrational bidding behavior in First-Price Sealed-Bid auctions. Furthermore, the theoretical set-up is easily applicable to various other canonical economic settings considering decision making under uncertainty.

The chapter proceeds as follows. In Section 1.2, I present a discussion of related literature, followed by the underlying assumptions and properties of the theory in Section 1.3. In Section 1.4, I show how naive reliance on formally unbiased estimators can lead to flawed parameter assessments. In Section 1.5, I depict how negligence concerning the random error affects an individual's certainty of future outcomes when predicting the next realization of one or more static random variables. Section 1.6 extends the analysis shown in Section 1.5 by also considering the case of predictions made via regression analysis. It exemplifies that the theory is also applicable to dynamic random processes. Section 1.7 shows how the theory of RENA can be applied in economic models dealing with bidding behavior in auctions. Section 1.8 concludes the chapter.

1.2 Related Literature

While overall contributing to the vast and steadily growing literature in psychology and (behavioral) economics about how humans form beliefs of uncertain events and how they might be influenced by cognitive biases, this section focuses on the areas this theory most directly contributes to. Generally, it contributes to research conducted to evaluate how human beings' probabilistic judgements about the future contrast with the basic implications of probability theory (Kahneman et al., 1982; Gilovich et al., 2002).

The RENA in this model relies only on the data in his own sample to predict future outcomes. Hence, the theory is in line with findings psychological literature emphasizing how individual's probabilistic beliefs are driven by observed frequencies (Dougherty et al., 1999; Sieck and Yates, 2001; Nilsson et al., 2005).

The literature most closely related to this theory is the one analyzing the "Law of Small Numbers" and the associated *gambler's fallacy* (Tversky and Kahneman, 1971; Rapoport and Budescu, 1997; Mullainathan, 2002; Rabin, 2002; Rabin and Vayanos, 2010). This brand of research states that individuals know the true data-generating-process (DGP) and believe that its stochastic characteristics will be present and reflected in all (small) samples. As a result, individuals are prone to believe in the "Law of Averages" (Ellenberg, 2014, p. 73): they discount the likelihood of observing unlikely events because they think unlikely events will always average out, even in the short run. Take the example of multiple flips of a fair coin: after observing five coin flips in a row resulting in heads, an individual who is subject to the gambler's fallacy believes the next toss yields tails with significantly higher probability than heads. Thus, such an individual knows the true probability distribution and draws flawed conclusions concerning the possible shape of random samples.

The approach presented here follows a related logic in the sense that individuals behave like all samples fully represent the distribution they were drawn from. However, the thought process and direction of inference are completely reversed: in my approach, the agent does not know the true DGP and its properties. All he knows is his sample data. His error then results from false *extrapolation*: the agent believes that the population's parameters cannot differ from those he calculated with the help of his random sample. Hence, the individual's mistake lies in the inferences he makes about the properties of the DGP, not vice versa. Additionally, in my approach, the actual significance of an agent's fallacy is dependent on the actual sample size which is unaccounted for by the theory of the "Law of Small Numbers".

Furthermore, the implications I derive are fully in line with Tversky and Kahneman (1974), stating that individuals appear to be insensitive to the fact that random variation is dependent on the sample size when comparing the informativeness of two samples with different size. Nevertheless, the context considered here is different. In this approach, no difference in informativeness between samples is evaluated. In fact, the RENA behaves as if no other sample but his own exists.

In a broader sense, this approach also contributes to the literature of Anchoring (Tversky and Kahneman, 1974) and, similarly, the central tendency bias (Crosetto et al., 2020). Generally speaking, both theories state that individuals tend to stick too much to their initial best guesses, for example, of the mean, and then fail to adjust their judgement sufficiently away from it. The theory of the RENA can be interpreted as providing a rationale on why people appear like they tend to focus “too much” on their initial estimates (here: their point predictor) and what might determine the extent of this bias. In this sense, this theory can serve as an endogenization about how an anchor is formed and why individuals appear to be excessively confident in it.

Further, this approach contributes to the extensive literature about overprecision as a form of Overconfidence (Moore et al., 2015; Simon and Kim, 2017). Overprecision has been used as an explanation for empirical findings/patterns in, for example, political decision making (Ortoleva and Snowberg, 2015), asset valuation and trading decisions at the stock market (Odean, 1999) or managerial behavior (Simon and Houghton, 2003; Malmendier and Tate, 2008). This theory contributes to this body of literature by offering an *endogenous explanation* not only for observations in line with this bias but also its origins. In this approach, overprecision does not arise from motivational factors like wishful thinking, which is in line with existing evidence (Logg et al., 2018). Neither does it arise due to the use of a biased estimation strategy concerning the data variability of a population (Kareev et al., 2002), for which the empirical findings are mixed (Kaesler et al., 2016). Rather, it may be a *symptom* originating in judgements suffering from negligence with respect to the random error.

1.3 The Random Error Neglecting Agent

The theory considered in this chapter specifies how an agent suffering from random (or sampling) error neglect forms a belief about the likelihood of a future event under uncertainty. I simply assume that an agent wants to make a probabilistic assessment as correctly as possible. No other preferences will be modeled explicitly, i.e. no explicit utility function is derived. In contrast to the model of fully rational agents which are most prominent in economic models, I assume that no agent has a perfect understanding of the underlying DGP and its parameters, be it the first and second moment of a probability distribution or the slope and intercept parameters of a regression function. The majority of situations in reality do not allow for perfect knowledge of the underlying distribution. Hence, the more realistic scenario is when the true DGP is unknown, or even cannot be known by a single agent. It is more interesting to consider how predictions based on random samples are forged since inferences, as well as forecasts, are generally derived by using limited data collected from the unobservable true distribution.

Thus, throughout this chapter, it is assumed that individuals do not perfectly understand all properties of a random variable X, i.e. its exact distribution and the respective parameters. They only consider a finite random sample, consisting of n fully random draws from this distribution. This sample is denoted by $G = \{A_1, ..., A_n\}$, with A_i representing a discrete data point of sample G.

Note that I do not make an explicit assumption about where G actually comes from. One possible source might be the notion of Gennaioli and Shleifer (2010) who use a purely "mental model" originating from Kahana (2014). It states that in their long run memory, individuals are capable of storing the true DGP. However, when making decisions, people rely on their operative or short-term memory which only has a bounded capacity due to cognitive limitations or external factors like time constraints. Thus, coming back to the model in this chapter, you could think of the random sampling process as being purely "mental" in this case. In contrast to Gennaioli and Shleifer (2010), this sampling process is not spoiled by a perceptive bias that lets the agent consider only specific data points. In my case, the likelihood of a discrete outcome to be considered by the agent depends only on its true probability of realization. Another way to interpret this model's set-up would be to think of an agent who actually has access to physical data in form of a random sample that he or somebody else has drawn from the population he does not or cannot know without incurring prohibitively high cost. Nevertheless, even though both notions appear rather different, the important point in both cases is that the agent does not suffer from any perception bias and thus, does not deliberately underweight or ignore some realizations. Thus, his sample can be considered to be fully randomly drawn from the population.

I further assume that individuals are capable of always applying the *best possible* estimation strategy for a problem in hand. In other words, they use the same estimation strategy as a sophisticated statistician who knows and understands the respective theoretical background perfectly. This implies, for example, the use of unbiased maximum likelihood estimators for making guesses about a distribution's parameters like mean and variance which in turn will then be used to predict a future outcome and its certainty. This setting serves the purpose to eliminate the simple use of heuristics or biased estimation strategies as potential drivers for the presented results.[1] The used optimal estimation strategies depend on the specific context and thus, will be stated explicitly in the following sections.

[1] There is an ongoing debate about whether individuals use biased estimation strategies. For example, in the psychological literature, it is frequently postulated that individuals systematically underestimate the true Variance of a Random variable (Kareev et al., 2002). However, the evidence for this claim is rather mixed and not conclusive (Konovalova and Le Mens, 2018).

So far, there is no difference between the assessment strategy of a rational, i.e. statistically sophisticated, agent considering a sample and the one of a RENA. However, the way how both kinds of agents treat and understand their estimates once they are calculated is different. The crucial assumption here is that the statistically sophisticated agent understands that her estimators are *random variables*: their exact realization depends on the specific sample data drawn from the population. Thus, those estimates follow their own distribution with their own dispersion, implying uncertainty that needs to be accounted for when using them for making statistical inferences. Take, for example, the simple case of a normally distributed random variable X with mean μ and variance σ^2. When estimating the mean of X, the maximum likelihood estimator is the arithmetic mean of a random sample, i.e. $m = \frac{1}{n}\sum_{i=1}^{n} A_i$. The sample mean m, however, is not a constant value, but normally distributed with mean μ and variance $\frac{\sigma^2}{n}$. Its variance term captures the random error, the potential variations in the realizations of m for different samples.

The RENA, however, naively thinks that his assessments of the parameters based on the sample data are equivalent to the true population parameters. Hence, he is either assumed to be not capable or just to be too ignorant to understand the statistical property that a random sample (with specific parameters) is not a perfect representation of the underlying population's distribution. In other words, he is negligent about the *random error* innate in estimates established via limited data. Statistically speaking, he confuses unbiasedness with correctness. This is per se problematic because unbiasedness does not mean that every single individual's guess is correct.[2] It just means that *on average*, this estimator will yield the true value. The RENA, however, falsely relies on his estimates as if they had no variability when forming an opinion about the likelihood of an outcome. In the context of the example from above, the RENA does not understand that m has a variance $\frac{\sigma^2}{n}$ which is larger than zero for finite n. One can think of this type of agent as some naive individual who just has read about optimal estimation strategies and wants to apply them to a data set without understanding the statistical properties of the methods.

[2] Taken the laws of probability as given, it is very likely that not a single estimate hits the *correct* expected value.

1.4 General problems of confusing random variables with constants

One could be tempted to think that when agents are using the maximum likelihood estimator, i.e. the best possible estimator, and the true parameters are not or cannot be known, they use the "second-best option" available and thus, the RENAs behavior and drawn inferences are always indistinguishable from those of a fully rational individual who is using limited data from a random sample $\{A_1, ..., A_n\}$. However, as I show in this section, this logic is only partially correct. Consider a random variable X that is normally distributed with mean μ and variance σ^2. The ML estimator of μ in this case is the sample mean, i.e. $m = \frac{1}{n}\sum_{i=1}^{n} A_i$. As a result, when just estimating the true mean, neither type of agent makes a systematic/biased estimation since $E(m) = \mu$. Another way to phrase this is that both, the estimates of large groups of RENAs in which the individuals aggregate all available information from the idiosyncratic samples should be as correct as those made by a large group of fully rational agents. This result, however, does not hold generally for all estimations made by using m:

> **Proposition 1**
>
> *Consider the case in which a RENA makes an estimate not based on the sample mean m but on a function f(m).*
>
> *1. If f(m) is linear, the RENA's estimate is still correct on average.*
>
> *2. If f(m) is strictly convex or strictly concave, the RENA's estimate is systematically biased.*

Proposition 1 shows a potential problem: what if the agents are tasked with making an estimation based on a function f of their sample mean, instead of just estimating μ via m?
Given f is *linear*, the unbiasedness result still applies.[3] If, on the other hand, f is *strictly convex* or *strictly concave*, this result no longer holds. As Jensen's inequality (Jensen, 1906) states, if f is *strictly concave*, $E(f(X)) < f(E(X))$ holds for X being a random variable, and vice versa, for f being *strictly convex*, $E(f(X)) > f(E(X))$. Thus, applying non-linear functions to random variables is non-trivial concerning biasedness.
The RENA, however, is ignorant about this property because he is ignorant about his estimate *being* a random variable. He believes that his sample estimate is equivalent to the true pa-

[3] It can be shown that for any linear function f, $E[f(m)] = f(E[m]) = f(\mu)$ (Fahrmeir et al., 2016).

rameter of the distribution since he neglects the random error. Thus, he falsely believes that applying a function to his estimate is equivalent to applying a function to a constant with no variation. Hence, when tasked with making a statistical inference by using a non-linear function of, for example, the sample mean, his estimate will be biased on average.

In contrast, the rational agent will not make such a mistake because she is aware of the above-stated statistical property.

This finding has two implications. First, it shows a limit to the phenomenon called "Wisdom of the Crowds" (WotC). WotC states that collecting information from various individuals cancels individual noise and leads to on average correct assessments and improved decisions, especially in comparison to choices made by a single decision maker (Surowiecki, 2005). For a group of individuals consisting of RENAs, however, this only holds for as long as the task in hand does not require the use of non-linear functions of their sample estimates. Else, the group's assessment will be systematically biased.

Second, related to the first point, this theory shows a potential channel about how framing effects (Tversky and Kahneman, 1989; Paese et al., 1993) concerning the specific way a question is asked may affect a decision maker. My theory states that two different questions both requiring the use of the same estimate may lead to different results with respect to average correctness of the responses. Consider the stylistic example in which question 1 asks: "What is the average return per month of Stock X?" while question 2 asks: "How many months will you on average have to hold Stock X to collect an overall return of B?" Given a finite sample of monthly returns for Stock X, the arithmetic mean is an unbiased estimator of the required mean rate of return and hence, a RENA's answer to question 1 will on average be correct. The second question, however, leads the agent to use the *inverse* of the arithmetic mean, which is a strictly convex function. Thus, a RENA's answer to question 2 will be on average *too high*, i.e. on average, the amount of months necessary will not be estimated correctly.

Hence, even without the necessity of stimulating an emotional response (Cassotti et al., 2012) or the use the availability heuristic (Tversky and Kahneman, 1973; Folkes, 1988), this theory states that framing might have an effect simply because of random error neglect and well-known statistical properties.

1.5 Flawed Conclusions about the Predictability of the Future

In this section, I show how the RENA's probabilistic assessment of a *future realization* of one or more random variables differs from the one made by a rational agent. In some sense, it is an extension to the previous section, where the results followed from agents using their estimated mean rate to answer *specific questions* about the central-tendency parameter of an unknown distribution, or, in other words, about the *average* outcome of a future random draw.

In contrast, the focus in this section lies on *interval predictions*, used to quantify the degree of uncertainty associated with an unbiased *point predictor* as the sample mean. There are several reasons for making this distinction. First, the true likelihood that a future realization of a continuously distributed random variable hits *exactly* its true mean value is equal to zero (given a non-zero population dispersion). Hence, some degree of uncertainty always needs to be accounted for.

Second, for some discrete distributions, the mean itself will *never* be an actual realization (for example, when rolling a dice). Hence, just relying on point predictions when forecasting the future is unsophisticated in most cases.

Considering a random sample drawn from the true distribution with sample size $n > 1$, it appears only reasonable that even the RENA who believes the sample to contain all necessary/available information will notice these uncertainties for his calculated mean rate, at least to some degree. The question then is, if and how much the assessment of a RENA about the uncertainty of his predictions differs from the one of a rational agent. An intuitive way to show those differences is to compare the respective average *prediction errors* both types of agents calculate. Those are an integral component for, e.g., the specification of all kinds of *certainty intervals*. Thus, another way of depicting the different assessments would be to show discrepancies in the calculation of the bounds of such intervals.

Confidence vs. Prediction Intervals

It is important to note at this point that, when speaking of certainty intervals in this context, I generally refer to *prediction intervals* (PIs), or as they are sometimes called *coverage intervals* (Vardeman, 1992; Poulson et al., 1997). The reason for using this type of interval instead of the more famous *confidence interval* (CI) in this context is that when dealing with an uncertain outcome of a future event, the primary goal is not to establish confidence bounds for the estimates of the parameters of a distribution. Given a sample mean m, when predicting the future, there is only limited insight in knowing the degree of certainty by which this sample-specific estimate captures the true mean.

A *prediction interval* on the other hand gives information about a range of values, the *next* observation will fall in at a given degree of certainty. Specifically, a 95% PI centered at the sample mean m states a 95% likelihood that the value drawn next from the same distribution will be contained in this interval. For further clarification of the differences, consider the following exemplary question. Assume you want to predict the time you need to go to work the next day based on your collected data from the last months. A confidence interval only tells you that your *true average travel time* lies – for a given level of certainty – between X and Y minutes. A prediction interval on the other hand will tell that the next trip will take between X and Y minutes for a given level of certainty. Hence, it is logical to use a PI instead of a CI in this context.

1.5.1 Overoptimism concerning own predictions

First, I show how a RENA's uncertainty-assessment is flawed when predicting the next realization of a single random variable, even though he uses an unbiased estimator as a point predictor.

Consider a random variable X that follows a normal distribution with mean μ and variance σ^2. The first parameter refers to the central tendency, the second one to the dispersion of the distribution. Further, assume a random sample drawn from this distribution consisting of n points of data. The task is to make a prediction for the next draw from this distribution, i.e. the value of A_{n+1} and to specify the associated prediction error. How does the assessment of the RENA look like in this context in comparison to a rational, i.e. statistically sophisticated, agent?

Proposition 2

Given that $X \sim N(\mu, \sigma^2)$ and $\{A_1, ..., A_n\}$ is a fully randomly drawn sample consisting of n realizations of X. In this case, the RENA underestimates the true prediction error for the next realization of X by a factor $\left(1 + \frac{1}{n}\right)$.

I assume here that the RENA makes no systematic estimation mistake for the first and second moment of the distribution. He always uses the best possible unbiased estimator as a foundation for his predictions. Thus, he calculates his sample mean according to $m = \frac{1}{n}\sum_{i=1}^{n} A_i$ and the sample dispersion according to $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(A_i - m)^2$. Hence, on average, neither mean nor variance will be systematically over- or underestimated.[4]

The problem is, however, that he believes that all relevant information is contained in his random sample. In other words, he believes he considers the whole data dispersion when in fact he only has an estimate in form of the sample dispersion. Since in his opinion, there is no other source of uncertainty in this case, he believes his prediction error to be just s^2. Or to phrase it differently, he believes that any future draw A_{n+1} has a mean m and a variance s^2.

As a result, on average, the RENA's belief about the prediction error is equal to $E[s^2] = \sigma^2$. This assessment, however, contains the problem: the RENA does not think about m and s^2 as sample dependent random variables but as constants.

A rational agent, when using the same sample for a prediction of the next random draw from this population while not knowing its true parameters will not only consider the randomness coming from the population dispersion (σ^2, or its estimate s^2). She will also incorporate the random error coming from *the estimate itself*: when using m as the center of a prediction, you have to account for the fact that it is a random variable following its own distribution and having its own variation. If $X \sim N(\mu, \sigma^2)$, then $m \sim N(\mu, \frac{\sigma^2}{n})$. This implies that the lower the sample size under consideration is, the more likely are substantial deviations from the true population characteristic, which carry over to forecasts made on this foundation.

Hence, as shown in von Auer (2005), the true uncertainty is captured by the variance of the point prediction, i.e. $Var[m - A_{n+1}]$. As is shown in Appendix (A.1), this *true* variance for a prediction concerning a future draw A_{n+1} is given by $\sigma^2\left(1 + \frac{1}{n}\right)$. By comparison, it follows that

[4] The proof of unbiasedness for both estimators is omitted here but can be found in any introductory statistics book, for example, Fahrmeir et al. (2016).

the true average prediction error is larger than the average of the RENA's assessment[5] by a factor $\frac{n+1}{n} > 1$, proving Proposition 2.

Thus, the smaller the sample size, the stronger is the effect of neglecting the random error and the more error prone are the probabilistic inferences drawn by the RENA. Generally speaking, for any finite $n > 1$, the RENA believes the future to be more predictable than is statistically justifiable because he neglects a significant source of the prediction error in form of the random error contained in his sample estimates. This neglect would, for example, manifest in Prediction Interval bounds being calculated too narrowly. Hence, they will differ from the ones made by a rational agent by $\sqrt{1+\frac{1}{n}}$, which can be quite significant given n is not too large.[6] To phrase it differently, for finite sample sizes, a RENA suffers from *overprecision* (Bazerman and Moore, 2012; Moore and Schatz, 2017), which in this context means that the RENA has statistically unjustified confidence about his knowledge of the outcomes of random events.

This result yields direct implications to various fields of economics dealing with decision making under uncertainty. A prominent example is contract theory, dealing with the optimal design of incentive schemes by a principal who requires an agent to complete a given task. The result presented above implies that efficient incentivization for a non-risk-neutral agent to exert effort when the outcome is uncertain requires more than knowledge about the agent's true degree of risk tolerance or aversion. In the standard case (Milgrom and Roberts, 1992, Chapter 7), the optimal "pay-for-performance" rate for a risk-averse agent does not only depend on his effort cost function and his degree of risk-aversion but also on the variation of the outcome. If the agent is a RENA, in addition to being risk-averse, his judgement of the uncertainty is flawed and the risk is underestimated. Thus, even though he is risk-averse, this theory states that an agent who neglects random errors in his probabilistic assessments is willing to accept more convex incentivization schemes and thus a higher risk than standard theory would suggest. In line with e.g. Silver (2012), this theory thus states that overprecision may take shape in form of understatements of risk and a failure of self-protection against it by decision makers, like managers or investors.

[5] The use of a biased estimator for the population variance, i.e. $s'^2 = \frac{1}{n}\sum_{i=1}^{n}(A_i - m)$, only amplifies this result.

[6] See Geisser (1993) on the correct specification of interval bounds for Prediction Intervals given population mean and variance are unknown.

1.5.2 Inordinate Scepticism concerning different views

The overprecision problem shown in the previous section, however, does not hold universally. Interestingly, it is only present when it comes to the evaluation of the *own* prediction's accuracy. When being tasked with the evaluation of a forecast provided, for example, by another agent, the theory states that the RENA will suffer from exactly the opposite problem:

> **Proposition 3**
> *Assume that a RENA is provided with an exogenously derived predictor z for the next realization of X. In this case, the RENA overestimates the true prediction error of z on average by a factor $\frac{n}{n-1}$.*

Proposition 3 states that the RENA will on average overestimate the prediction error of a forecast not made by himself. In other words, he will be unjustifiably sceptic about the accuracy of such a predictor.

To see this, assume that there exists an independent point predictor z to be evaluated. The agent has no information about how and on what basis z has been calculated, he just takes it as externally determined.[7] Further, as he believes to have calculated the true parameters already, he does not adjust his own prediction but just evaluates the likelihood – or error – of the external prediction by comparing it to the data points in his sample.

Appendix (A.2) shows that, as a result, the RENA on average underestimates the reliability of z by a factor $\frac{n}{n-1}$ in comparison to the correct evaluation. Again, especially for small n, this can be a quite significant deviation. It follows that the bounds of his intervals specified for z are too broadly calculated, in other words, he suffers from *underprecision* in this case. The reason for this result lies in the fact that his neglect of the random error in his estimation of the sample-moments, being the main driver for overprecision in the previous section, now works as a *counter-effect* for his belief in the accuracy of z. To phrase it differently: *Because* the RENA is too optimistic about his own prediction's accuracy, he is excessively pessimistic concerning

[7] Note further, that this also implies that no additional assumption about the quality of an externally provided predictor is made, i.e. the quality of the results does not depend on the actual "correctness" of z. You could for example assume, that an expert's assessment is significantly closer to the true parameter μ than a random guess provided by an individual totally unfamiliar with the problem in hand. The RENA, however, does not attach any weight for – potentially spurious – competence.

those made by others.

Stating this testable hypothesis about how different individuals treat a prediction made by themselves in contrast to those made by other agents, for example, their colleagues/other team-members, this theory also contributes to the economic literature regarding how agents perceive and evaluate the opinion/advice stated by others and how they update their beliefs thereupon (Harvey and Fischer, 1997; Bonaccio and Dalal, 2006). There is robust evidence that individuals tend to rely more on their own assessments and show higher doubt with respect to the judgmental accuracy of third party advisors (Van Swol and Sniezek, 2005; Gino and Schweitzer, 2008; Mannes, 2009). The most prevalent explanation attempts to explain this phenomenon relies on various forms of "personal motivation", for example, whether a decision maker is "powerful" (Tost et al., 2012), whether there exists some kind of personal bond between the decision maker and the other agent (Kadous et al., 2013) or whether this mindset is subconsciously created to sustain a feeling of one's superiority over another person, maybe even an expert. This theory offers a different explanation of why individuals may devaluate other people's beliefs that is not routed in motivational reasons: simply because people have too much faith in their own sparse data. They just do not take into account that their results might be random and there is more to be known.

1.5.3 The case of two random variables: inaccurate assessment of (future) differences

So far, I have only considered predictions made for the next realization of a *single* static random variable. The theoretical approach presented in this chapter, however, can easily be extended to the case in which an agent has the objective of assessing the differences of *two* distributions' properties as a foundation for a decision to be taken. Consider, for example, a manager whose goal it is to access the superiority of a product's quality over another. This can be the case in a pharmaceutical context, when a decision has to be made about what specific medication yields superior results. It also might be a provider of an internet platform who wants to evaluate the effectiveness of a newly adopted merchandising method concerning changes in user behavior.[8]

In both cases, it is important to estimate the first and the second moments of the respective

[8] Statistically speaking, a case like the first example is considered to be an *independent two-sample test*, the latter a *dependent (paired) two-sample test* (Imbens and Rubin, 2015).

distributions correctly in order to figure out potentially significant differences in the means of both distributions. Furthermore, it is vital to calculate the bounds of *future* mean-differences for *future* samplings, for example, conducted by a regulatory agency or simply the next customer who compares two goods produced by different firms (Hahn, 1977). Thus, in terms of prediction, the task in hand is to establish the upper and/or lower bounds which contain the future difference in (mean-)performance with a (reasonably) high certainty level.

Consider the case of two potentially related random variables X and Y, for example, the qualities of two products, which are normally distributed with given means and variances, i.e. $X \sim N_X(\mu_X, \sigma_X^2)$ and $Y \sim N_Y(\mu_Y, \sigma_Y^2)$. The co-variance between the two variables is denoted by $\sigma_{X,Y}$. [9]

The RENA thinks about both distributions only in the form of two randomly drawn samples. Each sample is drawn from the two distributions, respectively. For the sake of simplicity, it is assumed that both samples have the same size n. Let $G_1 = \{A_1, A_2, ...A_n\}$ be the sample drawn from N_X and $G_2 = \{B_1, B_2, ..., B_n\}$ from N_Y. As before, the RENA is fully capable of calculating the arithmetic means and sample standard deviations accurately and without bias. Let them be denoted as (m_X, s_X) for the first and (m_Y, s_Y) for the second sample. Comparing both sample means $(m_X - m_Y)$ yields again an unbiased estimator for the true mean difference $(\mu_X - \mu_Y)$.

However, the negligence concerning the random error, now present for both samples under consideration, again leads to flawed conclusions for the RENA:

Proposition 4

1. Assume that $\mu_X > \mu_Y$. Given that the distance between both means is not too high, there is a statistically significant chance that a RENA will be certain that $\mu_Y > \mu_X$.

2. Assume that $\mu_X = \mu_Y = 0$. Then, on average, the RENA will expect at least one mean to be significantly larger than zero, given the sample size n is not too large.

3. On average, the RENA underestimates the true prediction error of his forecast based on $m_X - m_Y$ by a factor $2\left(\frac{n+1}{n}\right)$.

[9] For independent random variables, it holds that $\sigma_{X,Y} = 0$, which does not change the results presented below.

First, Proposition 4 states that due to m_Y and m_X being random variables and given that there exist a sufficiently large range of values covered by *both* distributions,[10] there is a significant chance that a RENA will hold false beliefs about what average is higher. To state it intuitively by using the example from above in which an agent wants to find out which of two products has superior quality: there is a positive probability that he is certain that the product with lower average-quality is superior.

Second, using this example further, Proposition 4 states that there is a significant chance that the RENA will expect an average quality of one of the goods that is exaggeratedly high, given the samples are not too large. To show this, consider the case in which $\mu^x = \mu^y = 0$. Further, for simplicity and without loss of generality, assume that $\sigma_X^2 = \sigma_Y^2 = 1$, i.e. both products are qualitatively equivalent because both variables follow a standard normal distribution. This implies that both of the RENA's mean estimates, m_X and m_Y, are random variables that follow the same normal distributions, which is $N(0, \frac{1}{n})$. Hence, just comparing the two arithmetic means of the two samples is like conducting two independent random draws from the same distribution. Finding the highest mean is then statistically equivalent to finding the largest of two order statistics for this distribution. Using the approximation formula by Blom (1958) to calculate the respective expected value,[11] the expected mean of the highest of the two order statistic of a random variable distributed according to $N(0, \frac{1}{n})$ is equal to $0.58\frac{1}{\sqrt{n}}$. Thus, especially for small n, it is significantly higher than zero, the true average value of both random variables. Note, however, that this result does not contradict the statement of unbiasedness from before. The reason lies in the fact that the theory just states, that *one* of the products can falsely be judged superior, not that it has to be always *the same product* for all agents.[12] A RENA will expect a higher quality of one of the products than would be justified in this case. Hence, the final realization is likely to lead to some kind of "disappointment".

Third, Proposition 4 states that a RENA ignores the potential fluctuations in the difference of sample means. As a result, his predictive inference again suffers from overprecision in the sense that he *underestimates* the prediction error of his forecast based on $m_X - m_Y$. In com-

[10] I.e. $\mu_i \pm \mu_j \leq 3\sigma_i$, $i, j \in \{X, Y\}$ and $j \neq i$.

[11] Explicitly given for a random variable X following a normal distribution with mean μ and variance σ^2 by $\mu + \Phi^{-1}\left(\frac{r-\alpha}{N-2\alpha+1}\right)\sigma$, with N being the sample size, r the r^{th} largest order statistic and α being set equal to 0.375.

[12] In fact, the symmetry of the Normal distribution at hand guarantees that when aggregating all samples of all RENAs, these misestimations will cancel out perfectly.

parison to the prediction problem with only one random outcome the overprecision problem even increases because in this case, the RENA suffers from an increased effect of the random error neglect. As is shown in Appendix (A.3), under the simplifying assumption of equal variances of the two random variables, this implies that the miscalculation of the prediction error by the RENA is on average exactly twice as high as in the previous section. Speaking in terms of Statistical Intervals: to calculate the upper and lower bound of a prediction interval, the RENA will on average use a correction factor that is too small and needs to be multiplied by $\sqrt{2\frac{(n+1)}{n}}$.

Generally speaking, he specifies prediction intervals too narrowly and thus, exhibits too much confidence in the predictive power of his initial point-estimate of the mean-differences. Hence, this theoretical set-up postulates that agents who suffer from random error neglect tend not only to perceive but also to predict significant differences in outcomes/products more often than it would be statistically justified. Especially in combination with the possible assessment errors of the true parameters outlined above, this may lead to seriously flawed conclusions.

To illustrate that point, suppose the most extreme case of a RENA with $n = 1$, meaning that both samples consist of only one random draw (i.e. $G_1 = A_1$ and $G_2 = B_1$). Neglecting the fact that these are only samples of the true populations, the RENA believes that the two data points are the only realizations possible. He calculates $m_X = A_1$, $m_Y = B_1$ and $s_X = s_Y = 0$.

Assuming that both random variables follow a continuous (normal) distribution, the likelihood of $A_1 = B_1$ is equal to zero. As a result, such a RENA will always believe that (significant) differences between the two variables (for example, product qualities) exist and will exist *in all future samples.* In other words, not only is a RENA in danger of rejecting the hypothesis "$\mu_X - \mu_Y = 0$" too often, he also falsely believes that mean-differences *measured in future samples* will resemble his own estimate for this difference ($m_X - m_Y$) too strongly.

Such kind of overoptimism offers an explanation for, e.g., unjustified beliefs of economic agents in a product's superiority in quality (Spiller and Belogolova, 2017) or the persistence of gender-stereotypes/social norms (Bursztyn et al., 2018)

1.6 RENAs and Regression Analysis: False Faith in Trends

In the previous section, I focused on prediction problems concerning the next realization of one or two *static* random variables. In this section, I demonstrate that the theory of the RENA can also be applied for predictions made via regression analysis, especially time-series forecasts.

It is important to consider this area because, especially in business, when predicting, for example, the quarterly sales of a company, it is not only important to understand the factors influencing the outcome under consideration but also *to what extent* they do. Further, considering predictions of ex-ante unknown outcomes, it is often vital to specify the degree of certainty by which the estimated connection will uphold, or particularly, considering time-series analysis, how time-persistent it will be. Using this prediction method is particularly appealing because as Jayachandran (1983) or Box et al. (2015) show, autoregressive processes of the form $x_t = \rho x_{t-1} + \epsilon_t$, with $|\rho| < 1$ and $x_0 = 0$, can also be expressed in form of a linear regression model. Hence, the following section can be seen as an extension of the previously presented theory to prediction tasks considering the next outcomes of stochastic processes, too.

Assume that the true relationship between an outcome variable y and a predictor/explanatory variable x is captured by a simple linear regression model (Angrist and Pischke, 2008):[13]

$$y_t = \alpha + \beta x_t + \epsilon_t \tag{1.1}$$

where $t \in \mathbb{N}$ is the (time-)index for a specific pair of observations (y_t, x_t) and their corresponding error term ϵ_t. α denotes the intercept, β the slope parameter. For (1.1) being the true model, I implicitly assume that no relevant predictor variables are missing, x_t itself is not irrelevant for explaining movements in y_t, the relationship between (y_t, x_t) indeed is linear and α and β are constants for all observations. The reason to include a random error term in equation (1.1) even if it depicts the correct model is to account for unforeseeable (economic) shocks or idiosyncratic unobservables for an individual measurement. I assume that the error

[13] Even though the model here appears to be rather simplistic, the same analysis can be replicated by using a regression model with multiple explanatory variables, including, for example, a simple trend variable t or a quadratic term. The results presented do not change, they only become more distinct. Hence, for the sake of simplicity, I use the most basic version for my analysis.

terms ϵ_t are normally distributed with zero mean, are not auto-correlated, are unrelated to the predictor variables and all have a constant variance equal to σ^2. In line with the previous analysis, the RENA does not know the true model but relies on a finite sample of size T for his estimates that is drawn fully randomly. Further, I again assume that he makes no systematic estimation mistake, i.e. uses the best-possible estimation method for the task in hand. Thus, in this case, he is assumed to use the Ordinary-Least-Squares (OLS) method to estimate the parameters α and β. Formally, he chooses his parameter estimates $\hat{\alpha}$ and $\hat{\beta}$ to minimize the *Sum of Squared Residuals*, i.e. $\sum_{t=1}^{T} e_t^2 = \sum_{t=1}^{T}(y_t - \hat{\alpha} - \hat{\beta}x_t)^2$. As shown in Rao (2009), the OLS is the *Best Unbiased Estimator* given the assumptions made above. Hence, the agent does not miss-specify the real model (for example, assuming a linear regression while the true relation is non-linear), nor does his forecasting model suffer from *overfitting*.[14] Further, potential mistakes in prediction do not stem from the use of a biased estimator or a distorted sampling process but are solely based on random error neglect. Again, a problem arises due to the fact that, even though he realizes that there is a certain dispersion or error in his sample (here depicted by the residuals $e_t = y_t - \hat{y}_t$, with $\hat{y}_t = \hat{\alpha} + \hat{\beta}x_t$), he fails to understand that his estimates $\hat{\alpha}$ and $\hat{\beta}$ are random variables, following own distributions.[15] Since their value depends on the specific sample realizations, they are likely to differ for each random sample. This, however, is neglected by the RENA since he thinks he captured the true model parameters. In other words, he assumes that he fully understands the slope and the intercept for the relation of *regressand* (y) and *predictor* (x) and all uncertainty concerning future-predictions comes from individual random errors.

Again, he assumes his sample to represent all relevant properties, thus, he believes the variance of the residuals in his sample to be equal to the true dispersion of individual errors. As in previous sections, I again assume that he captures the sample-error dispersion via an unbiased estimator of the true variance of error terms. Under these circumstances, it is denoted by $s^2 = \frac{\sum_{t=1}^{T} e_t}{T-2}$ (von Auer, 2005). It may appear debatable that the agent is assumed to be sophisticated enough to understand that he has to control for degrees of freedom for an unbiased

[14] This term captures the problem of fitting the model more to the noise in the sample than to the real relationship. Generally speaking, an analysis suffers from *overfitting* if it corresponds too closely to a particular sample set of data (for example, assuming too many explanatory variables to specify a regression function that captures all data points with no residual error). Thus, it may fail to predict new observations reliably (Backhaus et al., 2006).

[15] Under the assumptions made above, $\hat{\alpha} \sim N\left(\alpha, \sigma^2\left[\frac{1}{T} + \frac{\bar{x}^2}{\sum_{t=1}^{T}(x_t-\bar{x})^2}\right]\right)$ and $\hat{\beta} \sim N\left(\beta, \frac{\sigma^2}{\sum_{t=1}^{T}(x_t-\bar{x})^2}\right)$.

estimation while in all other aspects, he neglects the sample properties. Nevertheless, the reason I chose this set-up follows the same logic as already outlined above. In particular, using a biased variant like $s_b^2 = \frac{\sum_{t=1}^{T} e_t}{T}$ only leads to the agent systematically *underestimating* the real dispersion of error terms, which in turn only leads to further strengthening the results.

Now, think of an agent whose task it is to make a prediction for a new outcome y_{T+1} for the relationship in place, given a certain value of x_{T+1}.[16] Once more, the goal is to compare the reliability of a forecast made by a RENA with the sophisticated (rational) approach.

Proposition 5

The RENA underestimates the true error for a prediction made via regression analysis to forecast the next outcome y_{T+1} based on x_{T+1} by a factor $\left(1 + \frac{1}{T} + \frac{(x_{T+1}-\bar{x})^2}{\sum_{t=1}^{T}(x_t-\bar{x})^2}\right)$.

The RENA bases his predictions on $\hat{\alpha}$ and $\hat{\beta}$ while accounting for a possible error term. He believes his possible prediction error just to be equal to his assessment of dispersion in error terms s^2 since all other sources of error are neglected by him. As a result, on average, a RENA will consider the prediction error to be equal to $E(s^2) = \sigma^2$. In other words, while he accounts for the possibility that in his prediction for y_{T+1} (given by $\hat{\alpha} + \hat{\beta}x_{T+1} + \epsilon_{T+1}$) the error term might take a value different from zero, he neglects the fact that both parameter estimates, even though they are unbiased, might differ from the true values.

The sophisticated calculation, however, accounts for that. As is shown in the Appendix (A.4), the correct expected prediction error is given by

$$\sigma^2 \left(1 + \frac{1}{T} + \frac{(x_{T+1} - \bar{x})^2}{\sum_{t=1}^{T}(x_t - \bar{x})^2}\right) \tag{1.2}$$

which proves Proposition 5.

In contrast to the RENA assessment, expression (1.2) shows the dependency of the true prediction error on sample size as well as on the dispersion of the predictor-realizations. A prediction can be considered more reliable, if the number of observations (T) and the squared dispersion

[16] Note that this forecasting procedure is called "Ex-post forecasting". "Ex- ante forecasting" refers to x_{T+1} as unknown and to be predicted as well (Armstrong, 2001). Since the latter case only increases the uncertainty under which a prediction has to be made, it only strengthens the results presented here and hence, is not considered.

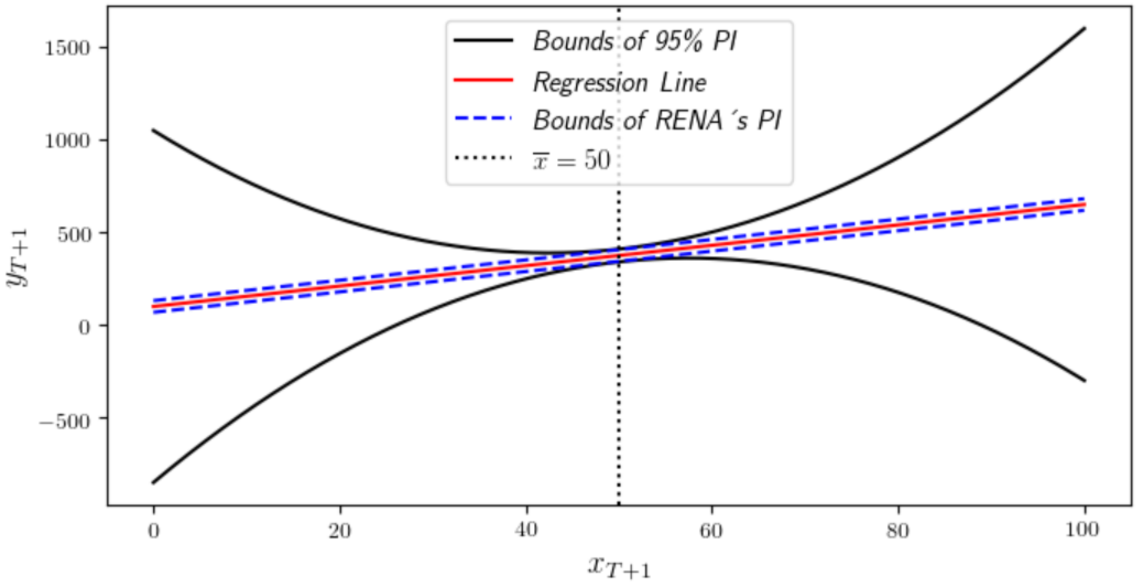


Figure 1.6.1: Simulated Prediction Intervals with $T = 20$, $\sigma = 16$, $\alpha = 100$ and $\beta = 5.5$.

around the sample mean, $\sum_{t=1}^{T}(x_t - \bar{x})^2$, are large. Further, this expression shows that the uncertainty of the prediction for y_{T+1} increases quadratically, the more distinct the respective predictor variable x_{T+1} is from the realizations within the sample. Intuitively speaking, one can be less certain of one's prediction when one leaves the *center of information* one's estimates are based on. Speaking in terms of boundaries of prediction intervals of a given level of significance, the bounds for a prediction of y_{T+1} resemble parabolas with a vertex at $\bar{x}$.

The problem is that the RENA does not account for either of these sources of uncertainty. Not only does he underestimate the prediction error because the random error is neglected, he also ignores the issue of predicting a value that relies on a predictor far away from $\bar{x}$, for example, a prognosis referring to an outcome lying far in the future. As he believes to capture the model accurately and his parameter estimates are constants, there is no reason to believe that this will change, even for rather extreme predictors. Again referring to the subject of prediction intervals: the bounds of certainty of which he believes contain a future value y_{T+1} given x_{T+1} at a given level of confidence are not only too small, they can be represented as straight lines running parallel to the estimated regression line. How large the discrepancies can actually be is graphically captured in Figure 1.6.1.

Hence, the consequence is an even stronger overprecision than in the sections before: it also increases the more distant x_{T+1} is from the sample mean of the explanatory variable. This is of particular importance in the context of a time-series analysis: the further away in the future a prediction based on a sample of current data lies, the more uncertainty must be taken into account. The RENA, however, does not account for the increased uncertainty.

This has some notable consequences. Especially for a small amount of data, a RENA exhibits faith in his trend-forecasts based on his estimates that is not statistically (and hence, not rationally) justifiable. Thus, such individuals are prone to fall for e.g. the *hot hand fallacy* (Gilovich et al., 1985; Barberis and Thaler, 2003) stating that humans tend to believe that a sequence of observed successes (basketball shots by a player, portfolio choices of an investor, etc.) is likely to be followed by another success. To phrase it differently, this theory states that people are prone to neglect the phenomenon called *regression towards the mean* first postulated by Galton (1886). Loosely speaking, this phenomenon refers to the fact that future (long-run) observations can tend towards the true mean of a stochastic process when current (small sample) measurements may show rather extreme realizations (i.e., outliers) due to randomness. As Siegel (2015) shows, this phenomenon is, for example, observable for stock returns which may show great volatility in the short run, with some being significantly above, others significantly below average, but are rather stable/similar in the long run. A RENA, however, is not able to account for such smoothing in the long run because he considers his sample (and his estimated parameters) to contain the ultimate truth about the respective connection. Hence, he perceives the concept of a less volatile long run and a (more or less) stable conversion as rather unlikely or implausible.[17]

To put this result in an economic context: managers using time-series-regression analysis while suffering from the random error neglect will be exceedingly in danger of perceiving the future rather as a replication of past events, ignoring that those particular events will – with great likelihood – not occur again. Relying too much on a small set of (time-series) data might provide a rationale for predictions like "There is no reason why anyone would want a computer in their home", made in 1977 by Ken Olson, the Co-Founder and President of Digital Entertainment Corporation (Stansberry, 2011), just a few years ahead of IBM introducing their first versions of personal computers and becoming undisputed market leaders. Another even more well-known example is the famous economist Robert Malthus who predicted in the

[17] Even if one does not account for the possibility that the underlying model might completely change for the long run.

18th century that in the future, food shortage and misery for mankind are inevitable (Malthus, 1798). Even though the trends at that time appeared to be supportive of Malthus' argument (linearly growing GDP while exponential population growth), that trend broke (in fact even reversed for many western civilizations) due to, for example, the industrialization in the centuries after Malthus' death.

Metaphorically speaking, RENAs naively extrapolating a trend based on one sample of (time-series) data are like car drivers who feel perfectly safe driving at high speed at night on an unlit road going straight because they cannot imagine a bend. Thus, they are only safe for as long as no curve appears. They do not consider that an estimated trend, even if captured correctly for a time, is only pointing in the right direction for as long as the trend does not break.

1.7 RENAs and First-Price Sealed-Bid Auctions

Even though the focus of this chapter has been so far on how neglecting random errors influences an individual's predictions of future outcomes, the theory presented is applicable to various other contexts due to its general setting. Whenever an economic agent has to make a decision based on an uncertain future and relies on a random sample instead of complete information of the random process, this theoretical approach may offer a possible explanation for discrepancies in empirically observed behavior and the theory of rational agents. In the following, this is illustrated by using the context of bidding behavior in auctions.

As is shown below, neglect of the random error in combination with limited data as choice-foundation offers some interesting insights about why auction participants might differ in their behavior from the theoretical predictions of the classical theory for single-object private-value auctions. The results presented below do not hinge on risk preferences or emotions as the most common explanatory factors used in the literature and for which existing evidence is rather mixed.

I focus on First-Price Sealed-Bid (FPSB) auctions in this section, i.e. auctions, in which all bids are handed in simultaneously and privately, the highest bidder wins the object and has to pay her bid. The underlying framework used in this section is the canonical model as, for example, described in Chapter 2 of Krishna (2009): M bidders with quasi-linear preferences[18] compete

[18] I.e., bidders are risk-neutral.

in an auction for an object they have an individual valuation for. Let v_i denote the valuation of bidder $i \in \{1, ...M\}$. This valuation is private information of each bidder. The agents have no budget constraints and want to maximize their expected utility with their bidding behavior. The model is *symmetric* in a sense that all values are independently and identically distributed according to some increasing distribution function F on some interval $[a, b]$, with $b > a \geq 0$. For simplicity, it is assumed that the valuations are *uniformly* distributed with $a = 0$ and $b = 1$. It can be shown that the symmetric utility maximizing bidding strategy played in equilibrium in an auction with only rational agents is bidding the expected value of the *second highest valuation*, given the agent's value is the highest. Mathematically, $b_i^* = E[X_1^{(N)} | X_1^{(N)} < v_i]$, with b_i^* being agent i's optimal bid, v_i her valuation and $X_1^{(N)}$ the highest order statistic out of N $(= M - 1)$ other values, which is a random variable following a specific distribution dependent on F. Note that this implies that no agent has an ex-ante incentive to deviate from this strategy as long as he believes all others to behave this way.

How would the RENA behave in such a situation? As in earlier sections, I do not assume that the RENA just follows a false model by assuming that he lacks important information about the underlying structure. In this case, this means that he has no misguided priors about, for example, the numbers of bidders, his own valuation, potentially correlated valuations or the format of the auction. Further, considering the case that he is competing with rational agents, he again uses the best possible action available and no "flawed heuristic" in the sense that he, for example, just always bids his own valuation. Rather, I assume that he knows the symmetric bidding equilibrium strategy played by rational players and why it is played.[19] Thus, even if an agent i is a RENA, he will decide to bid according to b_i^* from above. In this sense, he somewhat behaves like a naive reader of auction theory literature with no real understanding of the strategic reasons determining the optimal bid. He just wants to behave "optimally" given his prior that all others follow the same behavior. The following proposition states the consequences for the bidding behavior of RENAs:

[19] Implicitly, I thereby assume that he thinks of himself as rational and all others to be rational as defined in the standard theory.

Proposition 6
1. For a given valuation v_i, a majority of RENAs will overbid in a FPSB auction, i.e. place a higher bid than $E[X_1^{(N)}|X_1^{(N)} < v_i]$. In turn, a minority of RENAs will underbid.
2. The RENAs who underbid will do so by a significantly higher absolute amount than those who overbid.

Proposition 6 states that the deviation from the fully rational behavior arises when a RENA actually determines his bid: again in line with this model's assumptions, he does not (or cannot) know the true distribution of $X_1^{(N)}$ and as a result, he cannot calculate the exact value of $E[X_1^{(N)}|X_1^{(N)} < v_i]$, given his valuation v_i. Instead, he relies on a finite sample of $X_1^{(N)} < v_i$ and then calculates the mean of this sample.[20] In line with the potential sources of the sample presented in Section 1.3, he might either have a sample based on past experiences, is exogenously provided with the data, or has the true distribution in his long-term memory but can only consider finite data in this operative memory due to capacity constraints. Further, because he neglects the fact of only considering finite random draws, he believes his calculation to be equivalent to the true expectation of the order statistic $X_1^{(N)}$ given that it is smaller than his own valuation v_i. Hence, instead of using all information of the respective probability density $f(X_1^{(N)}|X_1^{(N)} < v_i)$, he again only considers a finite amount of discrete realizations $\{A_1, ..., A_n\}$ and naively believes to know it all.

This is particular problematic in this case because in contrast to e.g. the normal distribution, $f(X_1^{(N)}|X_1^{(N)} < v_i)$ is not symmetric around its mean but *negatively skewed*. This implies that the measures of central tendency of this probability function are not aligned but dispersed, i.e. median and mode are higher than the expected value. It is of particular importance in this case that the mode of the distribution is located on the right hand side of the expected value, because it points to the location with the highest concentration of probability mass for a particular density function. In other words, the mode marks the location of a – in case of a unimodal distribution global – maximum.

[20] To some extent, a theory recently presented in Kasberger (2020) makes a similar assumption. However, in contrast to my approach, it assumes rational bidders and focuses on different bidding strategies.

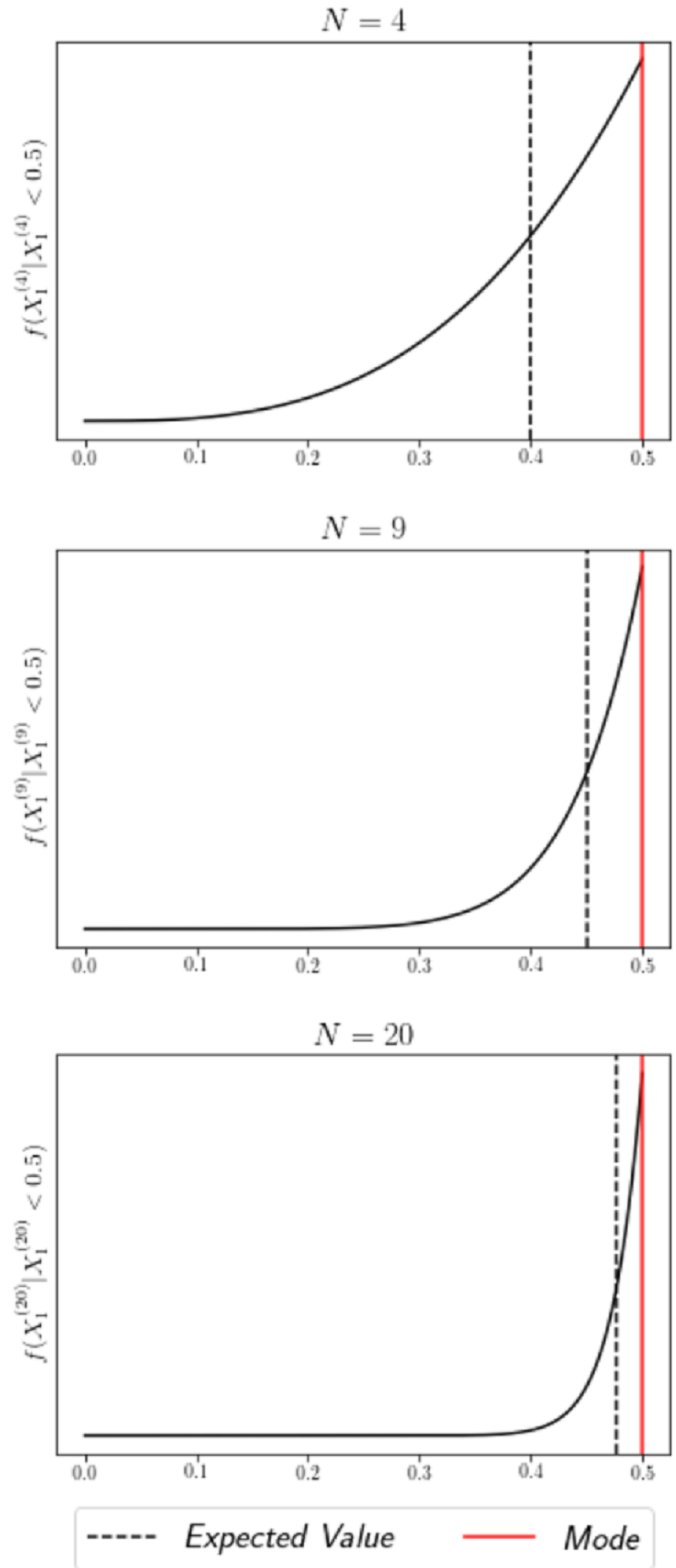


Figure 1.7.1: Probability density functions of $X_1^{(N)}$ for three different N, given $v_i = 0.5$.

Figure 1.7.1 exemplifies that for $X \sim U[0,1]$ and a given valuation,[21] $f(X_1^{(N)}|X_1^{(N)} < v_i)$ has this property for varying numbers of N: the red line, depicting the distributions mode, always lies on the right hand side of the dotted black line, depicting the expected value.

This statistical property is the foundation for the first point in Proposition 6: it states that random samples drawn from this distribution are more likely to contain realizations close to the mode. In terms of the graphic above, RENAs are more likely to randomly draw second-highest valuations close to their own, strictly higher than the correct expected value. Further, this implies that a majority of this type of agent will calculate a sample average that lies above the expected Value of $f\left(X_1^{(N)}|X_1^{(N)} < v_i\right)$. In turn, a minority will calculate a sample mean lower than the correct value. For simplicity and illustration, consider the corner case in which there is an infinite mass of RENAs with the same valuation who separately compete in an auction with N other bidders and who all consider individual random samples with only one data point, respectively. For example, for $N = 20$, the Law of Large Numbers states that around 62% of RENAs will draw a value higher than $E[X_1^{(20)}|X_1^{(20)} < v_i]$. Because they think their estimation of the true average to be fully correct, they follow the "optimal strategy" and place a bid as high as this particular value. Thus, they will *overbid*. In turn, the theory also states that a minority of RENAs (around 38% in this case) will *underbid*.[22]

The second part of Proposition 6 is directly related to the first point: Because overbidding is stated to be more prominent than underbidding, those who underbid must do so by a significantly higher absolute amount. The result arises because overall, the sample mean is an unbiased estimator of $E[X_1^{(N)}|X_1^{(N)} < v_i]$, which means that when all single data points of all RENAs are aggregated to one big sample, this overall sample's mean would be equivalent to the true aggregate. This, however, is statistically only possible if the data points of the above-mentioned minority represent more severe deviations from the true average than those of the majority. Thus, this theory states that one should observe more significant divergences from the symmetrically optimal bid b^* by participants who underbid.

To the best of my knowledge, no "meta-study" exists so far explicitly calculating the shares of over- and underbidders in all auctions using the same framework and comparing the respective amounts by which they deviate from the Risk-Neutral Nash-Equilibrium bidding strat-

[21] Stating it explicitly, the respective probability density is in this case equal to $\frac{1}{v^N} N x^{N-1}$.

[22] The qualitative results do not change for $N = 4$, $N = 9$, or any other finite number N.

egy. Although an empirical test might prove difficult due to the randomness, there exists some evidence weakly in line with the results presented above. Overbidding appears to be a prominent feature in FPSB auctions (Cox et al., 1988) and existing literature used explanations like risk-averse bidders (Harrison, 1989) or emotions (Filiz-Ozbay and Ozbay, 2007; Andreoni et al., 2007) as a potential explanation for this phenomenon. However, in particular the risk-aversion approach fails to explain (substantial) underbidding in such auction formats for which also exists some evidence (Kirchkamp and Reiss, 2004; Ivanova-Stenzel and Sonsino, 2004).

1.8 Conclusion

In this chapter, I derive a new theory about individuals neglecting random errors to analyze how they form beliefs about uncertain realizations of random events and how their resulting conclusions are flawed. The theory presented in this chapter represents a synthesis of statistical theory and decision making by emphasizing how awareness of statistical properties is essential for sophisticated probabilistic assessments and the resulting rational behavior.

Even though the estimation strategies in and of themselves are unbiased and hence, at least on average accurate, RENAs still tend to make systematic mistakes when forming beliefs about future outcomes. The problem can be considered to be some kind of misguided extrapolation: when predicting the range for a future outcome by forming confidence bands, those individuals underestimate the variation in potential realizations because they are too confident in their initial guesses made via the information from their samples. As a result, their interval-boundaries are defined too narrowly. In other words, they consider the future to be predictable to an extent that is rationally not justifiable. I further have shown how this mode of thought can be transferred to already existing economic models. Moreover, due to its broad specifications, the existing model can easily be modified by also including other biases. An example would be an availability bias which lets the RENA put different weights on different realizations in his sample such that some are just more salient. The resulting behavioral implications should strongly resemble those from the literature about selective and limited recall of information (Bordalo et al., 2016, 2017).

An interesting implication of this theory lies in a potential cure for the observed "misbehavior". In contrast to agents who are biased in form of a distorted perception of information such that some data points with certain properties are just ignored, this theory states that the mistakes an individual makes will decrease with the size of the sample he considers. Hence, simply increasing the amount of data an agent uses for his assessment would lead to significant improvements. Given, for example, an employer fears that his employees are suffering from this neglect, a potential remedy might be to require them to think of a certain number of potential results their actions may cause to improve the overall outcome.

Even though I have shown that the theoretical predictions are in line with a variety of established empirical findings, those present no conclusive evidence for the existence of random error neglect. Thus, a desirable next step would be the design and conduct of an experiment to either strengthen or debunk the central elements of this theory. Another potential area of future research concerning theoretical extensions is to evaluate how RENAs might learn from new data. The theory presented here is purely static and frequentistic without considering, for example, if and how individuals might reconsider their initial estimates after they learned that they actually made a mistake. If they incorporate the new realization in their sample, the following guesses should become more and more sophisticated. However, that is just one possible learning rule amongst many that may be considered.

The question on how humans form (and revise) beliefs about the future is and probably will remain fascinating. The goal of this work is to contribute to a better understanding of how deviations from fully rational behavior might be explained. My hope is that future experimental work is able to find evidence about the extent to which this theory's setup and predictions are valid.

Chapter 2

Autonomy Support and Innovation*

> "We don't need bosses. We need servant leaders.
> We need people to serve their teams and let their teams
> come up with the best ideas."
>
> V. Narasimhan, CEO Novartis
> in an interview with Gharib (2019)

2.1 Introduction

The aim of this chapter is to demonstrate optimal investment dynamics in leadership behavior that fosters innovation. Our focus is the intensity in which an innovation encouraging leadership behavior must be displayed over time and not one specific management practice. Management practices matter for an organization's success (Bloom et al., 2013; Bloom and Van Reenen, 2010), and economics has generally focused on beneficial complementarities between different practices (Ichniowski et al., 1995), Ichniowski and Shaw (1999). Rather little research in organizational economics (with only a few exceptions like Hermalin (1998)) focuses on the individual and how she ensures successful implementation of the practices: the leader and her behavior. Non-monetary incentives that enrich the traditional economics' toolkit of bonus schemes, e.g. providing workers with meaning to their job and raising awareness that their impact matters has been shown to increase worker motivation (Cassar and Meier,

* This chapter is based on joint work with Annemarie Gronau (LMU) and was first published in Gronau (2020).

2018; Levitt and Neckermann, 2014), are conveyed through the leader and her behavior towards the workforce. Ichniowski and Shaw (2003) find that in less traditional and productivity-enhancing practices, line workers interact more with supervisors and co-workers. As the effectiveness of management practices depends on their perceived quality and not their quantity (Edgar and Geare, 2005), attributes of leadership behavior that embody them are of particular interest.

Innovation is one important contributor to an organization's long-term competitiveness and success. Monetary incentives work for constrained problems but seem ineffective for open, unconstrained innovation characterized e.g. by a lack of ex-ante specified goals (Charness and Grieco, 2018). In general, management practices stimulate innovative activity (Shipton et al., 2006), but we are unaware of work on non-monetary incentives specifically aiming at encouraging innovative activities of the workforce. We argue that leadership behavior can be such a non-monetary incentive that instills innovative activity.

Workers in a hierarchical work relationship, who are not specifically employed to innovate, refrain from sharing novel ideas if they fear questioning the status quo and resulting negative consequences for themselves. They will speak up if they believe their ideas, and potentially critique, is welcomed and taken seriously by management. The leadership behavior of their supervisor can create a safe space where workers feel free to come up and share novel ideas.

This kind of leadership behavior is captured by the concept of Autonomy Support. Self-Determination Theory is an established construct in Social Psychology, and Autonomy herein is defined as a feeling of volition and freedom that a person experiences e.g. at work (Gagné and Deci, 2005). Autonomy Support is the degree to which the social context enables this feeling, which in our model is the leadership behavior of the principal. Autonomy Support strikes a balance between providing structure and granting freedom of thought, as it creates a safe space for experimentation and failure (Pisano, 2019) which inspires innovation, unleashes creative thinking and encourages workers to communicate their novel ideas.

Kaizen is an example for an Autonomy Supportive leadership behavior, and has received much attention in operations management of the manufacturing sector for some time (see Singh and Singh (2009) for a review). In the Kaizen philosophy, supervisors train workers in methods and Kaizen tools, invite ideas and offer feedback in a non-judgemental way, aiming at developing their workers to share their contributions for improvement and innovation in a safe environment (Quality-One International, 2015). Toyota adopted this innovation encouraging management style with a view to continuous improvement and lean management, which resulted in a strong comparative advantage over a competitor relying on controlling incentives (Helper and Henderson, 2014). Operations management and people management go hand in hand for a company's success (Bloom et al., 2015). We focus on the intensity in which Autonomy Support must be provided, whether it be in operations or people management practices, that encourages heterogeneous workers to thrive in innovative activities without a monetary component.

We formalize Autonomy Support in a principal-agent model as an investment in leadership behavior over time, as single interventions fade out eventually. We derive optimal investment patterns in Autonomy Support that incentivize effort in creative, small scale innovation of the workforce at the bottom of the hierarchy. The patterns are co-determined by the initial Autonomy Support levels of the agent and the rate at which support fades, as well as the benefit-cost ratio and time discount factor of the principal. We find that the principal invests in Autonomy Supportive leadership behavior for almost all parameter constellations. Only if the agent comes with extremely low or high initial levels of Autonomy Support does she refrain from investment. We demonstrate that the principal engages in Autonomy Supportive leadership behavior in accordance with the agent's need to be lifted up as to become active in innovation. Thereby, we contribute to the literature of managing innovation and the literature on economic incentives with a view to the current debate on work and whether monetary rewards are sufficient for performance (Shiller, 2019).

We proceed as follows. First, we show the limited effectiveness of monetary incentives to foster small scale innovation. We then introduce the concept of Autonomy Support review the literature on how it incentivizes innovative activity of the workforce. We then derive the theoretical model to trace optimal investments in Autonomy Support over time. We analyze the resulting investment patterns and derive policy recommendation and future avenues for research. Lastly, we conclude this chapter.

2.2 Limited effectiveness of monetary incentives for small scale innovation

Small scale innovation describes small step improvements of products and processes. It is also typically associated to originate from the part of the workforce who are not specifically employed to innovate, but employees who find creative ways to improve a product or process they are involved in. Small scale innovation can have an important impact on profits. For example, a Walmart worker's idea to use lighter one-step stools for loading trucks will save $ 30 millions in costs from inefficient packing (Thomas, 2019). Customer needs to adapt a product are identified through personal interaction of customer service executives (Bilsland and Cumbers, 2018). While small scale innovation may be of particular interest to small and medium sized companies that cannot maintain an R&D department (Rammer et al., 2009), it is interesting for companies of all sizes to encourage small scale innovation.

An important question then is whether small scale innovation can be incentivized financially. The seminal paper of Manso (2011) displays how an incentive scheme designed to tolerate early failure and reward late successes does indeed encourage innovation, a result that is backed up by a lab experiment in Ederer and Manso (2013). It is important to note that the authors speak of innovation in terms of the contrast between exploitation and exploration. Whether an agent is incentivized to exploit a given business situation or to explore a set of business opportunities: neither qualifies as a creative, unconstrained situation. The recent experimental literature suggests that monetary incentives increase the effort exerted and the quantity and quality of ideas, but do not have an impact on creativity or originality. In fact, they may even reduce creativity (Erat and Gneezy, 2016; Laske and Schröder, 2017). Gibbs et al. (2017) find that with monetary rewards fewer individuals submit more ideas, suggesting a trade-off where people refrain from sharing small and possibly far fetched ideas. The differential impact of monetary incentives on the number of ideas generated and their degree of creativity highlights that the innovation term is used broadly from new combinations of known elements to creative, out-of-the-box approaches. Charness and Grieco (2018) distinguish between constrained and open creativity, and conclude that financial incentives only work for constrained tasks. We adopt this distinction for the remainder of this chapter.

Further, we know that incentive schemes must be carefully designed as to avoid unintended consequences. One major factor is the measurability of effort or output on which the monetary incentive is conditioned. But small scale innovation that improves products and processes is hard to measure, in contrast to the launch of new products and techniques. For example, small process improvements are likely to be tested at some departments and then phased in for the entire organization, making it difficult to disentangle department effects. Continuously changing the production process with small improvements is likely to be disruptive, such that a product is relaunched with a set of small improvements as, for instance, Ikea's 'new' Billy bookshelf. This renders it impossible to trace better performance to one improvement alone. Toyota's productivity increase from Kaizen in the 1980s is attributed to its general ability to foster small scale innovation (Helper and Henderson, 2014), but not to one single improvement.

The operability of financial incentives for small scale innovation is therefore limited in addition to their inability to encourage open and creative ideas. This problem is amplified for the workforce on the ground whose core job is not to contribute ideas. These workers are the ones experiencing the production process first hand, have insight knowledge about the product and work in customer service, which makes them acutely aware of potential problems. But a variety of reasons can restrain them to voice a novel idea: they may fear repercussion for challenging the status quo (Zhou and George, 2003), worry that an idea that does not work out signals inability for the job, or are afraid that a successful idea can render their job useless. In order to unleash the small scale innovation potential of their workforce, companies must acknowledge and address these concerns, a task monetary incentives are unfit for.

2.3 How Autonomy Support incentivizes innovation

Leadership behavior is key for fostering small scale, open innovation of the workforce. Different management practices coalesce in providing a safe space for workers to raise their ideas and concerns, as well as welcoming and encouraging their ideas are characterized by this leadership behavior. This holds particularly true for workers who are not specifically employed for creative, innovative tasks, such as shop floor workers and customer advisers. These typically work within an organizational structure of controlling guidelines, work processes, and deadlines that does not give space for ideation. The pressure they experience prevents creative thinking, and "Numerous attempts at creativity get killed in their infancy because employees fall victim to these emotions" [caused by pressure] (Zhou and George, 2003). The authors make the case that a key determinant for creativity in the workplace is leadership and a supervision style that manages the workers' emotions through empowerment and encouragement. However, providing complete leeway, or no structure at all is unlikely to excite a worker to engage in some innovative activity that actually results in improvements for the company. Pisano (2019) rectifies this misunderstanding of an innovative organizational culture to be just encouraging by emphasizing the balance that management must strike to truly innovative improvements: there must be tolerance for failure, but not for incompetence; there must be a willingness to experiment, but in a highly disciplined way; there must be collaboration, but with individual accountability; there must be flat, but strong leadership: an innovative organizational culture, he writes, must be "psychologically safe but brutally candid".

The leadership behavior in an innovative organizational culture that achieves this balance is one of Autonomy Support. We introduce the concept of Autonomy Support and provide a literature review underlining its impact on innovation.

2.3.1 Autonomy Support

Autonomy Support is a concept from Self-Determination Theory. The underlying concept of Autonomy refers to a feeling of volition and freedom (Gagné and Deci, 2005) when engaging in an activity. Autonomy is thus overlapping but distinct from both its colloquial meaning of independence or the economic concept of intrinsic motivation. The latter two describe a person who chooses independent of others, or chooses what he wants. Autonomy however means that a person experiences a feeling of freedom while doing something even if the person may not have chosen it for himself. This is particularly important for the work environment as a worker has essentially never complete, independent choice of what to do. But he can experience a feeling of Autonomy because he is not controlled at work, or finds value in his work and work environment. It seems perspicuous that a feeling of volition and freedom positively underpins creative thinking, and most people who ever tried to think out-of-the-box under pressure and control (the opposite of Autonomy in Self-Determination Theory) would agree.[2]

> Autonomy Support then is "the degree to which socializing agents take the target individual's perspective; act in ways that encourage choice and self-initiation; provide meaningful rationales and relevance; and refrain from using language or displaying behaviours that are likely to be experienced as pressure toward particular behaviours" (Benita et al., 2014).

Autonomy Support goes beyond one or multiple specific management practices. The overall leadership behavior of the supervisor that permeates management practices however can be Autonomy Supportive. Reflecting the definition of Autonomy Support, Stone et al. (2009) expatiate the following points outlining how practitioners can create autonomous motivation in the long run:

- Asking open questions including inviting participation in solving important problems
- Active listening including acknowledging the employees' perspective
- Offering choices within structure, including the clarification of responsibilities
- Providing sincere, positive feedback that acknowledges initiative, and factual, non-judgmental feedback about problems
- Minimizing coercive controls such as rewards and comparisons with others
- Develop talent and share knowledge to enhance competence and autonomy

Indeed, this constitutes the balance of a successful the management style as described in Pisano (2019). We now turn to the empirical evidence of how such an Autonomy Supportive management style facilitates innovative activity of the workforce in real life.

[2] Assume the example of a Ph.D. student. He may research on whatever he chooses (independence), but that does not necessarily mean that he is motivated, or successful. What does help him succeed is Autonomy Support. Autonomy Support may come in the form of a supervisor who (a) enables and encourages him to take initiative and choose, (b) offers advice, (c) shares his/her perspective and experience when solving problems, and/or (d) offers the "bigger picture".

2.3.2 Literature review: Autonomy Support and innovation

Our introductory example of Toyota's Autonomy Supportive management style Kaizen ticks all boxes of Stone et al. (2009)'s list above. Workers are trained in statistical methods and the structure of the Kaizen process and Kaizen events such that they can make full use of it (Helper and Henderson, 2014; Quality-One International, 2015). They also receive feedback and encouragement both from their supervisor and colleagues (Helper and Henderson, 2014). Besides the continuously displayed leadership behavior, Toyota organized regular Kaizen events structured to invite and harness ideas and provide room to improve them. These events can be designed for broader knowledge gathering or brainstorming in a specific field, but in all cases, it is the task of the management to provide structure and create a safe environment (Early, 2012). Helper and Henderson (2014) attribute Toyota's increased productivity at a time when its rival's GM productivity decreased to the small scale innovation incentivized by this leadership behavior.

Autonomy Support is however not a direct result from an adopted management style alone, but conditional on the behavior of the individual leader, such that differences within a company can arise. Amabile et al. (2004)'s study on the impact of perceived leadership support on creativity showcases a vivid example for Autonomy Support, scrutinizing the widely diverging impact of two 'extreme' micro-managing team leaders in a firm. Both individuals are micromanaging in the sense that they are closely monitoring their team. One uses this management practice only to communicate top level decisions down to them, which puts pressure on the team. The other uses monitoring for immediate exchanges on upcoming challenges, consults the team in decisions and ensures smooth cooperation between team and top management. There is no independence in choice in either team, but in the latter the micro-manager displays Autonomy Support in his leadership behavior. As a consequence, the authors find that this team engages successfully in innovative activities. The controllingly monitored team however is unsuccessful. They further record a positive and a negative spiral in each team respectively. Autonomy Supportive leadership behavior manifests itself over time and has a lasting effect. The leader invests over time in the relationship with its team to constitute a coherent behavior. Besides the time component, this example pinpoints that Autonomy Support encompasses both "instrumental and socioemotional support" (Amabile et al., 2004) and that Autonomy Support and a tight structure are not mutually exclusive. One might argue that because of the structure, the leadership behavior of the Autonomy Supportive team leader was particularly important. This case also demonstrates that leadership behavior is not conditional on one specific management practice.

A leader achieves a structure that creates a safe space such that a feeling of Autonomy and innovative activity can arise by, for example, specifying only "issues to avoid" (Van de Ven, 1986), fostering a feeling of ownership (Dorenbosch et al., 2005) that leads to identification with a leader such that workers follow (Yoshida et al., 2014). Psychological safety promotes creativity, as shown e.g. in part time graduate students (Kark and Carmeli, 2009), and contributes to a firm's financial success through innovation, e.g. in a study of 163 Turkish firms (Akgün et al., 2009). Mumford et al. (2002) conclude that if a leader balances structure and encouragement (being both "cheerleader" as well as "the most demanding critic"), workers "can express their creative capacity". Interestingly, Zhou and George (2001) find that an Autonomy Supportive work environment encourages even dissatisfied workers to be creative. It allows them to use channels to change something about their current situation. At the team level, a leader provides and develops "a safe psychosocial climate and appropriate group processes" (West, 2002) on both the individual and the team level through consultative participation, clarifying objectives and encouraging positive feedback. The team members can then reinforce Autonomy Support among one another (Gagné and Deci, 2005). The team leader's role is "orchestrating" these efforts (Mumford et al., 2002), and like a conductor, becoming part of the group and ensuring that the members work harmoniously together. In the health care sector, for example, this has been shown to increase quality of work and innovation (Borrill et al., 2000).

An Autonomy Supportive leadership is beneficial also for workers who have a personal tendency for creativity. Indeed, it falls short to assume that personal creativity alone achieves innovative activity (Mumford et al., 2002). Rather, personal creativity and an appropriate work context complement each other in accomplishing this goal (Janssen et al., 2004). In the absence of non-controlling and supportive supervision creativity and patents actually decrease (Oldham and Cummings, 1996). Creativity of workers unsure of their capabilities is unleashed when management builds their confidence and serves as a role model (Tierney and Farmer, 2002), pointing towards the importance of Autonomy Support for blue collar workers. More complex jobs are assumed to spur worker's interest and creativity. Even for those Shalley et al. (2009) show that Autonomy Support has a positive impact on creativity regardless of whether workers have a high or low degree of Autonomy or work in complex or less complex jobs. For low Autonomy and less complex jobs the effect is just more pronounced.

As with other incentives, Autonomy Support must be properly designed and applied to ensure the desired effect and prevent unintended consequences. It is insufficient to give workers pro forma choice as to create an illusion of Autonomy. Experiments indicate that when workers' decisions are not taken seriously it discourages effort: when workers are delegated to choose which project to implement, but the leader overrules their decision, effort levels and transfers (Sloof and von Siemens, 2019; Corgnet and Hernán González, 2013) plummet. A successful manager must be able to change one's mind when involving workers in the decision process (Corgnet and Hernán González, 2013). Leadership behavior that is perceived as insincere or intrusive has a negative impact on innovative behavior (Bammens, 2016). Autonomy Support must also specifically aim at encouraging innovation. Ohly et al. (2006) find that while Autonomy Support inspires personal initiative, it does not increase creativity and innovation, and ascribe this finding to the fact that supervisor support was not clearly targeted at innovation. The company's system of processing suggestions might have been misused to communicate complaints.

This review highlights that Autonomy Support encompasses a variety of actions and behaviors of a leader, which can be displayed in a different operations and human management practices. Autonomy Supportive leadership behavior must be authentic and offer both instrumental and emotional support. A one time Autonomy Supportive intervention is unlikely to have a believable lasting effect, and fades over time. Therefore, it must be reinforced and renewed. Regular efforts to provide Autonomy Support are necessary, and different behaviors or intensities may be required at different points in time. While companies like Toyota incorporate a management style that targets at Autonomy Support, it eventually comes down to the quality of the leadership behavior of the principal in question. We formalize these notions in our model.

2.4 A model of investments in Autonomy Support

2.4.1 Model setup

We investigate optimal investments in providing Autonomy Support of a principal with a view to incentivize innovative activity of the worker. We expect optimal investments to vary over time, as leadership behavior and actions required differ. For example, extensive methods training or occasional feedback require behavior of different time intensity. We choose a two-period model to incorporate this notion in a principal-agent model. It allows us to keep the model as tractable as possible whilst capturing the dynamic aspect of Autonomy Support investments.

For the purpose of the research question, we concentrate only on innovative activity i of the agent as inventivized through Autonomy Support by the principal. When the agent (he) engages in innovative activity i the probability for his efforts to result in a successful innovation increases according to the probability function

$$Pr(i) = \frac{i}{1+i}$$

with $Pr(i) \in [0,1)$ for $i \geq 0$.

The agent experiences effort costs from innovative activity, which decreases in the amount of Autonomy Support available to the agent. Prior discounted and current Autonomy Support constitute this available amount. The discount factor δ accounts for the fact that Autonomy Support investments fade out over time. A one period intervention does not carry indefinitely into future periods with the same motivational power. We factor in the Autonomy Support provided previously, either in employment or personal relationships, by assuming a personal consolidated start level $\bar{s}$ of an agent. Alternatively, the start value can be interpreted as the agent's personal autonomous motivation level for innovation prior to employment (Shalley et al., 2009). Different agents therefore have different initial levels of $\bar{s}$. The total value of

Autonomy support $\bar{s}_1$ and $\bar{s}_2$ available to an agent in periods $t = 1, 2$ is therefore given by

$$\bar{s}_1 = \delta\bar{s} + s_{a,1}$$
$$\bar{s}_2 = \delta^2\bar{s} + \delta s_{a,1} + s_{a,2}$$

denoting the sum of the discounted start value $\bar{s}$ and the principal's investments in the prior, discounted, and the current period.

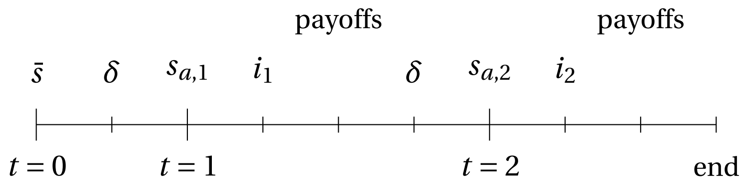
The payoff the agent experiences from successful innovation is denoted as v_A and normalized to one. We understand this payoff to be the utility derived from being creatively active and seeing one's innovative activity come to fruition, not as a monetary reward. Taken together, the agent's utility as a function of his innovative activity i for periods $t = 1, 2$ are

$$U_A(i,1) = v_A\frac{i}{1+i} - \frac{i}{\delta\bar{s} + s_{a,1}} = v_A\frac{i}{1+i} - \frac{i}{\bar{s}_1} \tag{2.1}$$
$$U_A(i,2) = v_A\frac{i}{1+i} - \frac{i}{\delta^2\bar{s} + \delta s_{a,1} + s_{a,2}} = v_A\frac{i}{1+i} - \frac{i}{\bar{s}_2} \tag{2.2}$$

The principal (she) receives a payoff v_P when the agent's innovation is successful, and bears costs $c(s_{a,t}) = \alpha s_{a,t}$ from providing Autonomy Support at time t. α denotes the principal's marginal cost. Alternatively, it can be interpreted as her ability to provide Autonomy Support. The principal is forward looking and accounts for the fact that Autonomy Support provided in period 1 impacts innovative activity and thus innovation profit in both periods. The principal discounts her profit in period 2 with β and includes it in her first period considerations. The profit functions for periods $t = 1, 2$ are

$$\Pi_P(s_{a,1}, 1) = v_P\frac{i}{1+i} - \alpha s_{a,1} + \beta\Pi_P(s_{a,2}, 2) \tag{2.3}$$
$$\Pi_P(s_{a,2}, 2) = v_P\frac{i}{1+i} - \alpha s_{a,2} \tag{2.4}$$

We assume that the principal can perfectly observe the initial amount of Autonomy Support $\bar{s}$ the agent enters the work relationship with. We look at situations in which the benefits from Autonomy Support and its associated costs are such that the principal considers investment. We achieve this by assuming that the benefit-cost ratio satisfies $\frac{v_P}{\alpha} > 2$. The time line of the model is as follows. In each period, the agent's total Autonomy Support is depreciated before the principal has the opportunity to invest anew. Based on the currently available Autonomy Support level, the agent chooses his innovation effort, and the payoffs of the period are realized.



The principal seeks to maximize her profits from innovation by optimally investing in the agent's Autonomy Support. We solve the maximization problem via Backward Induction because profits are time-interdependent through the investment choices.

2.4.2 Solving the model

Solving for period $t = 2$

While the principal takes future effects of her own current investment into account, the agent is backward looking. He considers each period separately and only takes past investments in his available Autonomy Support into account. In period $t = 2$, he maximizes $U_A(i, 2)$ from Equation (2.2) by his choice of innovative activity i, resulting in

$$i_2^* = (\delta^2 \bar{s} + \delta s_{a,1} + s_{a,2})^{\frac{1}{2}} - 1 \tag{2.5}$$

The agent's innovative effort i_2^* depends on his initial Autonomy Support level, the discounted previous and the current Autonomy Support investment by the principal. The agent only becomes active if there is overall enough Autonomy Support. If previous Autonomy Support is low or heavily discounted such that $\delta^2 \bar{s} + \delta s_{a,1} = 0$, no innovative activity takes place if the principal does not sufficiently invest in the current period $t = 2$. If the previous Autonomy Support is high and/or is only mildly discounted, innovative effort is exerted even if the principal does not invest in the current period at all, with $s_{a,2} = 0$.

The principal maximizes $\Pi_P(s_{a,2},2)$ from Equation (2.4) in order to derive her optimal investment $s_{a,2}$, which results in

$$s^*_{a,2} = \left(\frac{v_P}{2\alpha}\right)^{\frac{2}{3}} - \delta^2\bar{s} - \delta s_{a,1} \geq 0 \tag{2.6}$$

$s^*_{a,2}$ increases in profit v_P from successful innovation and decreases in the marginal cost parameter α for providing Autonomy Support. The benefit-cost ratio $\frac{v_P}{\alpha}$ also comes into play when determining if the principal invests at all in the second period. Only if the agent's initial Autonomy Support level and previous investment is low or heavily discounted such that it does not exceed the benefit-cost ratio $\frac{v_P}{\alpha}$, does the principal choose a positive investment $s_{a,2} > 0$. Otherwise, the principal does not need to replenish the stock of Autonomy Support and chooses $s_{a,2} = 0$.

The principal's profit in period $t = 2$ when she engages in current Autonomy Support investments $s^*_{a,2} > 0$ is

$$\begin{aligned}\Pi^*_{P,2}(s^*_{a,2}|s^*_{a,2} > 0) &= v_P\frac{\left(\frac{v_P}{2\alpha}\right)^{\frac{1}{3}} - 1}{\left(\frac{v_P}{2\alpha}\right)^{\frac{1}{3}}} - \alpha\left(\frac{v_P}{2\alpha}\right)^{\frac{2}{3}} + \alpha(\delta^2\bar{s} + \delta s_{a,1}) \\ &= v_P - 3\alpha\left(\frac{v_P}{2\alpha}\right)^{\frac{2}{3}} + \alpha(\delta^2\bar{s} + \delta s_{a,1})\end{aligned}$$

and for her optimal choice of $s^*_{a,2} = 0$

$$\Pi^*_{P,2}(s^*_{a,2}|s^*_{a,2} = 0) = v_P - \frac{v_P}{(\delta^2\bar{s} + \delta s_{a,1})^{\frac{1}{2}}} > 0$$

Solving for period $t = 1$

The agent maximizes his utility $U_A(i,1)$ of period $t = 1$ in Equation (2.1) by choosing

$$i^*_1 = (\delta\bar{s} + s_{a,1})^{\frac{1}{2}} - 1 \geq 0$$

The agent's innovation effort in period $t = 1$ increases in both his initial Autonomy Support level and the principal's investment in the current period.

When maximizing $\Pi_P(s_{a,1},1)$ in Equation (2.3), the principal takes this and the discounted future impact of her investments in period $t=1$ for the subsequent period into account. Her optimal investment choice is

$$s^*_{a,1} = \left(\frac{v_P}{2\alpha(1-\beta\delta)}\right)^{\frac{2}{3}} - \delta\bar{s} \geq 0 \tag{2.7}$$

Her investment increases when the benefit-cost ratio increases and increases in how strongly she values future periods, as described by β. The effect of the agent's discount parameter δ is ambiguous and will be part of the discussion on the different investment patterns. Contingent on the parameter constellation, the principal may or may not invest in Autonomy Support in period $t=1$.

The resulting profit in period $t=1$, if she invests in Autonomy Support in both periods such that $s^*_{a,1}>0$ and $s^*_{a,2}>0$, is

$$\begin{aligned}\Pi^*_{P,1}(s^*_{a,1}>0) = v_P - \frac{1}{\frac{v_p}{(2\alpha(1-\beta\delta)}^{\frac{2}{3}}} - \alpha\left(\frac{v_p}{(2\alpha(1-\beta\delta)}\right)^{\frac{2}{3}} + \alpha\delta\bar{s} \\ + \beta\left[v_P - 3\alpha\left(\frac{v_P}{2\alpha}\right)^{\frac{2}{3}}\right] + \beta\delta\left(\frac{v_p}{(2\alpha(1-\beta\delta)}\right)^{\frac{2}{3}}\end{aligned} \tag{2.8}$$

and for $s^*_{a,1}=0$

$$\Pi^*_1(s^*_{a,1}=0) = v_P - \left(\frac{1}{\delta\bar{s}}\right)^{\frac{1}{2}} + \beta v_P\left(\frac{1}{\delta^2\bar{s}}\right)^{\frac{1}{2}} \tag{2.9}$$

2.4.3 Results: Autonomy Support investment patterns

We derive the investment patterns that emerge depending on the parameter constellations of the benefit-cost ratio $\frac{v_P}{\alpha}$, the agent's depreciation rate of Autonomy Support δ, and the principal's discount factor β of future profits.[3]

The first broad distinction for the different patterns is the relationship between the agent's and principal's future benefits from investing in Autonomy Support in the current period. The

[3] The specific derivations for each case are stated in the Mathematical Appendices B.1 and B.2.

condition core to this distinction is given by

$$\frac{1}{\delta} - \sqrt{\delta} \lesseqgtr \beta \tag{2.10}$$

as depicted in Figure 2.4.1, where $\frac{1}{\delta} - \sqrt{\delta} = \beta$ is delineated.

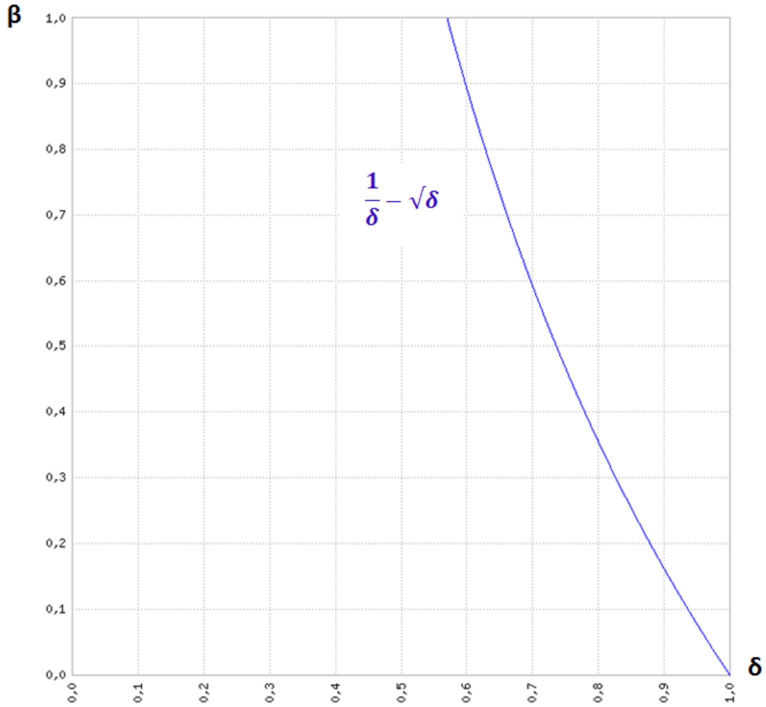


Figure 2.4.1: β- δ relationship

To the left of this line, $\frac{1}{\delta} - \sqrt{\delta} > \beta$ holds when either δ or β, or both simultaneously are low. The agent discounts Autonomy Support at a high rate, and investments fade out strongly. The principal does not value future payoffs highly. Intuitively, this translates to lower incentives to invest in Autonomy Support and results in lower investments.

To the right of this line, $\frac{1}{\delta} - \sqrt{\delta} < \beta$ holds when β and δ are simultaneously high. This means that the agent discounts Autonomy Support at a low rate, and investments last. The principal values future payoffs strongly. Investing in the current period is beneficial for both. Intuitively, this translates to higher incentives to invest Autonomy Support and results in higher investments in the given period.

In accordance with this distinction, we now describe the specific investments for low and for high future benefit investment patterns.

Low future benefit investment patterns under $\frac{1}{\delta} - \sqrt{\delta} > \beta$

Pattern I: Investment in each period

The principal chooses a positive investment in period $t = 1$, $s^*_{a,1} > 0$, if the initial Autonomy Support amount $\bar{s}^i$ does not exceed (derived from Equation (2.7)) a threshold of

$$\frac{1}{\delta}\left(\frac{v_P}{\alpha}\frac{1}{2(1-\beta\delta)}\right)^{\frac{2}{3}} > \bar{s}^i \tag{2.11}$$

This means that the benefit-cost ratio $\frac{v_P}{\alpha}$ of innovation is high enough to interest the principal in encouraging innovative activity. However, the initial Autonomy Support amount is so low that she chooses to invest in the first period. After the depreciation of $s^*_{a,1}$ (in Equation (2.7)) the principal further invests $s^*_{a,2} > 0$ in period $t = 2$, if Equation (2.6) satisfies

$$\begin{aligned} s^*_{a,2} &= \left(\frac{1}{2}\right)^{\frac{2}{3}}\left(\frac{v_P}{\alpha}\right)^{\frac{2}{3}} - \delta^2\bar{s} - \delta s^*_{a,1} \\ &= \left(\frac{1}{2}\right)^{\frac{2}{3}}\left(\frac{v_P}{\alpha}\right)^{\frac{2}{3}}\left(1-\delta\left(\frac{1}{1-\beta\delta}\right)^{\frac{2}{3}}\right) > 0 \end{aligned}$$

which is always true for $\frac{1}{\delta} - \sqrt{\delta} > \beta$ as assumed for this section.
Rearranging the time-interdependent profit function of period $t = 1$ (Equation (2.8)) yields a lower bound of $\bar{s}^i$ that ensures positive expected profits from innovative activity for the principal:

$$\bar{s}^i \geq \frac{v_P}{\alpha}\frac{1}{\delta}\left[3\left(\frac{1}{2}\right)^{\frac{2}{3}}\left(\frac{v_P}{\alpha}\right)^{-\frac{1}{3}} - 1\right]\left[\frac{1+\frac{\beta\delta}{\alpha}}{(1-\beta\delta)^{\frac{2}{3}}} + \beta\right] \equiv X_{\Pi_{P1,1}} \geq 0 \tag{2.12}$$

The initial amount of Autonomy Support $\bar{s}^i$ must exceed this threshold, otherwise the principal does not find it optimal to add to it in the first and subsequent second period. Intuitively, there must be some, but not too much Autonomy Support of the agent to build on such that the principal, although inclined to invest little, finds it optimal to invest in both periods. In Figure 2.4.2, we see that depending on the initial Autonomy Support, the principal matches her investment such that the agent achieves the $\bar{s}_1$ necessary to optimally exert effort in innovative activity. The higher the initial level, the less must the principal touch it up. The emerging patterns are steadily decreasing (Pattern Ia in Figure 2.4.2) or hump-shaped (Pattern Ib). The agent's utilities in the respective periods are

$$U_A(i^*,1) = 1 - \frac{2}{\left(\frac{v_P}{2\alpha(1-\beta\delta)}\right)^{\frac{1}{3}}} + \frac{1}{\left(\frac{v_P}{2\alpha(1-\beta\delta)}\right)^{\frac{2}{3}}}$$
$$U_A(i^*,2) = 1 - \frac{2}{\left(\frac{v_P}{2\alpha}\right)^{\frac{1}{3}}} + \frac{1}{\left(2\frac{v_P}{2\alpha}\right)^{\frac{2}{3}}}$$

Pattern II: Investment only in second period

If the agent enters the company with an initial amount of Autonomy Support $\bar{s}^{ii}$ that exceeds the threshold in Equation (2.11), the principal does not invest in period $t = 1$ in the $\frac{1}{\delta} - \sqrt{\delta} > \beta$ environment. However, she does invest in $t = 2$ as per Equation (2.6) if

$$\frac{1}{\delta^2}\left(\frac{1}{2}\right)^{\frac{2}{3}}\left(\frac{v_P}{\alpha}\right)^{\frac{2}{3}} > \bar{s}^{ii} \tag{2.13}$$

or, phrased differently, if the initial Autonomy Support amount is sufficient to encourage innovative activity in the first period, but not in the second period.
The agent's utilities become

$$U_A(i^*,1) = 1 - \frac{2}{(\delta\bar{s})^{\frac{1}{3}}} + \frac{1}{(\delta\bar{s})^{\frac{2}{3}}}$$
$$U_A(i^*,2) = 1 - \frac{2}{\left(\frac{v_P}{2\alpha}\right)^{\frac{1}{3}}} + \frac{1}{\left(2\frac{v_P}{2\alpha}\right)^{\frac{2}{3}}}$$

Pattern III: No investment

If the agent's initial amount of Autonomy Support $\bar{s}^{iii}$ exceeds the threshold in Equation (2.11), such that

$$\bar{s}^{iii} \geq \frac{1}{\delta^2}\left(\frac{1}{2}\right)^{\frac{2}{3}}\left(\frac{v_P}{\alpha}\right)^{\frac{2}{3}} \tag{2.14}$$

it is sufficiently high to encourage innovative activity even if the principal does not invest in either period. The agent's utilities are

$$U_A(i^*,1) = 1 - \frac{2}{(\delta\bar{s})^{\frac{1}{3}}} + \frac{1}{(\delta\bar{s})^{\frac{2}{3}}}$$
$$U_A(i^*,2) = 1 - \frac{2}{\left(\delta^2\bar{s}\right)^{\frac{1}{3}}} + \frac{1}{\left(\delta^2\bar{s}\right)^{\frac{2}{3}}}$$

In contrast, if the initial level is so low that it falls short of the threshold $X_{\Pi_{P1,1}}$ in Equation (2.12), then the principal optimally chooses not to invest in either period and no innovation takes place. The agent's utilities then are

$$U_A(i^*,1) = 0$$
$$U_A(i^*,2) = 0$$

No investments for a high and a low initial amounts and the resulting development of Autonomy Support levels are depicted in Pattern IIIa and Pattern IIIb in Figure 2.4.2.

Pattern	Thresholds					Investment
IIIb			$\bar{s}$	$<$	$X_{\Pi_{P1,1}}$	no
I	$X_{\Pi_{P1,1}}$	$\leq$	$\bar{s}^i$	$<$	$\frac{1}{\delta}\left(\frac{v_P}{\alpha}\frac{1}{2(1-\beta\delta)}\right)^{\frac{2}{3}}$	$t = 1,2$
II	$\frac{1}{\delta}\left(\frac{v_P}{\alpha}\frac{1}{2(1-\beta\delta)}\right)^{\frac{2}{3}}$	$\leq$	$\bar{s}^{ii}$	$<$	$\frac{1}{\delta^2}\left(\frac{1}{2}\right)^{\frac{2}{3}}\left(\frac{v_P}{\alpha}\right)^{\frac{2}{3}}$	$t = 2$
IIIa	$\frac{1}{\delta^2}\left(\frac{1}{2}\right)^{\frac{2}{3}}\left(\frac{v_P}{\alpha}\right)^{\frac{2}{3}}$	$\leq$	$\bar{s}^{iii}$			no

Table 2.4.1: Thresholds for low future benefit investment patterns

Table 2.4.1 displays an overview of the thresholds for all low future investment patterns. For given parameter values the principal prefers investing only in the second period when the agent's initial level of Autonomy Support is high enough.

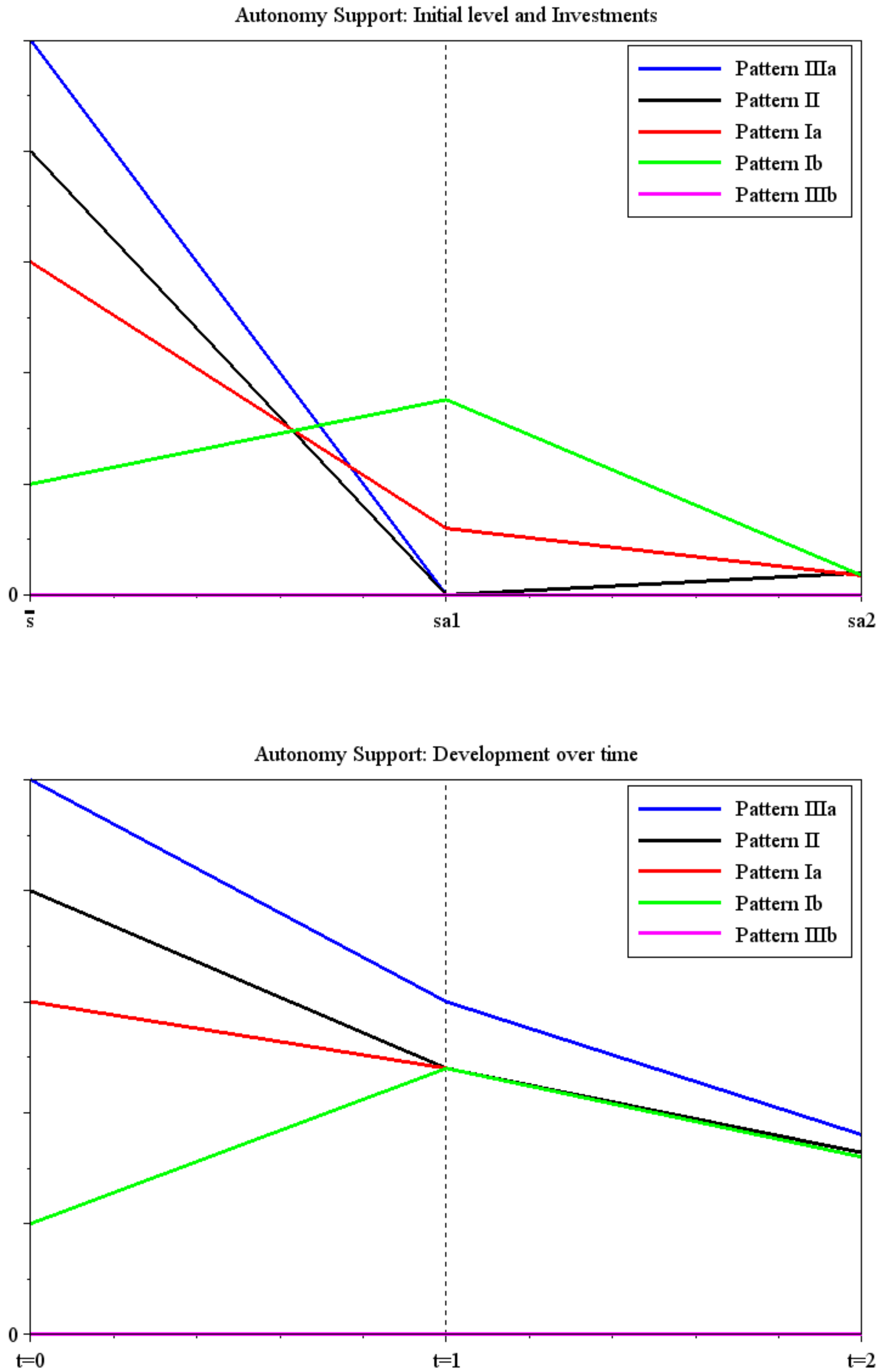


Figure 2.4.2: Development of Autonomy Support for low future benefit investment patterns

High future benefit investment patterns under $\frac{1}{\delta} - \sqrt{\delta} \le \beta$

Pattern IV: Investment only in first period

With the relationship $\frac{1}{\delta} - \sqrt{\delta} \le \beta$, it holds that

$$\begin{aligned} s^*_{a,2} &= \left(\frac{1}{2}\right)^{\frac{2}{3}} \left(\frac{v_P}{\alpha}\right)^{\frac{2}{3}} - \delta^2 \bar{s} - \delta s^*_{a,1} \\ &= \left(\frac{1}{2}\right)^{\frac{2}{3}} \left(\frac{v_P}{\alpha}\right)^{\frac{2}{3}} \left(1 - \delta \left(\frac{1}{1-\beta\delta}\right)^{\frac{2}{3}}\right) \le 0 \end{aligned}$$

and the principal always optimally chooses $s^*_{a,2} = 0$.
Further, the principal invests in Autonomy Support in period $t = 1$ only if

$$\frac{1}{\delta}\left[\frac{v_P}{2\alpha}\left(1 + \frac{\beta}{\sqrt{\delta}}\right)\right]^{\frac{2}{3}} > \bar{s}^{iv} \tag{2.15}$$

however, this equation takes a different value than in Pattern I.
Inserting $s^*_{a,1}, s^*_{a,2}$ in the time interdependent profit function of period $t = 1$, we can derive a lower bound $X_{\Pi_{P1,2}}$ that ensures that the principal invests in $t = 1$:

$$\bar{s}^{iv} \ge \frac{1}{\delta}\frac{v_P}{\alpha}\left[\left(\frac{1}{2(1-\beta\delta)}^{\frac{2}{3}}\left(\frac{v_P}{\alpha}\right)^{-\frac{1}{3}}\right)(2(1+\frac{\beta}{\sqrt{\delta}})^2 + 1) - (1+\beta)\right] \equiv X_{\Pi_{P1,2}} \tag{2.16}$$

The resulting utilities for the agent are

$$\begin{aligned} U_A(i^*, 1) &= 1 - \frac{2}{\left(\left[\frac{v_P}{2\alpha}\left(1+\frac{\beta}{\sqrt{\delta}}\right)\right]\right)^{\frac{1}{3}}} + \frac{1}{\left(\left[\frac{v_P}{2\alpha}\left(1+\frac{\beta}{\sqrt{\delta}}\right)\right]\right)^{\frac{2}{3}}} \\ U_A(i^*, 2) &= 1 - \frac{2}{\left(\left[\delta\frac{v_P}{2\alpha}\left(1+\frac{\beta}{\sqrt{\delta}}\right)\right]\right)^{\frac{1}{3}}} + \frac{1}{\left(\left[\delta\frac{v_P}{2\alpha}\left(1+\frac{\beta}{\sqrt{\delta}}\right)\right]\right)^{\frac{2}{3}}} \end{aligned}$$

The principal matches his investment to the agent's initial level in the first period such that he exerts optimal innovative effort in both periods. With varying initial levels, the pattern for development over time can take the shape of Pattern IVa or Pattern IVb, steadily decreasing or hump-shaped, as depicted in Figure 2.4.3.

Pattern V: No investment

If the initial Autonomy Support exceeds the threshold in Equation (2.15)

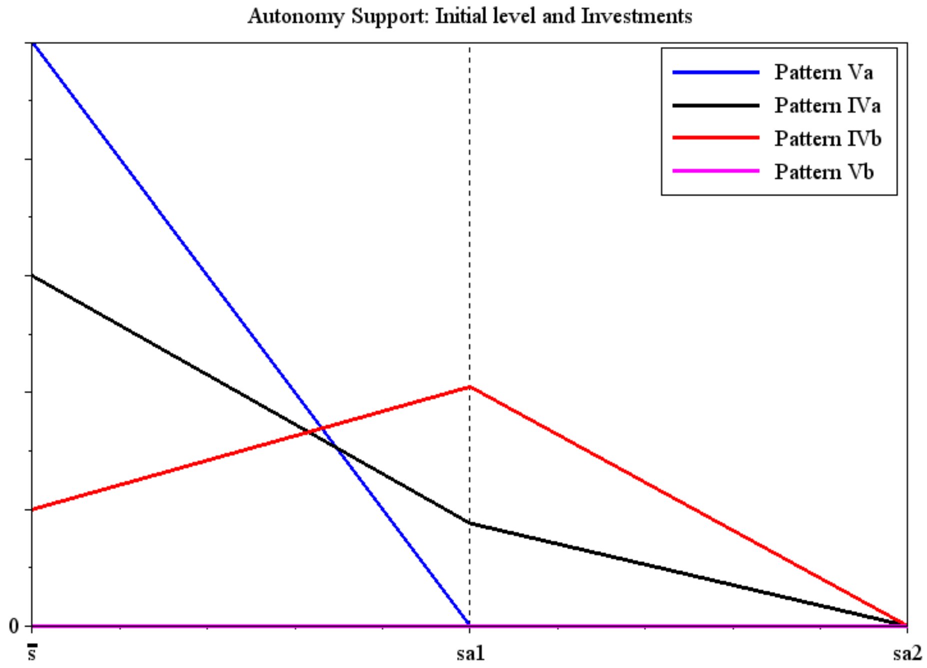
$$\bar{s}^{v} \geq \frac{1}{\delta}\left[\left(\frac{v_P}{2\alpha}\right)\left(1+\frac{\beta}{\sqrt{\delta}}\right)\right]^{\frac{2}{3}} \tag{2.17}$$

then it sufficiently big to encourage innovative activity in both periods without investments by the principal. In contrast, if the initial amount of Autonomy Support is so small that it falls short of the threshold in Equation (2.15), the principal does not find it worthwhile to invest in Autonomy Support at all. For high or low initial levels, the investments are depicted in Patterns Va and Vb in Figure 2.4.3, respectively.
The resulting utilities are as in Pattern III, albeit with parameters satisfying $\frac{1}{\delta} - \sqrt{\delta} \leq \beta$.

Table 2.4.2 displays an overview of the thresholds for high future benefit investment patterns. For given parameter values the principal prefers investing only in the first period when the agent's initial level of Autonomy Support is high enough.

Pattern	Thresholds				Investment
Vb			$\bar{s}$	$<$ $X_{\Pi_{P1,2}}$	no
IV	$X_{\Pi_{P1,2}}$	$\leq$	$\bar{s}^{iv}$	$<$ $\frac{1}{\delta}\left[\frac{v_P}{2\alpha}\left(1+\frac{\beta}{\sqrt{\delta}}\right)\right]^{\frac{2}{3}}$	$t=1$
Va	$\frac{1}{\delta}\left[\frac{v_P}{2\alpha}\left(1+\frac{\beta}{\sqrt{\delta}}\right)\right]^{\frac{2}{3}}$	$\leq$	$\bar{s}^{v}$		no

Table 2.4.2: Thresholds for high future benefit investment patterns



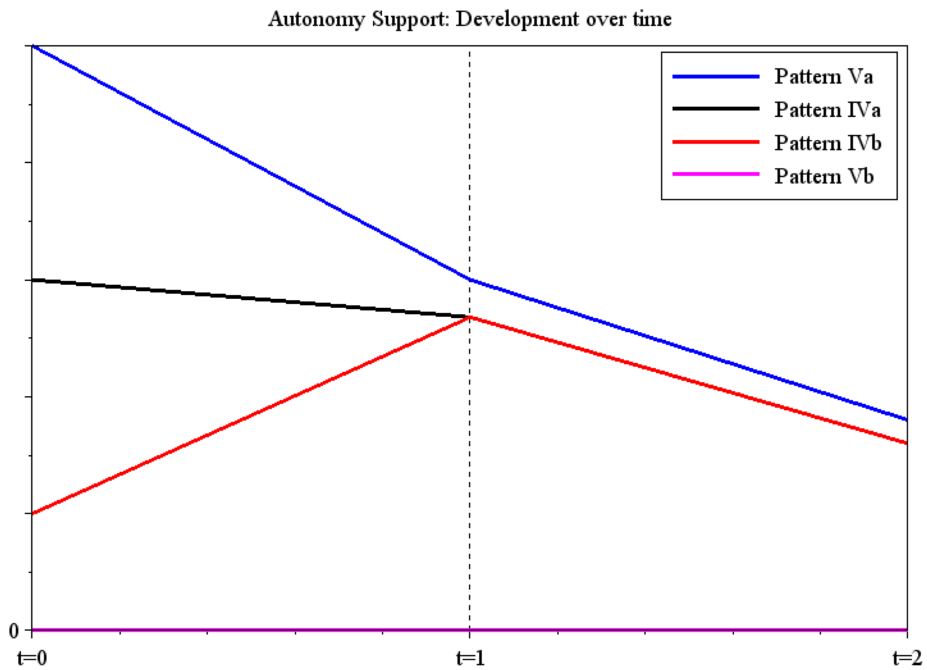


Figure 2.4.3: Development of Autonomy Support for high future benefit investment patterns

2.5 Discussion and implications

More attractive investments in Autonomy Support, let that be because of higher payoff v_P or lower cost α, naturally result in higher investments, and potentially investments in more time periods, as can be seen in Tables 2.4.1 and 2.4.2.
For both β-δ relationships and a given benefit-cost ratio $\frac{v_P}{\alpha}$, we find a certain substitutability[4] between $\bar{s}$ and the principal's investments, following the literature that investments additively impact a worker's creative inclination (Hagger et al., 2015). The more the agent feels encouraged to be active in innovation at the start the work relationship, the less must the principal invest in order to maintain that motivation. Substitutability does not infer, however, that for individuals with high initial levels Autonomy Support investments are futile. Even for rather high levels of initial Autonomy Support we find positive investments in at least one period by the principal. This reflects the finding that it cannot be taken for granted that workers creative at the beginning of a work relationship remain so (Mumford et al., 2002).
In those cases where the agent's initial level and the level required to achieve the optimal innovation effort choice do not diverge too strongly, the principal matches their difference with her investment as depicted in Figures 2.4.2 and 2.4.3. The difference is co-determined by the rate at which the initial level and further investments are depreciated as well as the valuation of the principal for future periods. Our model hence captures the notion that individuals have different Autonomy Support requirements to succeed at different points in time, and that appropriate intensity of Autonomy Support investments position them to be successful (Shalley et al., 2009).
For those situations where the principal only invests in one of the two periods, the β-δ relationship determines in which period the investment is done. When the principal's valuation of the next period exceeds the depreciation term ($\frac{1}{\delta} - \sqrt{\delta} \leq \beta$), she invests 'in advance' in the first period such that the Autonomy Support given carries over to the second period and still encourages innovative activity. This reflects a result in the innovation management literature that Autonomy Support in the form of encouragement is particularly important at early stages of an innovative project, while Autonomy Support in the form of partaking in the deci-

[4] The model with complementarity between the initial level and investments is derived in the Mathematical Appendix B.3. We show that complementarity only eliminates the u-shaped Pattern II and no investment Pattern V for high initial levels of Autonomy Support. Else, our qualitative results remain unchanged.

sion process and control over the implementation becomes more important in later periods (Axtell et al., 2006). It seems plausible to assume that the latter is less time-intensive for the principal and can be described by her lower investments in Autonomy Support at the later period. This is also in accordance with a situation where the principal has a high valuation of future periods: if she manages the entire project from the first creative idea up to the implementation, she likely is forward looking and rather invests in advance instead of supporting her worker only at later stages of the project.

We observe the possibility of a u-shape of Autonomy Support investments in Pattern IIb when either or both of β and δ are low ($\frac{1}{\delta} - \sqrt{\delta} \geq \beta$). Although relying on the agent's initial level in the first period, the principal touches up the depreciated Autonomy Support in the second period. This result is surprising: intuitively, one expects that the principal refrains from investments in the last period of the work relationship. The expected profit from innovation in the second period is enough for her to make that investment nonetheless. Considering the notion that the principal only provides Autonomy Support as required to achieve innovative activity, the principal can optimally decide to let initial levels deteriorate and invest only in the last period. As this result is possible even when the principal has all information and acts rationally, it stands to reason that this case gains importance when the principal is not fully informed. For example, a principal in a company that has not done Autonomy Support before may not be fully aware of the potential benefits v_P from innovative activity or the actual costs α of providing support, or how quickly Autonomy Support discounts at rate δ. In the course of managing her workforce, she learns about these dimensions and adjusts her optimal investment decision accordingly.

We find that the principal optimally chooses not to invest if the agent's initial Autonomy Support level exceeds a high threshold (Equations 2.14, 2.17), that allows innovative activity without any investments. These thresholds are key: the principal does not choose to invest because she does not see worth in Autonomy Support, but for a given benefit-cost ratio $\frac{v_P}{\alpha}$ that co-determines the thresholds it is not optimal for her to do so. This adds nuance to the notion that some companies rely on agent's creativity because they do not want to invest (Mumford et al., 2002). Instead, they may not find it optimal, and workers still exert innovation effort. On the reverse, the benefit-cost ratio may not allow the principal to invest and build up Autonomy Support in agents that arrive with a very low initial level.

The intuition of the investment dynamics of our two-period model can easily be transferred to a multiple period model. Investments then carry over to more than one subsequent period and positively influence the principal's profit potentially longer. As in our two-period model, the principal only aims for a lower level of innovative activity i in the last period. With more periods, her investments in earlier periods are higher and fade out towards the end of the relationship. The distinction in our results with respect to the β-δ relationship becomes less important. In either case, we expect a wave pattern of investments. In the low investment case, the principal restocks the Autonomy Support depreciated from the previous period, and in the high investment case, the principal invests in advance and lets it depreciate in the next period. But this only matters for the initial investments. Afterwards, the principal ensures the optimal long term innovative effort by maintaining it with her investments. This means that at some point, the principal replenishes the depreciated Autonomy Support even of an agent with a very high initial level. In the intermediate periods of the model, we expect rather stable expected profits and utilities from innovative activity. The picture is less clear for agents who enter the company with such a low level of Autonomy Support that in our version, the principal refrains from investment. With more periods, the threshold for not investing would be lower, as the principal would forego more profits over time. This renders it potentially interesting to build such a worker's Autonomy Support "from scratch". Considering real life employment relationships and life cycles, this seems plausible. Only under the conceivably worst starting condition do investments in Autonomy Support not positively impact a person's activity and unleash creative potential. The notion that growth is inherently possible, and desirable, for everyone is perfectly in line with Self-Determination Theory.

Thus far, the depreciation rate is assumed to be agent-specific. However, we can also understand δ to be an environmental variable. Supervisors' choice of providing Autonomy Support has a strong impact on workers' innovative efforts that can diverge even within a company (Amabile et al., 2004), and Autonomy Support is relationship-specific between supervisor and worker. If the company has an organizational structure prone to disrupting this relationship, e.g. a tendency for unexpected job rotation or restructuring, at least part of the Autonomy Support is lost because it is not necessarily attributed the company, but the specific supervisor. The company then influences innovative activity not only through investments via the supervisor, but also whether it allows these investments to last. This suggests that δ as an environmental variable is correlated with investments in Autonomy Support: the more easily the organizational structure disrupts supervisor-worker relationships, the more important become supervisors' investments. This interpretation can be incorporated in the model by allowing investments in Autonomy support to also influence a period-specific depreciation rate.

Our model further indicates that agents with varying start values of Autonomy Support derive similar utilities from being innovatively active. Only those with start values so high that they prevent further investments thrive above, while those with start values so low that no investments are made have no utility. This underlines our understanding that people value being allowed to be creative, but also adds to the bigger notion that they value being autonomous: not only appreciating results, but also the processes that lead to results (Benz and Frey, 2008).

Transferring our findings to real world scenarios indicates that for longer relationships, reflected in a multiple period scenario, Autonomy Support is more likely provided continuously. The first investment, as in the first period of our model, depends on the β-δ relationship. When they are high, such that both parties expect high future benefits, the supervisor engages in sizable Autonomy Support investments right from the start. When one or both are low, the principal anticipates low future benefits, such that the first investment is rather small. As such, high Autonomy Support investments can be understood as part of an on-boarding process where both parties expect the work relationship to last. We believe that screening for Autonomy Support in the recruitment process occurs even in the absence of specific measurements. Previous Autonomy Support is reflected in previous behaviors and choices of the agent, which may at least be in part observable, e.g. in the CV or recommendation letters. In an ongoing work relationship, the current need for Autonomy Support investments may be detected by employing questionnaires for perceived Autonomy Support (Hagger et al., 2007; Mageau et al., 2015) such that the principal can react to it. Even creative workers are in need of an Autonomy Supportive leadership behavior (Mumford et al., 2002) to preserve their efforts.

Regarding the benefit-cost ratio, companies may be concerned about costs in training team leaders in Autonomy Support[5] as to introduce Autonomy Support. We would expect trained team leaders to then actually have a lower marginal cost α. Even then, assessing the potential benefit from small scale innovation υ_P is hard, especially when a company has no prior

[5] Fixed costs do not change the qualitative results of our model, but slightly shift the thresholds.

experience in innovating. We believe this may be the most constraining factor in providing Autonomy Support. In the example of Toyota and its opponent GM, GM apparently did not believe its workforce capable of innovation resulting in profits, such that the thought of providing Autonomy Support may have never occurred.
In this chapter, we focus on investments in Autonomy Support that encourage innovative activities. Autonomy and Autonomy Support however have been generally found to positively impact well-being and performance of individuals (Gagné and Deci, 2005). A natural extension of our model is therefore to include a standard task. We make the case that Autonomy Support must be properly designed to be encourage innovation. But as this involves looking deeply into the task at hand and how to improve it, it seems obvious that even Autonomy Support tailored towards innovative activity would to some extend spill over to effort in the standard task. As this increases the principal's profits from innovation and regular business from the same investments, it amounts to an increase in v_P in our model, resulting in higher investments, investments in more periods, and investments for agents with smaller initial amounts of Autonomy Support. The interaction between standard and innovative task could, for example, be captured by allowing successful innovation to directly reduce effort costs in the standard task. Additionally to the agent's benefit from innovation v_A, currently the utility from being active, this generates a direct utility advantage. Generally, a model incorporating the standard task needs to include a monetary exchange. As shown in the literature review, monetary incentives do not facilitate generating creative ideas, posing the interesting question of whether innovation efforts should be compartmentalized as to not be perceived to have monetary rewards.

2.6 Conclusion

Our research contributes to the discussion on optimal management practices. Instead of isolating the effect of a single management practice, we focus on the impact of Autonomy Supportive leadership behavior on innovative effort. Leadership behavior is not contingent on a specific management practice, it can or cannot arise within one structure (Amabile et al., 2004), but may more readily appear in a management practice designed for support, such as in our introductory example of Toyota's Kaizen. Nonetheless, Autonomy Support fosters a feeling of Autonomy, an innovative action, only when given frequently, and in accordance to the worker's need. Providing this Autonomy Support is therefore an investment on behalf of the supervisor.

In this chapter, we demonstrate the optimal investment patterns in Autonomy Support over time that fosters small scale innovation. Small scale innovation describes creative, unconstrained innovative efforts of an agent. We review research showing that monetary incentives are restricted in their effectiveness for unconstrained innovation but that an Autonomy Supportive leadership behavior successfully instills innovative efforts. Leadership behavior unfolds over time, as a single intervention's effectiveness fades out over time. Informed by the literature, we build a two-period model of the principal's optimal investments in Autonomy Support. We find that when the principal values future benefits from innovative effort, she tends to invest already in the first period. When future benefits are not strongly valued, the principal tends to retouch depreciated Autonomy Support levels in the second period. Our model also suggests a certain substitutability between investments and the agent's initial Autonomy Support level, which he has accrued in previous (personal or professional) relationships. However, screening for agents with high initial Autonomy Support, or creativity levels, does not resolve the need to invest in Autonomy Supportive leadership behavior, echoing the literature (Mumford et al., 2002). Expanding the results of our two-period model indicates that even individuals with relatively high initial levels should be invested in at some point to preserve their innovative efforts. Workers endowed with almost zero initial Autonomy support receive no investments in our model; with a wide time horizon, investments in them become more likely.

One can argue that some supervisors already manage their workforce intuitively in this way. However, as a non-monetary incentive, Autonomy Support may have interactions with other incentives and should be provided consciously and in a structured way. As is the case with incentives in general, Autonomy Support must be tailored to be conducive to the desired outcome, in our case innovation, and given in the required intensity. Our research contributes thus to the discussion on non-monetary incentives, but we do not add yet another tool to the incentive toolbox. The concept of Autonomy Support touches upon motivation through decision rights, rewards, verbal praise and knowledge sharing, but encompasses these factors to feed into a feeling of Autonomy that successfully instills motivation (Gagné and Deci, 2005).

In this, Autonomy Support allows us to expand our understanding of how known incentives effectively work together. We further point towards time dependent need and effectiveness of incentives.

We also demonstrate that different leadership behavior intensity can constitute optimal Autonomy Support, as optimal investments depend on the agent's initial level. Leadership of a team then entails that the same action inspires different levels of Autonomy Support for each group member. It is therefore the responsibility of the team leader to understand which actions and behaviors must be taken on a team and which on an individual level in order to provide the Autonomy Support needed for each team member. With an Autonomy Supportive leadership style, the supervisor outgrows the role of controlling the workforce; she becomes a service provider who helps her team achieve the best for the entire organization.

Chapter 3

The Dry Powder Paradox of Monetary Policy*

3.1 Introduction

Should central bankers be rewarded for keeping interest rates up? And should monetary policy "keep its powder dry" in anticipation of deteriorating economic conditions? The notion of dry powder assumes that the maximum potential effectiveness of monetary policy is a state variable that follows a stochastic process dependent on past (and future) use of it. In typical models designed for studying "optimal" monetary policy and its role for macroeconomic stabilization, monetary policy cannot be stored and thus, interest rate setting is a state-contingent equilibrium process, not a goal of stabilization itself. In this chapter, we shed light on the theoretical underpinning of this dry powder view and answer the question of whether a central bank should set interest rates precautionarily according to the New Keynesian paradigm. A central bank is said to set interest rates precautionarily if it has asymmetric preferences concerning deviations of interest rates from their natural level.

The narrative of dry powder is prevalent across professionals and the public regarding both monetary and fiscal policy. There are several instances where headlines after meetings of the governing council of the Fed, the ECB, or the Bank of England make explicit reference to the

* This chapter is based on joint work with Markus Epp (University of Freiburg).

argument.[2] The argument resounds when economic conditions are expected to deteriorate in the near future so that further interest rate cuts are expected to be necessary. While there is some consensus about the notion of “fiscal space”, i.e. the limited capacity of governments to run deficits (due to borrowing limits and Laffer curves in taxation revenues or seignorage), the potential of interest rate cuts to amplify economic activity is, presumably, restrained by a binding lower bound on the nominal interest rate as in (Brunnermeier and Koby, 2016), also referred to as the *reversal rate*. Hence, a potential rationale for promoting the "dry-powder-argument" might be that setting interest rates (precautionarily) above their natural level may offer some leeway with respect to such a binding constraint.

Note that in this chapter, we do not attempt to answer the question of whether central banks actually "keep their powder dry" and set interest rates precautionarily. This clearly is an empirical question that goes beyond the scope of the present work. Further, we acknowledge that finding conclusive evidence for this point may prove challenging because several factors impair the distinction of central banks’ motives: First, conflicting regional interests obfuscate a clear distinction of fundamental components from deliberate precaution in driving observed monetary policy, a problem particularly associated with heterogeneous monetary unions like the eurozone. Second, central banks’ might exhibit hidden preferences for interest rate smoothing.[3]

We do not attempt to settle the controversy around deviations of actual interest rate setting from natural interest rates with this work, not to mention the associated difficulties arising with estimations of the latter. Instead, we analyze institutional circumstances under which precautionary interest rate setting (PIRS) may or may not occur as an equilibrium outcome in the New Keynesian core model, as well as the welfare consequences when it does.

Given the central bank has a dual mandate with a certain degree of freedom in putting weights on price stability and economic activity, our results are as follows: First, letting a discretionary central banks’ objective function also host explicit preferences for keeping positive interest rates, the associated PIRS yields the *dry powder paradox*: it leads to a deflationary bias and lower output gaps which in turn lead to on average *lower* nominal interest rates.

[2] See The Economist (2014) on Feb 6th: “Keeping its powder dry”; Goodman (2019) in *Bloomberg.com* on Dec 19th: “Bank of England Keeps Powder Dry as Brexit Moves Into Next Stage”, or Brzeski (2020) on Apr 30th: "ECB: Keeping its powder dry".

[3] Woodford (2003) shows that such behavior might be optimal even when society has no explicit preference for smooth interest rates.

Nevertheless, the created deflationary bias can be helpful when addressing the time inconsistency problem, which arises when the central bank cannot commit to a policy plan and the economy's steady-state is distorted (or alternatively: perceived to be distorted). We show that the welfare-optimal incentive for positive nominal interest rates has closed-form expressions. If the central bank and society have asymmetric preferences with respect to their weights on the output gap, depicted in their loss functions, the optimal incentive is strictly positive if the central bank is less "conservative" in comparison. Further, if the preferences perfectly coincide, the size of the optimal incentive is still strictly positive due to the welfare gains it provides by being a tool to overcome the time inconsistency problem. Further, our analysis shows that introducing the interest rate objective explicitly allows the central bank to balance its response to shocks with its goal to obtain a certain deflationary bias.

However, the welfare gains due to PIRS arising from overcoming the time inconsistency problem vanish when a central bank can credibly commit. In this case, a higher incentive for positive nominal interest rates yields on average higher nominal interest rates which in turn lead to higher inflation in the long-run. Hence, when the objective actually is to create some leeway away from a binding zero lower bound – a case in which the central bank is unable to match the decrease in the natural rate of interest with a proportional reduction in its policy interest rate due to a binding constraint for the latter – such a reward might help to accomplish that goal. In other words, given the central bank can credibly commit and thereby anchor expectations, our theory states that a reward on positive nominal interest rates can be used to avoid potential market fragilities at the zero(or effective)-lower bound (Fischer, 2016; Orphanides, 2020). Nevertheless, not-considering the existence of an effective lower bound, there is no welfare-argument in this case for a positive weight on keeping interest rates at higher levels.

Additionally, our analysis elucidates the potential for PIRS to be a valid strategy for a central bank having the objective of keeping control over its instruments to fulfill its mandate(s): it can be a tool to avoid the scenario of "fiscal dominance", in which interest expenses of a government cannot be covered by tax collection or rollovers quickly enough, such that the central bank might be tempted or even forced to sacrifice its main policy instrument to sustain government-solvency.

The chapter is structured as follows: in Section 3.2, we review the related literature. Section 3.3 highlights the role of incentivizing a central bank for positive nominal interest rates in the

New Keynesian model with discretionary and/or commitment policies, bringing about the dry powder paradox as well as a resulting welfare analysis. Section 3.4 presents PIRS in a setup of potential institutional conflict between a central bank and fiscal policy, in which its role for disciplining public spending is analyzed. Section 3.5 concludes this chapter.

3.2 Related Literature

Our approach builds on several branches of literature analyzing welfare effects of interest rate policy regimes by using the framework of the New Keynesian model, which presents the core amongst various strands of DSGE models. First, our conducted welfare analysis is located inside the area of research assessing macroeconomic effects of commitment of a central bank and the potentially inefficient outcomes implied by purely discretionary policies, in particular, in the presence of a (suboptimal) positive inflation bias due to time inconsistency (Kydland and Prescott, 1977; Barro and Gordon, 1983; Walsh, 2003; Galí, 2018). As, for example, the analysis of Clarida et al. (1999) shows, a central bank that is able to credibly commit to a choice of an optimal monetary policy plan that simply is followed through afterwards can overcome the inconsistency in hand and reduce the resulting welfare losses due to the possibility of stimulating the (rational) expectations of the economic agents about future output gaps and inflation rates. In contrast, a central bank only capable of acting discretionary, i.e. period-per-period, does not take into account the effect of its policy at one point in time on its own objective in previous and later periods and thus, generates outcomes which are inferior with respect to welfare, i.e. higher positive output gaps only a the expense of higher average inflation.

In our approach, we show that the inclusion of an artificial optimally specified reward (or weight) on keeping interest rates on average on a higher level for a central banker might serve as a remedy for the welfare losses in the discretionary case by providing an incentive to reach the desired lower average level of inflation.

Further, our analysis also contributes to the game-theoretical literature about how a government/a society optimally delegates monetary policy to a central bank (Illing, 1997) and the kind/degree of a central bank's autonomy. According to Fischer (1995), one can distinguish between *goal independence* and *instrument independence*. The first expression refers to a

regime in which a central banker is elected by a government but afterwards is autonomous in the determination of her goals. The latter expression refers to a scenario in which the goals for a central bank are pre-specified, but it is free in its choice of instruments to reach them.
In its results, our approach can be interpreted in both ways: First, our set-up may represent an attempt to optimally design a contract for central bankers (Persson and Tabellini, 1993; Walsh, 1995; Svensson, 1997; Rostagno et al., 2019). The roots of this modelling approach lie in contract theory, referring to an efficient incentivization of an agent by a principal. In the context of monetary policy, this implies that a central bank gets an explicitly formulated contract, specifying rewards dependent on a specified set of goals (Reis, 2013). The inclusion of an explicit reward for higher interest rates is in line with such a modelling approach.
Nevertheless, our set-up can also be related to Rogoff (1985) who proposed the election of a conservative central banker in presence of the issue of dynamic inconsistency. In contrast to specifying a contract rewarding higher nominal interest rates, a government/society might simply elect a central banker exhibiting preferences consistent with that goal. This point of view is more in line with goal independence. After the delegation of monetary policy, such a central banker is independent in picking her goals according to her preferences as well as the instruments to reach it. Both types of modelling are in line with our analysis and do not change the results presented below. Thus, we refrain from making an explicit choice about the exact scenario.
In its results, our approach contributes to the literature evaluating optimal policies in the presence of a binding effective/zero lower bound (Eggertsson and Woodford, 2003; Jung et al., 2005; Adam and Billi, 2007; Nakov, 2008; Nakata and Schmidt, 2019). With a zero lower bound constraint, there is no guarantee that the central bank is capable of stabilizing the economy in a downturn, leading, for example, to a reduction in output and an overall decrease in economic welfare (Galí, 2018). The existence of a binding lower bound might be one of the intuitive rationales for rewards for central bankers to sustain higher nominal interest rates in the first place. Hence, our work contributes to this literature by showing that this policy, even though it appears intuitive, potentially causes adverse effects on the average level of interest rates. Thus, it cannot unequivocally be considered as appropriate to achieve the underlying goal of gaining distance to a binding lower bound.
Furthermore, our approach contributes to the literature concerning optimal monetary policy

design in the New Keynesian model in the presence of fiscal dominance, or the threat of unsustainable fiscal policies (De Resende and Rebei, 2008; Kumhof et al., 2010; Leith and Wren-Lewis, 2013; Dufrénot et al., 2018). Our focus thereby is not so much on potential policies providing a remedy given the state of fiscal dominance already occurred but about potential measures a central bank can take to prevent such a situation in the first place.

3.3 The Dry Powder Paradox in the New Keynesian Model

To illustrate the dry powder paradox, we base our analysis on Gali's structural framework (Galí, 2015, ch. 3-5) which became the industry standard of linear monetary policy analysis and the point of departure for more intricate DSGE models. It consists of the following equations:

$$x_t = \mathbb{E}_t x_{t+1} - \sigma^{-1}[i_t - \mathbb{E}_t \pi_{t+1} - r_t^e] \tag{3.1}$$

$$\pi_t = \beta \mathbb{E}_t \pi_{t+1} + \kappa x_t \tag{3.2}$$

where $x_t = y_t - y_t^e$ is the output gap defined as the deviation of the log of output from its efficient level y_t^e, and where r_t^e is the natural rate of interest consistent with the efficient level of output. Our analysis thus allows for deviations of the desirable output gap from its natural level y_t^n. Both $y_t^e \geq y_t^n$ are assumed to be determined by structural parameters and technological capacities consistent with a zero-inflation steady state. σ is the intertemporal elasticity of substitution, β is the steady state discount factor and κ is function of technological parameters.

This model is usually derived from households' optimization on their intertemporal allocation of consumption plus the labor-leisure trade-off, entering the optimal staggered price setting of firms as a measure of marginal cost. After market-clearing (and log-linearization around a zero-inflation steady state), the household's Euler equation prescribes an IS-relationship (3.1) while optimal price setting implies the well-known New Keynesian Phillips curve (3.2).

A second-order approximation of the consumer's corresponding optimization problem yields that social welfare losses relevant for monetary policy are proportional to the following functional expression (Woodford (2003), ch. 6; Galí (2015), ch. 5) after normalization:

$$\mathbb{L} = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{1}{2} \left(\pi_t^2 + \hat{\alpha} x_t^2 \right) - \hat{\lambda} x_t \right] \tag{3.3}$$

where $\hat{\alpha}$ and $\hat{\lambda}$ are coefficients of structural parameters.[4] Note that this welfare loss can be expanded by using the variance decomposition of the unconditional second moment of inflation and output:

$$\mathbb{L} = \sum_{t=0}^{\infty} \beta^t \left[\frac{1}{2} \left([\mathbb{E}_0 \pi_t]^2 + \hat{\alpha} [\mathbb{E}_0 x_t]^2 + \hat{\alpha} \mathrm{Var}[x_t] + \mathrm{Var}[\pi_t] \right) - \hat{\lambda} \mathbb{E}_0 x_t \right] \tag{3.4}$$

which allows us to disentangle the losses created by the stochastic environment (represented by u_t and r_t^e in the structural description (3.1) and (3.2)) from the average biases in policy-making. These are relevant for comparing monetary policy frameworks consistent with identical responses to exogenous events.

In practice, central banks do not observe the structural parameters and thus have either different preferences than $\hat{\alpha}$ and $\hat{\lambda}$ or incomplete knowledge about them. We denote the central bank's (potentially differing) preferences by α and λ. A central bank's loss function consistent with the dual mandate[5] that has been mostly investigated in the literature is given by

$$\mathbb{L}^T = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{1}{2} \left(\pi_t^2 + \alpha (x_t - \bar{x})^2 \right) \right] \tag{3.5}$$

which is equivalent up to a constant with

$$\mathbb{L}^T = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{1}{2} (\pi_t^2 + \alpha x_t^2) - \lambda x_t \right] \tag{3.6}$$

[4] In Galí (2015), they are given by

$$\hat{\alpha} = \frac{\hat{\omega}}{\epsilon} \left(\sigma + \frac{\varphi + \theta}{1 - \theta} \right), \; \hat{\lambda} = \frac{\hat{\omega}}{\epsilon} \Phi$$

where Φ is the steady state distortion, $1 - \theta$ is the production elasticity of labor, φ is the Frisch-elasticity of labor supply, ϵ the substitution elasticity of goods and $\hat{\omega}$ is a function that is strictly decreasing in the degree of price stickiness, in the return to labour as well as the demand-elasticity. Similar proofs can be found in King and Kerr (1996), Bernanke and Woodford (1997), Rotemberg and Woodford (1997), McCallum and Nelson (1997).

[5] Central banks' legitimacy in most developed economies is mandated to price stability and the support of economic growth aside maintenance of financial stability by means of liquidity provision.

where $\lambda = \alpha\bar{x}$, and $\bar{x}$ is the socially desired output gap.[6]

We will exemplify below whether a central bank could find it optimal to deliberately minimize

$$\mathbb{L}^P = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{1}{2}(\pi_t^2 + \alpha x_t^2) - \lambda x_t - \delta i_t \right] \tag{3.7}$$

subject to (3.1) and (3.2). Since we are not interested in the dynamics at the lower bound itself but in the leeway a central bank has towards a lower bound, we neglect the explicit zero-lower bound constraint. The existence of such a lower bound only affects the relative comparison between discretionary and commitment solutions. Note, however, that the term δi_t can be interpreted as a reward for setting interest rates above the zero-level, i.e. above a lower bound $\bar{i} = 0$.

We stick to the usual denomination of α as a measure of how "conservative" the central bank is. In contrast, the parameter δ can be understood as a measure of how "Fisherian" the central bank is: it determines the nominal interest rate that the central bank is willing to set on average in order to anchor inflation expectations, given that long-run real interest rates are determined exogenously.[7] As will be illustrated, the Fisherian parameter allows the central bank to disentangle long-run biases from a desired amount of "dry powder", i.e. a long-run distance from hitting the zero-lower bound.

In order to verify these results, we will compare welfare losses and desirability of parameter δ in the central banks' loss function to achieve socially preferred outcomes under uncertainty concerning $\hat{\alpha}$ and $\hat{\lambda}$. We start our analysis for two popular regimes of credibility: in the first regime, the central bank cannot commit to paths of the interest rate, often referred to as "discretionary" monetary policy; in the second regime, the central bank commits to state-contingent paths of interest rates.

First, we will illustrate that the central bank can mitigate its time inconsistency problem that occurs under discretion. This will allow the central bank to get closer to the optimal inflationary (and thus output-)bias.

[6] In the New Keynesian model with steady state distortions, i.e. a long-run equilibrium deviating from the flexible price equilibrium, this "efficient" output gap arises naturally as a means to address inefficiently low production due to monopolistic competition.

[7] Put differently, δ is the relative weight a central banker puts on keeping the powder dry vis-a-vis its dual mandate.

3.3.1 Rational expectations equilibrium: The discretionary case

Suppose the central bank cannot credibly commit to paths of interest rates. As a result, the central bank is only able to engage in per-period optimization. Following Galí (2015), the central bank takes expectations to be predetermined and thus the structural constraints as given

$$\pi_t = \kappa x_t + \nu_t,$$
$$x_t = -\sigma^{-1}(i_t - r_t^e) + \gamma_t.$$

which assumes that supply-side cost-pressure $\nu_t = \beta \mathbb{E}_t \pi_{t+1} + u_t$ and demand-pull inflation $\gamma_t = \mathbb{E}_t x_{t+1} + \sigma^{-1}\mathbb{E}_t \pi_{t+1}$ are exogenously determined and thus are fixed values from the central bank's perspective.

The central bank thus solves

$$\mathbb{E}_t \left[\min_{\{x_t, \pi_t, i_t\}_{t \geq 0}} \left\{ \mathbb{L}_t^i \,\middle|\, \pi_t = \kappa x_t + \nu_t,\ i_t = r_t^e + \sigma(\gamma_t - x_t) \right\} \right]$$

where $i = \{P, T\}$ indicates the central bank's interest rate regimes as in (3.7) and (3.6). We neglect asymmetric inflation targets. Discretionary solutions can be found by looking at per-period optimality. Hence, for PIRS we have

$$\min_{x_t, \pi_t, i_t} \left\{ \frac{1}{2}(\pi_t^2 + \alpha x_t^2) - \lambda x_t - \delta i_t \,\middle|\, \pi_t = \kappa x_t + \nu_t,\ i_t = r_t^e + \sigma(\gamma_t - x_t) \right\} \tag{3.8}$$

Note that the New Keynesian IS curve is now required as a constraint since it is directly relevant for the losses incurred by the central bank due to the direct effects of i_t via δ.

Interior solutions are characterized by

$$(\alpha + \kappa^2) x_t + \delta\sigma - \lambda + \kappa \nu_t = 0$$

which determines the output gap, inflation and the interest rate as

$$x_t = \frac{\lambda - \delta\sigma}{\alpha + \kappa^2} - \frac{\kappa}{\alpha + \kappa^2}\nu_t \tag{3.9}$$

$$\pi_t = \frac{\kappa(\lambda - \delta\sigma)}{\alpha + \kappa^2} + \frac{\alpha}{\alpha + \kappa^2}\nu_t \tag{3.10}$$

$$i_t = r_t^e + \frac{\delta\sigma^2 - \sigma\lambda}{\alpha + \kappa^2} + \sigma\gamma_t + \frac{\sigma\kappa}{\alpha + \kappa^2}\nu_t \tag{3.11}$$

Note that $\delta = 0$ generates the usual discretionary solution for a distorted steady state and $\lambda > 0$. Further, (3.11) can be interpreted as a Taylor-type interest rate setting rule. Note also that expressions (3.9) - (3.10) exhibit the usual leaning-against-the-wind properties: aside from the biases (constant terms in π_t and x_t), the expressions multiplied with supply-side inflation pressures ν_t imply the usual sacrifice ratio $x_t = -\frac{\kappa}{\alpha}\pi_t$.[8]

For comparison, the common solutions under discretion and in absence of steady state distortions are given by

$$x_t = -\frac{\kappa}{\alpha + \kappa^2}\nu_t$$

$$\pi_t = \frac{\alpha}{\alpha + \kappa^2}\nu_t$$

which also illustrates that the reaction to variations of supply-sided inflation will be virtually equivalent under the assumed new structure of the central bank's loss function.[9]

The biases imply permanent positive/negative output gaps created by keeping the interest rate significantly below/above the natural rate of interest. The reactions to supply-side driven inflation $\nu_t = \beta\mathbb{E}_t\pi_{t+1} + u_t$ and "demand shocks" $\gamma_t = \mathbb{E}_t x_{t+1} + \sigma^{-1}\mathbb{E}_t\pi_{t+1}$ also mean that the Taylor principle is applied, since the coefficient on inflation expectations exceeds unity by $\alpha\beta/(\alpha + \kappa^2)$, which is larger than zero under the usual assumptions on parameters.

Inserting the explicit expressions for ν_t and γ_t, the dynamics of the equilibrium under the new

[8] Hence, this type of monetary policy is observationally equivalent to typical inflation targeting: the model predicts a negative empirical relationship between inflation and output gaps, as in McLeay and Tenreyro (2020).

[9] Without structural distortions or preferences over the instrument i_t, the interest rate rule needs to be specified by the researcher such that the model produces a unique solution. The proof of this result can be found in Bullard and Mitra (2002). Galí (2015) discusses the regions of determinacy for several rules in ch. 4.

specifications of $\lambda, \delta > 0$ can be expressed independently from γ_t as a system

$$(\alpha + \kappa^2) \begin{bmatrix} x_t \\ \pi_t \end{bmatrix} = A \begin{bmatrix} \mathbb{E}_t x_{t+1} \\ \mathbb{E}_t \pi_{t+1} \end{bmatrix} + B u_t + C$$

where

$$A = \begin{bmatrix} 0 & -\kappa\beta \\ 0 & -\alpha\beta \end{bmatrix}, \; B = \begin{bmatrix} -\kappa \\ -\alpha \end{bmatrix}, \; C = [\lambda - \delta\sigma] \begin{bmatrix} 1 \\ \kappa \end{bmatrix}$$

which implies that A has eigenvalues that solve

$$\epsilon \left(\frac{\alpha\beta}{\alpha + \kappa^2} + \epsilon \right) = 0 \; \Rightarrow \epsilon_{1,2} = \left\{ 0, -\frac{\alpha\beta}{\alpha + \kappa^2} \right\}$$

and hence the solution is unique when $\alpha(\beta - 1) < \kappa^2$, which is true for the standard assumptions of the parameters, i.e. $\alpha > 0$, $\beta \in [0, 1)$ and κ being non-negative. This property carries over to variations of the model where v_t has some exogenous stochastic component additional to inflation expectations.

The model's solution can thus be computed by iterating the inflation solution forward and inserting it back into the equation for the output gap, or simply using the method of undetermined coefficients. Using the latter, the solutions of π_t and x_t can be expressed as

$$x_t = x(\delta) + \chi_u u_t + \chi_r r_t^e$$
$$\pi_t = \pi(\delta) + \psi_u u_t + \psi_r r_t^e$$

with intercepts $x(\delta)$, $\pi(\delta)$ and coefficients $\chi_u, \chi_r, \psi_u, \psi_r$.

We highlight that the intercepts are functions of δ by specifying them as $x(\delta)$ and $\pi(\delta)$ which are determined independently of shocks. To see this, note that the constant non-stochastic solutions $x_t = x$ and $\pi_t = \pi$ are consistent with values

$$x(\delta) = \frac{(1-\beta)(\lambda - \delta\sigma)}{\alpha(1-\beta) + \kappa^2} \tag{3.12}$$

$$\pi(\delta) = \frac{\kappa(\lambda - \delta\sigma)}{\alpha(1-\beta) + \kappa^2} \tag{3.13}$$

which are not dependent on the stochastic structure of the shocks. Note also that the typical long-run Phillips curve holds and implies that there is no inflation-output trade-off in the long run since $x = \frac{1-\beta}{\kappa}\pi$.

Under the assumption of u_t and r^e_t following auto-regressive processes with persistence ρ_u and ρ_r, respectively, expectations of π_t and the output gap x_t are given by

$$\mathbb{E}_t x_{t+1} = x(\delta) + \chi_u \rho_u u_t + \chi_r \rho_r r^e_t$$
$$\mathbb{E}_t \pi_{t+1} = \pi(\delta) + \psi_u \rho_u u_t + \psi_r \rho_r r^e_t$$

Plugging these solutions into (3.11) , we find that the interest rate rule can be expressed as

$$i_t = \frac{\lambda - \sigma\delta}{\alpha + \kappa^2}\Delta + \phi_r r^e_t + \phi_u u_t \tag{3.14}$$

where $\Delta = \frac{\kappa(\alpha+\kappa^2)}{\alpha(1-\beta)+\kappa^2}$. We abstract here from stating parameters $\chi_u, \chi_r, \psi_u, \psi_r$ and the resulting ϕ_r, ϕ_u explicitly since their exact specification is unimportant for the results in hand.

This solution for the equilibrium interest rate illustrates the dry powder paradox: the equilibrium interest rates are *lower* despite the central bank's incentive to set higher interest rates. The reason is that the direct incentive to set higher interest rates, as expressed in (3.11), is more than offset by reduced inflation and output expectations, anchored at lower average levels due to the deflationary effect caused by $\delta > 0$. Thus, the central bank will set lower interest rates than it otherwise would. In other words, δ works as a counter to the inflationary bias created by the steady state distortion and yields an incentive to tolerate deflation and negative output gaps.

By iterating optimality conditions (3.9) and (3.10), it also becomes clear that δ affects only the intercepts, not the responses of endogenous variables to exogenous variations in u_t and r^e_t. Note that this solution is consistent with the Fisher-equation holding in the shock-free (that is non-stochastic) steady state. In the long-run, the PIRS-Taylor-rule and the Fisher equation both imply

$$i_t = r^e_t + \pi(\delta)$$

which implies that nominal interest rates are determined by the stochastic process underlying the natural rate of interest and the choice of δ. However, since $\pi'(\delta) < 0$, not only the average inflation rate is smaller under precautionary interest rate setting but also the nominal interest rate, which is the essence of the dry powder paradox. There is no trade-off between the Fisher effect and the typical impact of higher interest rates on inflation since in the rational expectations equilibrium both effects point in the same direction.

3.3.2 Welfare analysis: Overcoming the time inconsistency problem

In the present framework, the inflationary bias is a result of the temptation to address steady state distortions using monetary policy, which will give rise to Kydland and Prescott's popular time inconsistency problem of discretionary policies (Kydland and Prescott, 1977). Consistent with the model above, the central bank is said to have an inflationary bias when $\lambda > 0$. Note that this inflationary bias is more aligned with society's preferences the closer it is to $\hat{\lambda} > 0$.

In what follows, we illustrate that the inflationary bias in the New Keynesian model arises primarily because of the distortions to the steady state, which makes some deviation of inflation from zero optimal. We thereby exploit that the reaction to exogenous shocks is equivalent under all values of δ under discretionary policy (as illustrated by our results above).

First, note that the differences in welfare losses are proportional to the biases in (3.4), so that welfare comparisons of different discretionary policy schemes come down to assessing different values of δ in societies' intertemporal loss function:

$$\mathbb{L} = \sum_{t=0}^{\infty} \beta^t \left[\frac{1}{2} \left([\mathbb{E}_0 \pi_t]^2 + \hat{\alpha} [\mathbb{E}_0 x_t]^2 \right) - \hat{\lambda} \mathbb{E}_0 x_t \right] = \frac{\Psi(\hat{\alpha})}{2(1-\beta)} \bar{\pi}^2 - \frac{\hat{\lambda}}{\kappa} \bar{\pi} \tag{3.15}$$

where we exploit the following two points: (i) unconditional expectations $\mathbb{E}_0 \pi_t$ coincide with both the average inflation $\bar{\pi}$ in a certainty-equivalent steady state, and the non-stochastic steady state of the model (ii) the long-run Phillips curve prescribes $x = \frac{1-\beta}{\kappa} \bar{\pi}$, implying that $\Psi(\hat{\alpha}) := 1 + \hat{\alpha} \left[(1-\beta)/\kappa \right]^2 > 1$ and, most importantly, that society cannot attain $x_t = \bar{x}$ and $\pi_t = 0$ in the long-run in the presence of rational expectations.

As we have shown above, the discretionary central bank sets $\bar{\pi}$ for a given δ and λ according to expression (3.13).

When society and central bank share the perception of an undistorted steady state $\bar{x} = 0$ (or alternatively, $\hat{\lambda} = \lambda = 0$), expression (3.15) immediately implies that $\delta = 0$ is unambiguously optimal for welfare. If the natural level of output is not inefficiently low, a central bank is always capable to achieve zero inflation and no output gap on average, commonly referred to as the *Divine Coincidence* (Blanchard and Galí, 2007). The inclusion of an incentive to keep nominal interest rates higher on average in this case only leads to an artificially created deflationary bias and negative output gaps which clearly is welfare inferior.

In contrast, suppose that the steady state is distorted so that society's preferences have $\hat{\lambda} > 0$, which implies that policy frameworks with discretionary leaning-against-the-wind are evaluated with regard to the trade-off expressed in (3.15). When the central bank adopts $\lambda > 0$ to conform $\hat{\lambda} > 0$, optimal discretionary monetary policy produces the well-known inflationary bias which for the given model is given by $\lambda/((1-\beta)\alpha + \kappa^2)$ (see the discretionary solution (3.10) for $\delta = 0$). Under these circumstances, the deflationary bias created by $\delta > 0$ can help to offset the time inconsistency problem that arises.

Using $\hat{\lambda} = \hat{\alpha}\bar{x}$, the bias optimal for society in inflation (and the associated bias in the output gap) are:

$$\pi^* = \frac{1-\beta}{\kappa}\frac{\hat{\alpha}}{\Psi(\hat{\alpha})}\bar{x}, \tag{3.16}$$

$$x^* = \left(\frac{1-\beta}{\kappa}\right)^2 \frac{\hat{\alpha}}{\Psi(\hat{\alpha})}\bar{x} \tag{3.17}$$

where we find π^* by minimizing (3.15) with respect to $\bar{\pi}$. Afterwards, we use the long-run Philips curve to find x^*.

The results illustrate that in the presence of the long-run correlation between inflation and output deviations, society cannot reach the allocation $x_t = \bar{x}$ and $\pi_t = 0$ and thus picks the middle ground with some inflation and some output stimulus.

Time inconsistency implies that a non-committing central bank reaching out for $\bar{x}$ will fail to achieve this goal if agents with rational expectations take the central bank's per-period incentives into account and will only achieve (3.16) by coincidence. Nevertheless, picking an optimal incentive scheme δ can remedy this situation: according to the biases in discretionary solutions, equations (3.12), (3.13) show how the average inflation and output gap are determined for $\delta > 0$.

Hence, we can use, e.g., expressions (3.16) and (3.13) to solve $\pi^* = \pi(\delta)$ for the optimal value of δ. We find:

$$\delta^* = \frac{\bar{x}}{\sigma}\left[\frac{\alpha\Psi(\hat{\alpha}) - \hat{\alpha}\Psi(\alpha) + \hat{\alpha}\beta}{\Psi(\hat{\alpha})}\right] \tag{3.18}$$

which is weakly positive if $\alpha \geq \hat{\alpha}$, i.e. when the central bank puts more weight on output gaps than society. The inequality holds only weakly because, as already mentioned above, $\delta^* = 0$ if $\bar{x} = 0$: the time inconsistency problem does not arise if society and central bank agree that the steady state is undistorted.

Another interesting result arises when the case of $\alpha = \hat{\alpha}$ is considered, i.e. society and central bank agree on the weight they put on a non-negative output gap. Expression (3.18) can be used to quantify the time inconsistency bias since then

$$\delta^* = \frac{\bar{x}}{\sigma}\left[\frac{\alpha\beta}{\Psi(\alpha)}\right]$$

holds. This result shows that the optimal reward for positive interest rates, depending on the weight the central bank puts on the output gap aside structural parameters, is still positive. The reason is that the deflationary bias created by δ represents a welfare improving countereffect to the inflationary bias innate to the time-inconsistency problem: if a central bank cannot credibly commit, it fails to achieve the optimal inflation level π^* without the additional bias induced by δ because agents with rational expectations know that the central bank is subject to its per-period incentive to achieve $\bar{x}$. A discretionary central bank with $\delta = 0$ that would announce to pursue an inflation target π^* has an incentive to deviate to a higher level of inflation to reach the desired output gap $\bar{x}$. Since the economic agents hold rational expectations, however, they would anticipate this deviation, leading to the inflation bias. Thus, a positive δ^* addressing this bias by providing a credible incentive for the central bank to pursue π^* is still welfare improving, even when the preferences of society and central bank formally coincide.

Note that for the economy converging to the flexible price equilibrium, i.e. $\kappa \to 0$, this expression simply boils down to $\delta^* = 0$. The reason is that in this case, the New Keynesian Phillips curve (expression (3.2)) states that there is no connection between a positive output gap and inflation. Thus, a discretionary central bank can reach the desired output gap $\bar{x}$ without inflationary bias. This implies that there is no need for a positive δ in this case.

3.3.3 Commitment solution: Fisher effects of interest rate rewards

When the central bank can commit, it will take implications for future inflation and output gaps into consideration when choosing inflation and output today. It is well-known that credible commitments of the central bank to state-contingent paths of future interest rates help the central bank to replenish the solution to the Ramsey problem of minimizing (3.6) subject to (3.2), typically referred to as forward guidance solution. Hence the question here is not so much about potential improvements of the reward δ but about the qualitative implications such a reward may have on interest rate setting as well as on the "stock of dry powder", i.e. the distance from the zero/effective lower bound a central bank will have on average.
Again using the set-up presented in Galí (2015) as a foundation, the central bank's optimization scheme can be expressed by the following Lagrangian, when confronted with the incentive to keep interest rates positive and the ability to commit to state-contingent plans:

$$\mathscr{L}_0 = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{1}{2}(\pi_t^2 + \alpha x_t^2) - \lambda x_t - \delta[\sigma(x_{t+1} - x_t) + \pi_{t+1} + r_t^e] + \mu_t(\pi_t - \kappa x_t - \beta\pi_{t+1} - u_t) \right] \tag{3.19}$$

where $\{\mu_t\}_{t=0}^{\infty}$ denotes a sequence of Lagrange multipliers.
Thus, the necessary conditions to characterize optimality are given by

$$\alpha x_0 - \lambda + \delta\sigma = \kappa\mu_0$$
$$\pi_0 + \mu_0 = 0$$

and, $\forall\, t > 0$

$$\beta^t[\alpha x_t - \lambda + \delta\sigma - \kappa\mu_t] - \beta^{t-1}[\delta\sigma] = 0$$
$$\beta^t[\pi_t + \mu_t] - \beta^{t-1}[\delta + \beta\mu_{t-1}] = 0$$

which is equivalent to

$$\alpha x_t - \lambda - \delta\sigma\frac{(1-\beta)}{\beta} = \kappa\mu_t$$
$$\pi_t + \mu_t - \frac{\delta}{\beta} = \mu_{t-1}$$

Eliminating the Lagrangian multipliers, we get a single equation correlating inflation and output gaps for all time periods $t \geq 0$:

$$\pi_0 = \frac{\lambda - \delta\sigma}{2} - \alpha x_0,$$
$$\pi_t = \frac{\alpha}{\kappa}[x_t - x_{t-1}] + \frac{\delta}{\beta},\ \forall t > 0$$

which shows that precautionary interest rate setting survives in the commitment solutions as a permanent inflationary bias in all periods.

Given the ex ante equality of unconditional expectations $\mathbb{E}_0 x_t = \mathbb{E}_0 x_{t-1}$ inflation's unconditional expectation must be given by

$$\mathbb{E}_0 \pi_t = \frac{\delta}{\beta} \tag{3.20}$$

which illustrates that the central bank has an inflationary bias in this case. This can be attributed directly to the Fisher effect: since the central bank can engineer higher nominal interest rates in the future by increasing inflation expectations *today* it is tempted to move along a higher average inflation path. This "backfiring" of the Fisher effect can be directly attributed to the dependence of the IS curve (3.1) on real, not nominal interest rates: since the incentives of central bankers to raise inflation are permanent, inflation expectations will be higher in every equilibrium.

Now it is straightforward to show that the standard log-linearization of the inflation rate, i.e. the definition $\pi_t = p_t - p_{t-1}$, the starting value p_{-1} and the starting bias $\frac{\lambda - \sigma\delta}{2\kappa}$ can be used to iterate for a solution of the price level, such that the price level definition that suffers from variations in the output gaps is given by:

$$\tilde{p}_t = \frac{\alpha}{\kappa} x_t \tag{3.21}$$

where $\tilde{p}_t = p_t - \hat{p}_t$, and $\hat{p}_t = p_{-1} + \frac{\delta}{\beta} t + \frac{\lambda - \sigma\delta}{2\kappa} - \frac{2\alpha}{\kappa}$ which captures the result that the central bank pursues a positive inflation target and has a sacrifice ratio between output gaps and the deviation of the price level p_t from its deterministic growth path $\frac{\delta}{\beta} t$. This is to say that the central bank expects and commits to some constant price level appreciation.

Reformulating the IS-curve depicted in (3.1), note also that we can express the interest rate equation as

$$i_t = r_t^e + \mathbb{E}_t \pi_{t+1} + \sigma \mathbb{E}_t [x_{t+1} - x_t]$$

which can be rewritten with the help of the expectations of the optimality condition under commitment. Thus, the following expression for the nominal interest rate can be derived:

$$i_t = r_t^e + \frac{\delta}{\beta} + \left(\sigma + \frac{\alpha}{\kappa}\right) \mathbb{E}_t [x_{t+1} - x_t] \tag{3.22}$$

Using once again the ex-ante equality of unconditional expectations $\mathbb{E}_0 x_t = \mathbb{E}_0 x_{t-1}$, the unconditional expectation of i_t is given by

$$\mathbb{E}_0 i_t = r_t^e + \frac{\delta}{\beta} \tag{3.23}$$

Hence, due to the positive reward δ, the central bank gains on average higher nominal interest rates at the cost of a positive inflationary bias when it can commit to a policy plan.

3.3.4 Discussion of results

The analysis presented above sheds some light on the different effects of a positive δ on the optimal policy of a central bank and the corresponding welfare effects. In particular, when considering the introduction of such an incentive, an important point to measure its effectiveness is the ultimate goal for which it was introduced.

First, consider the case in which PIRS is considered to avoid interest rates coming to close to an effective lower bound. Using our analysis from above, we have shown that whether the resulting effect of introducing an incentive δ is inflationary or deflationary crucially hinges on the central bank's ability of commitment. I.e., whether a reward on higher interest rates actually leads to higher interest rates depends on a central bank's ability to commit to a policy plan.

Given a central bank's inability to commit, $\delta > 0$ anchors inflation and output gap expectations on a lower average level. As a result, those effects overcompensate the direct effect on i_t and lead to lower nominal interest rates, on average. Thus, when the goal of the policymaker actually is to get some leeway away from an effective lower bound and the central bank can only act discretionary, $\delta > 0$ is a counter-productive instrument because it leads to the above mentioned dry powder paradox.

On the other hand, the commitment case obviates the Fisher effect, since commitment removes the time inconsistency so that only the permanent effect of precautionary interest rate setting inflation survives. If a $\delta > 0$ is used in this case, the central bank is able to engage in some "fine-tuning" of long-run inflation to its desired level, accompanied by higher average nominal interest rates.

Second, consider the case of introducing PIRS to improve welfare. From a purely welfare-oriented point of view, without considering the existence of a zero/effective lower bound, a positive reward for nominal interest rates makes no sense in the commitment case: As e.g. Galí (2015) shows, a central bank can address the time inconsistency problem without $\delta > 0$ in such a case and reaches the social optimum. Using its optimal plan $\{\pi_t, x_t\}_{t=0}^{\infty}$, it is able to (asymptotically) reach an equilibrium with zero average inflation and no average output gap. Since the public anticipates this outcome, this policy enables the central bank to achieve positive output gaps at lower inflation in the short-run.

If commitment is not an available option and assuming a distorted steady state, however, the analysis above demonstrates that the deflationary bias a $\delta > 0$ causes is welfare-enhancing because it presents a counter effect to the inflationary bias of a central bank and thereby addresses the time inconsistency problem. This, of course, only holds for as long as a distorted steady state, i.e. the positive inflationary bias, exists. In absence of this case, any incentive to artificially set higher nominal interest rates is welfare inferior in comparison to e.g. the standard Taylor interest rate rule that is able to establish on average zero inflation and no output gap in equilibrium (Woodford, 2003).

3.4 Strategic Reasons for Precautionary Interest Rate Setting

Does a reward for higher interest rates help the central bank to keep fiscal policy in check? The underlying assumption of this question is that fiscal policy tends to run unsustainable deficits, antagonizes stabilization attempts of central banks, or at least induces an amount of economic activity inconsistent with the dual mandate of the central bank. A rich literature has studied these hypotheses and continuous to do so.[10] Without taking a stand in these debates, we study an extreme situation which arises when the central bank has to keep fiscal policy afloat by monetizing deficits in the treasury: a situation that arises naturally even when the government keeps deficits contained in expectations, but where the price stability mandate urges increases in interest expenses which cannot be covered by tax collection or rollovers quickly enough so that the central bank has to sacrifice its policy instrument to maintain the solvency of the government. This situation is usually referred to as *fiscal dominance* and its likelihood might depend positively on expansionary monetary policy followed by a quick and significant turn to contractionary monetary policy. Piled-up fiscal deficits financed by debt issuance during expansions then translate into higher interest expenses once the stance of monetary policy is reversed. If the increase in public debt cannot be rolled over at the interest rate set by the central bank, the treasury might face a self-enforcing spiral of insolvency. Hence the question arises, whether the central bank can diminish chances for such a situation, for example by providing additional incentives for debt consolidation, thereby encompassing its dual mandate in a larger set of circumstances.

To answer this question within the New Keynesian model, we have to disclose the description of the government sector as is standard in the literature. First note that the structural equations (3.1)-(3.2) above can be generalized to adhere fiscal policy, when interpreting x_t as the difference between deviations of total spending and potentials from their steady states (or balanced growth paths).

The aggregate market-clearing condition (households' budget identity) is

$$Y_t = C_t + G_t \tag{3.24}$$

In its log-linear version, it accommodates the New Keynesian model's market-clearing

$$y_t = \zeta c_t + (1-\zeta) g_t \tag{3.25}$$

[10] An early and very influential treatment of optimal fiscal policy was provided by Arrow and Kurz (1970). Crucial contributions to the discussion of optimal fiscal policy rules are discussed in Leith and Wren-Lewis (2000) and Leith and Wren-Lewis (2013).

where ζ and $1-\zeta$ are the relative sizes of the private sector and the government spending in a steady state (since $c_t = \frac{C_t - C}{Y}$, $g_t = \frac{G_t - G}{Y}$), respectively. The typical Phillips curves (3.2) and IS-curves (3.1) apply but feature the adjusted output gap $x_t = y_t - y_t^e$ where

$$x_t = \zeta c_t + (1-\zeta) g_t - y_t^e \tag{3.26}$$

and where consumption is determined via the Euler equation

$$c_t = \mathbb{E}_t c_{t+1} - \sigma^{-1}[i_t - \mathbb{E}_t \pi_{t+1} - r_t^e] \tag{3.27}$$

The market-clearing condition obviates that by means of accounting, the central bank has to rely both on public and private sector demand in achieving the (price-)stabilizing output gap by exhausting production potentials y_t^e. For a given level of private activity, the output gap will only be closed for sufficient public demand. The interpretation of the IS-curve (3.1) is usually such that a sequence of consumption expenditures $\{c_t\}$ enables the path of output gaps $\{x_t\}$ in a Ricardian way, where paths of government expenditures, taxation, deficits and and public debt levels are exogenous, i.e. do not affect the real allocation $\{c_t\}$. This is consistent with public debt not reflecting real wealth to the private sector such that private activity is fully determined by monetary policy. Public debt titles are thus the (only) riskless vehicle available to households to smooth consumption over the infinite lifetime of the representative agent. This will allow us to let government expenditures be determined exogenously, which is consistent with optimal fiscal policy as found by Benigno and Woodford (2003) as well as the stochastic processes estimated for fiscal policy in the empirical literature (Jonsson and Klein, 1996; Traum and Yang, 2011; Barhoumi et al., 2016).

When all expenditures are financed using nominal short-term debt B_t and the central bank ensures solvency of the treasury in all states of the world, public debt evolves as

$$B_t^n = (1 + i_{t-1}) B_{t-1}^n + P_t G_t \tag{3.28}$$

for a given B^n_{-1}. G_t denotes real public expenditure (or primary deficits). Define real debt $B_t \equiv \frac{(1+i_t)B^n_t}{P_t}$ so that the law of motion (3.28) implies

$$B_t = \frac{1+i_t}{1+\pi_t} B_{t-1} + (1+i_t)G_t \tag{3.29}$$

which deviates from its stationary value according to a log-linear approximation (see Woodford (1998), equ. (2.9)) as

$$\beta b_t = \beta i_t + b_{t-1} - \pi_t + (1-\beta) g_t \tag{3.30}$$

We will assume below that monetary policy keeps the treasury afloat at all times in order to maintain the treasury's ability to cover real interest expenses to the public. The central bank will thus be subject to a constraint on interest rate setting, which will also consistently peg the default probability to zero in every equilibrium.

Fiscal Dominance

We follow Blanchard (2004) by specifying a risk premium on government debt which arises due to a default probability p_t on government debt.[11] Given this probability, expected returns on government bonds would equal $(1+i^B_t)(1-p_t)$, where bonds are priced via i^B_t such that they compensate risk averse market participants:

$$1 + i^B_t = 1 + i_t + \sigma p_t \tag{3.31}$$

where σ is the measure of risk aversion. The central bank sets $p_t = 0$ by credibly committing to bail out outstanding (nominal) claims on the treasury and thus ensures $i^B_t = i_t$ in this equilibrium. We assume that all such equilibria can be indexed by a maximum capacity for rolling over debt, Ω_t:

$$B_t - B_{t-1} \leq \Omega_t \tag{3.32}$$

[11] This default probability might alternatively capture the quasi-default event of financial repression and other pay-off relevant remedies to sovereign insolvency.

which by (3.29) immediately translates into a constraint on real interest expenses, given public spending and the inflation shield

$$(i_t - \pi_t)B_{t-1} \leq (1+\pi_t)[\Omega_t - (1+i_t)G_t] \tag{3.33}$$

where we think of Ω_t as the time-varying capacities of the government to issue new debt on financial markets and service old debt by means of (non-distortionary and revenue-enhancing) taxation.

As we have seen above, the discretionary central bank is subject to the dry powder paradox, which lowers both average inflation and nominal interest rates. Alternatively, as in the commitment case, both average inflation and nominal interest rates rise in response to an increase in incentives δ. This same-sign characteristic implies that the total effect of δ on the likelihood to breach this condition becomes a quantitative question.

For discretionary monetary policy, we obtain the following result: as is shown in the Appendix, increasing δ will reduce the likelihood to enter the stage of fiscal dominance if the size of the government sector is sufficiently large in comparison to the maximum rollover capacity in the steady state ($G > \frac{\Omega}{2}$). This finding is a result of the independence of real interest rates with respect to δ and the fact that inflation affects the rollover limit and government expenditures equally, while nominal interest rates affect the real value of government expenditures only. Hence, there are two competing effects of higher values of δ: a common effect reducing government solvency due to lower inflation (first term on the RHS in (3.33)) but also a diminished real cost of given government expenditures due to lower interest cost (second term in brackets in (3.33)). Both effects are driven by the dry-powder paradox.

At first glance, these results appear to be somewhat counter-intuitive: in the discretionary case, given a high level of government expenditures, the desirability of high values of δ does not come from a disciplining effect of a positive δ on public expenditures. Rather, they originate from the lower average nominal interest rates, decreasing the direct costs of public expenditures. Thus, these results hold because of the dry powder paradox we derived earlier.

Under commitment, the effects exactly revert: the increased cost of expenditures due to the on average higher nominal interest rates only lead to an increased likelihood of entering the stage of fiscal dominance for a central bank. Thus, under these circumstances, our theory

states that a positive δ is an invalid tool for avoiding a situation of fiscal dominance.
Hence, there is again a trade-off involved with adopting rewards δ: under discretion, the present framework predicts that δ helps reducing the likelihood of fiscal dominance scenarios, but comes at the cost of a higher chance of encountering the zero-lower bound. Under commitment, a central bank can gain additional leeway away from the lower bound, but only at the expense of reduced government solvency.

3.5 Conclusion

Even though its validity is disputed by many economists, the argument that central banks should "keep some powder dry" has sparkled the debate of monetary policy in the recent episode of binding lower bounds on the nominal interest rate. Not assuming that the central bank or professionals are irrational, the question arises why the narrative of storable nominal interest rates is so resilient, in particular, when macroeconomic models suggest that higher interest rates do not accomplish better stabilization, even in the presence of a binding lower bound.
In this chapter, we have studied institutional and strategic rationales for the use of an explicit incentive for a central bank to honor the DPA. We find that the structural rationale based on its welfare improvement hinges on the commitment ability of a central bank as well as the existence of a distorted steady state in the discretionary case. But even though an explicit incentive to increase nominal interest rates may help to overcome the time inconsistency problem, it counter-intuitively is an invalid instrument to increase average nominal interest rates (i.e., gain some leeway away from a binding lower bound) because of the existing dry powder paradox: rational expectations with respect to inflation and the output gap are anchored at lower average levels due to the deflationary bias this reward provides. This effect more than offsets its direct positive effect on the nominal interest rate. Thus, PIRS only leads to average nominal interest rates potentially moving even closer to an effective lower bound. Only if the central bank can credibly commit to a policy plan, such a positive incentive can help to accomplish higher average inflation rates due to higher nominal interest rates, as a result of the Fisherian feature of the New Keynesian model.
Further, we have shown how setting interest rate precautionary can decrease the threat of

fiscal dominance for a discretionary central bank. Interestingly, this result does not arise because of a disciplinary effect of PIRS on government expenditures. Rather, because it leads to decreased real costs of expenditure for a government due to the existence of the dry powder paradox. This result, however, was developed under the assumption that the path of government spendings follows an exogenous process, rather than being determined endogenously. Assuming that the government minimizes its own loss function would yield the best response function, directly linking the optimal level of spending with the nominal interest rate level the central bank sets. Thus, we think that considering feedback effects of interest rate setting on public spending and strategic interaction between the government and fiscal policy is an interesting avenue for extending the presented analysis.

To the best of our knowledge, there is no empirical study so far providing evidence on whether central bankers actually honour the argument of keeping some "dry powder" when setting their interest rates. We look forward to assessing the explanatory value of an interest rate preference for the actual conduct of monetary policy in future research.

Appendices

Appendix A

Random Error Neglect

A.1 Proposition 2: One sample prediction error

The statistically "correct" prediction error

According to von Auer (2005), the uncertainty associated with a prediction is captured by the variance of the point prediction, i.e. $Var[m - A_{n+1}]$. It can be expressed as

$$E[(m - A_{n+1})^2] - E[(m - A_{n+1})]^2 \tag{A.1}$$

Since m is an unbiased estimator of the Expected value of A_{n+1}, and m and A_{n+1} are independent, this can be simplified to

$$E[(m - X_{n+1})^2] \tag{A.2}$$

which is equivalent to

$$\begin{aligned} &E[m^2 - 2mX + X_{n+1}^2] \\ =&E[m^2] - 2E[m]E[X_{n+1}] + E[X_{n+1}^2] \end{aligned}$$

Knowing that for a random variable X, it holds that $Var[X] = E[X^2] - E[X]^2$, substitution leads to the following expression

$$Var[m^2] + E[m]^2 - 2E[m]E[X_{n+1}] + Var[X_{n+1}] + E[X_{n+1}]^2$$

With $m \sim N(\mu, \frac{\sigma^2}{n})$ and $X_{n+1} \sim N(\mu, \sigma^2)$, I find that the above expression equals

$$\frac{\sigma^2}{n} + \mu^2 - 2\mu^2 + \sigma^2 + \mu^2$$

which is equivalent to

$$\frac{n+1}{n}\sigma^2$$

The RENAs assessment

In contrast, the RENA believes to consider all possible data points in his sample and uses the true distribution parameters with no innate estimation uncertainty. Thus, as, for example, Geisser (1993) shows, all prediction uncertainty of using the mean as a predictor for a future outcome comes from the variance of the respective random variable to be predicted. The RENA (naively) assesses it via his sample-dispersion. Thus, the RENA believes his prediction error to be equivalent to his sample's variance, i.e.

$$\frac{1}{n-1}\sum_{i=1}^{n}(m - A_i)^2$$

or, on average

$$E\left(\frac{1}{n-1}\sum_{i=1}^{n}(m - A_i)^2\right) = \sigma^2$$

Hence, comparing both errors, the RENA underestimates the true uncertainty by a factor $\frac{n}{n+1}$, proving Proposition 2.

A.2 Proposition 3: Externally provided estimates

A.2.1 Correct assessment

In similar fashion to above, the accuracy of a fixed external predictor z is assessed. The correct evaluation of the goodness of this constant as a predictor is

$$E[(z - X_{n+1})^2] \tag{A.3}$$

which since z is exogenously given is equivalent to

$$\begin{aligned}
&z^2 - 2zE[X_{n+1}] + E[X_{n+1}^2] \\
=&z^2 - 2z\mu + Var[X_{n+1}] + E[X_{n+1}]^2 \\
=&z^2 - 2z\mu + \mu^2 + \sigma^2 \\
=&\sigma^2 + (z - \mu)^2
\end{aligned}$$

A.2.2 The RENA's assessment

In turn, the RENA evaluates z similar to the previous section by using his sample data only, i.e. in expectation, his assessment can be expressed as

$$E[\frac{1}{n-1}\sum_{i=1}^{n}(z - A_i)^2]$$

applying some zero-trick, this expression can be rewritten as

$$\begin{aligned}
&E[\frac{1}{n-1}\sum_{i=1}^{n}(z - m + m - A_i)^2] \\
=&E[\frac{1}{n-1}\sum_{i=1}^{n}[(z-m)^2 + 2(z-m)(m-A_i) + (m-A_i)^2] \\
=&E[\frac{1}{n-1}\sum_{i=1}^{n}[(z-m)^2 + \frac{1}{n-1}\sum_{i=1}^{n}2(z-m)(m-A_i) + \frac{1}{n-1}\sum_{i=1}^{n}(m-A_i)^2] \\
=&E[\frac{1}{n-1}\sum_{i=1}^{n}[(z-m)^2]] + E[\frac{1}{n-1}\sum_{i=1}^{n}2(z-m)(m-A_i)] + E[\frac{1}{n-1}\sum_{i=1}^{n}(m-A_i)^2]
\end{aligned}$$

since $m = \frac{1}{n}\sum_{i=1}^{n} A_i$ and $E[\frac{1}{n-1}\sum_{i=1}^{n}(m-A_i)^2] = \sigma^2$, this can be simplified to

$$\sigma^2 + E[\frac{1}{n-1}\sum_{i=1}^{n}[z^2 - 2zm + m^2]] + E[2(z-m)\frac{1}{n-1}\sum_{i=1}^{n}(m-A_i)]$$

Due to the definition of m, $\sum_{i=1}^{n}(m-A_i)] = 0$. Thus, what remains is

$$\begin{aligned}
&\sigma^2 + E[\frac{1}{n-1}\sum_{i=1}^{n}[z^2 - 2zm + m^2]]\\
=&\sigma^2 + E[\frac{n}{n-1}[z^2 - 2zm + m^2]]\\
=&\sigma^2 + \frac{n}{n-1}[E[z^2] - 2zE[m] + E[m^2]]\\
=&\sigma^2 + \frac{n}{n-1}[z^2 - 2z\mu + Var[m^2] + E[m]^2]\\
=&\sigma^2 + \frac{n}{n-1}[z^2 - 2z\mu + \frac{\sigma^2}{n} + \mu^2]
\end{aligned}$$

which finally can be simplified to

$$\frac{n}{n-1}\left(\sigma^2 + (z-\mu)^2\right) \tag{A.4}$$

Hence, comparing both errors, the RENA overestimates the true uncertainty by a factor $\frac{n}{n-1}$, proving Proposition 3. In other words, he suffers from *underprecision*.

A.3 Proposition 4: Two sample case

The statistically "correct" prediction error

The complete proof strongly resembles the one from the section before. The statistically correct calculation of the prediction error of $m_A - m_B$ in this case is

$$E\{[(m_A - m_B) - (X - Y)]^2\} \tag{A.5}$$

because $E[m_A] = \mu_x$ and $E[m_B] = \mu_y$ and hence $\{E[(m_A - m_B) - (X - Y)]\}^2 = 0$.
Expression (B.1) can be rewritten as

$$E[(m_A - m_B)^2] - 2E[(m_A - m_B)(X - Y)] + E[(X - Y)^2]$$

$$\Leftrightarrow Var(m_A - m_B) + E[m_A - m_B]^2 - 2E[m_A X + m_B Y - m_B X - m_A Y] + Var(X - Y) + E[X - Y]^2$$

It holds that $Var(X - Y) = Var(X) + Var(Y) - 2\sigma_{X,Y}$.
Further, $Var(m_A - m_B) = Var(m_A) + Var(m_B)$ since $Cov(m_A, m_B) = 0$, because both estimates come from independent samples.
As a result, the sophisticated prediction error is equal to

$$[\sigma_X^2 + \sigma_Y^2]\,\frac{(n+1)}{n} - 2\sigma_{X,Y}$$

or for $\sigma_X^2 = \sigma_Y^2$:

$$[\sigma_X^2]\,2\frac{(n+1)}{n} - 2\sigma_{X,Y} \tag{A.6}$$

The RENAs prediction error

In contrast, the RENAs again (implicitly) calculates only the sample-dispersion around his mean-differene estimate. Hence, on average:

$$E\left[\frac{1}{n-1}\sum[(m_A - m_B) - (A_i - B_i)]^2\right] \tag{A.7}$$

Following a similar proof as in the earlier section, this is equal to

$$\sigma_X^2 + \sigma_Y^2 - 2\sigma_{X,Y} \tag{A.8}$$

or simply $Var(X - Y)$.
The direct comparison shows that the RENA underestimates future variation in sample means by a factor of $2\frac{n+1}{n}$ on average.

A.4 Proposition 5: Predictions via Regression Analysis

The proof follows the same logic as the one presented in Chapter 11 in von Auer (2005). Let $\hat{y}_{t+1}$ be the forecast of the unknown outcome y_{t+1} based on OLS-parameter estimates $\hat{\alpha}$ and $\hat{\beta}$ as well as the known predictor x_{t+1}. Mathematically, $\hat{y}_{t+1} = \hat{\beta}x_{t+1}+\hat{\alpha}$ and $y_{t+1} = \alpha+\beta x_{t+1}+\epsilon_{t+1}$. The true prediction error can be calculated by

$$Var(y_{t+1} - \hat{y}_{t+1}) = E\left[[(y_{t+1} - \hat{y}_{t+1}) - E(y_{t+1} - \hat{y}_{t+1})]^2\right] \tag{A.9}$$

Because $E[\epsilon_{t+1}] = 0$, $E[\hat{\alpha}] = \alpha$ and $E[\hat{\beta}] = \beta$, (A.9) can be simplified to

$$Var(y_{t+1} - \hat{y}_{t+1}) = E\left[(y_{t+1} - \hat{y}_{t+1})^2\right] \tag{A.10}$$

And thus,

$$\begin{aligned} & E\left[[x_{t+1}(\beta - \hat{\beta}) + (\alpha - \hat{\alpha}) + \epsilon_{t+1}]^2\right] \\ \Leftrightarrow & E\left[[x_{t+1}(\beta - \hat{\beta}) + (\alpha - \hat{\alpha})]^2 - 2\epsilon_{t+1}(\alpha - \hat{\alpha})] - 2\epsilon_{t+1}(\beta - \hat{\beta}) + \epsilon_{t+1}^2\right] \\ \Leftrightarrow & E\left[[x_{t+1}(\beta - \hat{\beta}) + (\alpha - \hat{\alpha})]^2\right] - 2E[\epsilon_{t+1}(\alpha - \hat{\alpha})] - 2E[\epsilon_{t+1}(\beta - \hat{\beta})] + E[\epsilon_{t+1}^2] \end{aligned}$$

Because the sample-parameters as well as the true parameters are independent of the future random error as well as $E[[\epsilon_{t+1}] = 0$, it must hold that $E[\epsilon_{t+1}(\alpha - \hat{\alpha})] = 0$ and $E[\epsilon_{t+1}(\beta - \hat{\beta})] = 0$. Further, $E[\epsilon_{t+1}^2] = Var(\epsilon_{t+1}) + E[\epsilon_{t+1}]^2 = Var(\epsilon_{t+1}) = \sigma^2$
Hence, the expression above boils down to

$$\sigma^2 + E\left[[x_{t+1}(\beta - \hat{\beta}) + (\alpha - \hat{\alpha})]^2\right]$$

which is equivalent to

$$\begin{aligned} & \sigma^2 + E\left[[x_{t+1}(\beta - \hat{\beta})]^2 + (\alpha - \hat{\alpha})^2 + 2x_{t+1}(\beta - \hat{\beta})(\alpha - \hat{\alpha})\right] \\ \Leftrightarrow & \sigma^2 + x_{t+1}^2 E[(\beta - \hat{\beta})^2] + E[(\alpha - \hat{\alpha})^2] + 2x_{t+1}E[(\beta - \hat{\beta})(\alpha - \hat{\alpha})] \end{aligned}$$

With $E[(\alpha-\hat{\alpha})^2] = Var(\hat{\alpha})$ and $E[(\beta-\hat{\beta})^2] = Var(\hat{\beta})$, it transforms to

$$\sigma^2 + x_{t+1}^2 Var(\hat{\beta}) + Var(\hat{\alpha}) + 2x_{t+1}E[(\beta-\hat{\beta})(\alpha-\hat{\alpha})] \tag{A.11}$$

which leaves the last part to be calculated explicitly.

$$\begin{aligned}
&E[(\beta-\hat{\beta})(\alpha-\hat{\alpha})]\\
\Leftrightarrow &E[\alpha\beta-\hat{\beta}\alpha-\hat{\alpha}\beta+\hat{\beta}\hat{\alpha}]\\
\Leftrightarrow &-\alpha\beta+E[\hat{\beta}\hat{\alpha}]
\end{aligned}$$

with $\hat{\alpha}$ being the OLS estimator as thus $\hat{\alpha} = \bar{y}-\hat{\beta}\bar{x}$:

$$\begin{aligned}
&-\alpha\beta+E[\hat{\beta}(\bar{y}-\hat{\beta}\bar{x})]\\
\Leftrightarrow &-\alpha\beta+\bar{y}E[\hat{\beta}]-\bar{x}E[\hat{\beta}^2]\\
\Leftrightarrow &-\alpha\beta+\bar{y}\beta-\bar{x}[Var(\hat{\beta})+\beta^2]\\
\Leftrightarrow &-\alpha\beta+\beta(\bar{y}-\beta\bar{x})-\bar{x}[Var(\hat{\beta})]\\
\Leftrightarrow &-\alpha\beta+\beta\alpha-\bar{x}[Var(\hat{\beta})]\\
\Leftrightarrow &-\bar{x}[Var(\hat{\beta})]
\end{aligned}$$

Plugging this into expression (A.11), I get

$$\sigma^2 + x_{t+1}^2 Var(\hat{\beta}) + Var(\hat{\alpha}) - 2x_{t+1}\bar{x}[Var(\hat{\beta})]$$

with $Var(\hat{\alpha}) = \sigma^2\left(\frac{1}{T}+\frac{\bar{x}^2}{\sum_{t=1}^{T}(x_t-\bar{x})^2}\right)$ and $Var(\hat{\beta}) = \frac{\sigma^2}{\sum_{t=1}^{T}(x_t-\bar{x})^2}$, we get finally get

$$\sigma^2\left[1+\frac{1}{T}+\frac{(x_{t+1}-\bar{x})^2}{\sum_{t=1}^{T}(x_t-\bar{x})^2}\right] \tag{A.12}$$

which concludes the proof.

Appendix B

Autonomy Support and Innovation

B.1 Low future benefits

Pattern I: Conditions

For Pattern I with investments in each period, $s^*_{a,1} > 0$ and $s^*_{a,2} > 0$, given by Equations (2.6) and (2.7), must hold. For this to be the case, neither can $\frac{v_P}{\alpha}$ be too low (which holds by assumption $\frac{v_P}{\alpha} > 2$), nor can $\bar{s}$ be too high, as is stated by condition in Equation (2.11), or too low, as is stated in Equation (2.12). Also, the optimal $s^*_{a,1}$ in Equation (2.6) cannot be too high, otherwise the principal will not find it optimal to invest in period $t = 2$ as well. For $s^*_{a,1}$ (Equation (2.7)) not to exceed the threshold level in period $t = 2$, the following condition must hold:

$$\left(\frac{v_P}{2\alpha(1-\beta\delta)}\right)^{\frac{2}{3}} - \delta\,\bar{s} < \frac{1}{\delta}\left(\frac{v_P}{2\alpha}\right)^{\frac{2}{3}} - \delta\,\bar{s} \tag{B.1}$$

where the RHS is derived by the inequality of $s^*_{a,2} > 0$. Simplification leads to

$$\frac{1}{\delta} - \sqrt{\delta} > \beta$$

which is the necessary condition for Pattern I with investments in each period to exist.

The threshold condition in Equation (2.11) follows directly from $s^*_{a,1} > 0$ in Equation (2.7). The threshold condition in Equation (2.12) follows directly from plugging the optimal investment levels $s^*_{a,1} > 0$ and $s^*_{a,2} > 0$ into the principal's time-interdependent profit function (Equation (2.3)) and solving for the $\bar{s}$ guaranteeing the principal non zero profits for her investments.

Pattern II: Conditions

An investment in only the second period requires $s^*_{a,1} = 0$. Following Equation (2.7), this implies $\frac{1}{\delta}\left(\frac{v_P}{\alpha}\frac{1}{2(1-\beta\delta)}\right)^{\frac{2}{3}} \geq \bar{s}^i$. Further, it requires $s^*_{a,2} > 0$, implying that the condition in Equation (2.13) must hold.

Hence, Pattern II only exists if $\bar{s} \in \left[\frac{1}{\delta}\left(\frac{v_P}{\alpha}\frac{1}{2(1-\beta\delta)}\right)^{\frac{2}{3}}, \frac{1}{\delta^2}\left(\frac{1}{2}\right)^{\frac{2}{3}}\left(\frac{v_P}{\alpha}\right)^{\frac{2}{3}}\right)$. This set in only non-empty if

$$\frac{1}{\delta}\left(\frac{v_P}{\alpha}\frac{1}{2(1-\beta\delta)}\right)^{\frac{2}{3}} < \frac{1}{\delta^2}\left(\frac{1}{2}\right)^{\frac{2}{3}}\left(\frac{v_P}{\alpha}\right)^{\frac{2}{3}}$$

Simplifying this inequality leads to the same condition as above:

$$\frac{1}{\delta} - \sqrt{\delta} > \beta$$

Pattern III: Conditions

If $\bar{s} > \frac{1}{\delta^2}\left(\frac{1}{2}\right)^{\frac{2}{3}}\left(\frac{v_P}{\alpha}\right)^{\frac{2}{3}}$, Equation (2.6) states that $s^*_{a,2} = 0$ is optimal for the principal. Further, since $\frac{1}{\delta} - \sqrt{\delta} > \beta$, this fulfills the condition in Equation (2.14) and $\bar{s}$ lies above the threshold level ensuring $s^*_{a,1} = 0$. Hence, under this condition the principal will not invest in any period.

B.2 High future benefits

Pattern IV: Conditions

When future benefits are high for the principal, she wants to invest a positive amount in $t = 1$. With $\frac{1}{\delta} - \sqrt{\delta} \leq \beta$, $s_{a,1}$ is such that it exceeds Equation (B.1) and there are no investments are made in $t = 2$. With $s^*_{a,2} = 0$ the optimization problem of the principal in $t = 1$ becomes

$$max_{s_{a,1}} \frac{(\delta\bar{s} + s_{a,1})^{\frac{1}{2}} - 1}{(\delta\bar{s} + s_{a,1})^{\frac{1}{2}}} v_P - \alpha s_{a,1} + \beta \left[v_P - \frac{v_P}{(\delta^2\bar{s} + \delta s_{a,1})^{\frac{1}{2}}} \right] \tag{B.2}$$

The resulting optimal level of $s_{a,1}$ is then:

$$s^*_{a,1} = \left[\frac{v_P}{2\alpha} \left(1 + \frac{\beta}{\sqrt{\delta}} \right) \right]^{\frac{2}{3}} - \delta\bar{s} \tag{B.3}$$

We check our result by plugging $s^*_{a,1}$ into the threshold level in period $t = 2$

$$s^*_{a,1} \geq \frac{1}{\delta} \left(\frac{v_P}{2\alpha} \right)^{\frac{2}{3}} - \delta\bar{s}$$

which simplifies to $\delta + \sqrt{\delta}\beta \geq 1$ or $\beta \geq \frac{1}{\sqrt{\delta}} - \sqrt{\delta}$, respectively. This holds as $\frac{1}{\delta} - \sqrt{\delta} \leq \beta$ holds and $\frac{1}{\delta} \geq \frac{1}{\sqrt{\delta}}$ is true for $\delta \in (0, 1)$.

The condition in Equation (2.15) follows directly from equation (B.3). The condition in Equation (2.16) follows directly from plugging the optimal investment levels $s^*_{a,1} > 0$ and $s^*_{a,2} = 0$ into the principals time-interdependent profit function (Equation (2.3)) and solving for the $\bar{s}$ guaranteeing the principal nonzero profits for her investment. These conditions therefore describe Pattern IV.

B.3 Model version with complements

Model setup

We show that except for the u-shaped Pattern II and the no investment Pattern V, the investment patterns in the main section emerge also under the assumption of complementarity. We introduce complementarity by assuming that the agent's initial level and the investments by the principal multiply, implying that the principal only invests if the agent arrives with at least a minimal positive $\bar{s}$. The total value of Autonomy Support $\bar{s}_1$ and $\bar{s}_2$ available to the agent in periods $t = 1, 2$ becomes

$$\bar{s}_1 = \bar{s} \cdot s_{a,1}$$
$$\bar{s}_2 = \bar{s} \cdot \delta s_{a,1} + \bar{s} \cdot s_{a,2}$$

For the sake of simplicity, we assume that only Autonomy support provided by the principal is discounted. The agent's utility as a function of his innovative activity i for periods $t = 1, 2$ are

$$U_A(i,1) = v_A \frac{i}{1+i} - \frac{i}{\bar{s} \cdot s_{a,1}} \qquad = v_A \frac{i}{1+i} - \frac{i}{\bar{s}_1}$$
$$U_A(i,2) = v_A \frac{i}{1+i} - \frac{i}{\bar{s} \cdot (\delta s_{a,1} + s_{a,2})} \qquad = v_A \frac{i}{1+i} - \frac{i}{\bar{s}_2}$$

The principal's profit functions for periods $t = 1, 2$ are

$$\Pi_P(s_{a,1}, 1) = v_P \frac{i}{1+i} - \alpha s_{a,1} + \beta \Pi_P(s_{a,2}, 2)$$
$$\Pi_P(s_{a,2}, 2) = v_P \frac{i}{1+i} - \alpha s_{a,2}$$

We solve the model by Backward Induction.

Solving the model

Period 2

The agent chooses the optimal level of i to maximize his utility $U_A(i,2)$, resulting in

$$i_2^* = (\bar{s}(\delta s_{a,1} + s_{a,2}))^{\frac{1}{2}} - 1$$
$$i_2^* \geq 0$$

Innovative activity in period $t = 2$, i_2^*, is only greater than zero if $s_{a,2} \geq 1 - \frac{1}{\bar{s}} - \delta s_{a,1}$.
If $\frac{1}{\bar{s}} + \delta s_{a,1}$ is large enough, the principal does not have to provide additional Autonomy Support in the second period to induce innovative effort by the agent. If investments in the first period have vanished such that $\delta s_{a,1} = 0$, then $s_{a,2}$ must be high enough to instill innovation in the second period.

The principal maximizes $\Pi_P(s_{a,2}, 2)$, which with the agent's choice becomes

$$max_{s_{a,2}} \frac{\bar{s}(\delta s_{a,1} + s_{a,2}))^{\frac{1}{2}} - 1}{(\bar{s}(\delta s_{a,1} + s_{a,2}))^{\frac{1}{2}}} v_P - \alpha s_{a,2}$$

in order to derive her optimal investment in the second period by taking the agent's innovation effort into account, which results in

$$s_{a,2}^* = \left(\frac{1}{\bar{s}}\right)^{\frac{1}{3}} \left(\frac{v_P}{2\alpha}\right)^{\frac{2}{3}} - \delta s_{a,1} \geq 0 \tag{B.4}$$

Similar to the main section, the optimal value of $s_{a,2}^*$ decreases in α and increases in v_P. The principal only invests in Autonomy Support in the second period if her investment in the first does not exceed the threshold $\frac{1}{\delta}\left(\frac{1}{\bar{s}}\right)^{\frac{1}{3}} \frac{v_P}{2\alpha}^{\frac{2}{3}}$ making additional investments unnecessary.

Period 1

Optimizing $U_A(i,1)$ leads to the agent's optimal choice of innovative activity in period $t = 1$:

$$i_1^* = (\bar{s} \cdot s_{a,1})^{\frac{1}{2}} - 1 \geq 0$$

The principal takes i_1^* as well as the discounted future consequences of her choice of $s_{a,1}$ into account. She maximizes $\Pi_P(s_{a,1},1)$ which has become

$$max_{s_{a,1}} \frac{(\bar{s} \cdot s_{a,1})^{\frac{1}{2}} - 1}{(\bar{s} \cdot s_{a,1})^{\frac{1}{2}}} v_P - \alpha s_{a,1} + \beta \bar{\Pi}_2 + \alpha \beta \delta s_{a,1}$$

with $\bar{\Pi}_2 = v_P - \frac{v_P}{\bar{s}^{\frac{1}{3}} \left(\frac{v_P}{2\alpha}\right)^{\frac{1}{3}}} - \alpha \frac{1}{\bar{s}^{\frac{1}{3}}} \left(\frac{v_P}{2\alpha}\right)^{\frac{1}{3}}$ as a constant, and optimally chooses

$$s_{a,1}^* = \left(\frac{v_P}{2\alpha(1-\beta\delta)}\right)^{\frac{2}{3}} \frac{1}{\bar{s}}^{\frac{1}{3}} \geq 0 \tag{B.5}$$

where $(\alpha - \beta\delta\alpha) > 0$, because $\delta, \beta < 1$.

Results: Autonomy Support investment patterns

As in the main section, the investment patterns that emerge depend on the benefit-cost ratio $\frac{v_P}{\alpha}$, the agent's depreciation rate of Autonomy Support δ and the principal's discount factor β, and can be distinguished by the relationship between β and δ.

Low future benefit investment patterns under $\beta < \frac{1}{\delta} - \sqrt{\delta}$

When either or both β and δ are low, the principal has lower incentives to invest which results in lower investments.

Pattern A: Investment in each period

When the principal wants to lower investments, she may have to invest in each period. For this, $s^*_{a,1}$ must be smaller than the threshold level $\frac{1}{\delta}\left(\frac{1}{\bar{s}}\right)^{\frac{1}{3}} \frac{v_P}{2\alpha}^{\frac{2}{3}}$ such that $s^*_{a,2} > 0$ is optimal. It follows that

$$\begin{aligned} s^*_{a,2} &= \left(\frac{1}{\bar{s}}\right)^{\frac{1}{3}} \left(\frac{v_P}{2\alpha}\right)^{\frac{2}{3}} - \delta\left(\frac{v_P}{2\alpha(1-\beta\delta)}\right)^{\frac{2}{3}} \frac{1}{\bar{s}}^{\frac{1}{3}} \\ &= \left(\frac{1}{\bar{s}}\right)^{\frac{1}{3}} \left(\frac{v_P}{2\alpha}\right)^{\frac{2}{3}} \left[1-\delta\frac{1}{(1-\beta\delta)}^{\frac{2}{3}}\right] > 0 \end{aligned}$$

Since $\bar{s}$ and $\left(\frac{v_P}{2\alpha}\right)$ are strictly positive, this expression holds for as long as $\beta < \frac{1}{\delta} - \sqrt{\delta}$ which is the same necessary condition as in the main section with substitutability. However, there is no explicit upper threshold $\bar{s}$ that renders $s^*_{a,1}$ to zero. Therefore, there is no u-shaped pattern with complementarity.

Inserting $s^*_{a,1}$ and $s^*_{a,2}$ (from Equations (B.5) and (B.4)) into the profit function we derive the principal's payoff

$$\Pi^*_1 = v_P - \frac{v_P}{\frac{v_p}{(2\alpha(1-\beta\delta)}^{\frac{2}{3}}} - \alpha\left(\frac{v_P}{2\alpha(1-\beta\delta)}\right)^{\frac{2}{3}} \frac{1}{\bar{s}}^{\frac{1}{3}} + \beta\delta\alpha\left(\frac{v_P}{2\alpha(1-\beta\delta)}\right)^{\frac{2}{3}} \frac{1}{\bar{s}}^{\frac{1}{3}} + \beta\left[v_P - 3\alpha\left(\frac{1}{\bar{s}}\right)^{\frac{1}{3}}\left(\frac{v_P}{2\alpha}\right)^{\frac{2}{3}}\right]$$

which we use to derive the sufficient condition that guarantees positive profits for the principal

$$\bar{s} \geq \frac{9}{4}\frac{\alpha}{v_P}\left[\frac{(1-\beta\delta)^{\frac{2}{3}}+\beta}{(1+\beta)}\right]^3$$

The higher $\frac{v_P}{\alpha}$, the lower $\bar{s}$ can be without setting $s^*_{a,1} = 0$. The threshold depends negatively on δ and β, as long as $\delta, \beta \in (0,1)$, which by assumption, they are. Hence, if $\bar{s}$ is not too low the principal invests a positive amount in both periods in the low future benefit investment pattern.

Pattern B: No investment

In Pattern A, we derive that the principal does not invest in period $t = 1$, $s^*_{a,1} = 0$, when

$$\bar{s} < \frac{9}{4}\frac{\alpha}{v_P}\left[\frac{(1-\beta\delta)^{\frac{2}{3}}+\beta}{(1+\beta)}\right]^3$$

Taking this into account in Equation (B.4), the optimal investment in period $t = 2$ must satisfy $s^*_{a,2} = \left(\frac{1}{\bar{s}}\right)^{\frac{1}{3}}\left(\frac{v_P}{2\alpha}\right)^{\frac{2}{3}}$ leading to the condition for a positive investment $s^*_{a,2} > 0$

$$\bar{s} \geq \frac{9}{4}\frac{\alpha}{v_P}$$

We know that $\bar{s} < \frac{9}{4}\frac{\alpha}{v_P}\left[\frac{(1-\beta\delta)^{\frac{2}{3}}+\beta}{(1+\beta)}\right]^3$, therefore we know that $\bar{s} < \frac{9}{4}\frac{\alpha}{v_P}$ must hold as well, because $\left[\frac{(1-\beta\delta)^{\frac{2}{3}}+\beta}{(1+\beta)}\right]^3 \in [0.125, 1]$ for $\delta, \beta \in [0, 1]$.
It follows that for $\bar{s} < \frac{9}{4}\frac{\alpha}{v_P}\left[\frac{(1-\beta\delta)^{\frac{2}{3}}+\beta}{(1+\beta)}\right]^3$ the principal does not invest in either period, such that $s^*_{a,1} = s^*_{a,2} = 0$.

High future benefit investment patterns under $\beta \geq \frac{1}{\delta} - \sqrt{\delta}$

Pattern C: Investment only in first period

With high future benefits when both β and δ are high, the principal wants higher investments in the first period. For $\beta \geq \frac{1}{\delta} - \sqrt{\delta}$,

$$s^*_{a,2} = \left(\frac{1}{\bar{s}}\right)^{\frac{1}{3}}\left(\frac{v_P}{2\alpha}\right)^{\frac{2}{3}}\left[1-\delta\frac{1}{(1-\beta\delta)}^{\frac{2}{3}}\right] \leq 0$$

such that the principal never wants to invest in period $t = 2$ and chooses $s^*_{a,2} = 0$. Her maximization problem in period $t = 1$ becomes

$$max_{s_{a,1}} \frac{(\bar{s}\cdot s_{a,1})^{\frac{1}{2}}-1}{(\bar{s}\cdot s_{a,1})^{\frac{1}{2}}} v_P - \alpha s_{a,1} + \beta\left[v_P - \frac{v_P}{(\delta\bar{s}\cdot s_{a,1})^{\frac{1}{2}}}\right]$$

resulting in the optimal investment in the first period of

$$s^*_{a,1} = \left(\frac{1}{\bar{s}}\right)^{\frac{1}{3}} \left(\frac{v_P}{2\alpha}\right)^{\frac{2}{3}} \left[1 + \frac{\beta}{\sqrt{\delta}}\right]^{\frac{2}{3}} \tag{B.6}$$

This also exceeds the threshold $\frac{1}{\delta}\left(\frac{1}{\bar{s}}\right)^{\frac{1}{3}} \frac{v_P}{2\alpha}^{\frac{2}{3}}$, guaranteeing that $s^*_{a,2}$ in Pattern C is zero.

Comparison to the model version in the main section

For the low future benefits case, the principal either invests in both periods when the agent arrives with a sufficiently high initial amount of Autonomy Support, or invests in neither period if the agent does not. The u-shaped Pattern II in the main section does not emerge. Once the agent's Autonomy Support stock is too low, the principal has no incentive to top it up under complementarity.

For high future benefits, the principal always invests, but only in the second period. The no investment Pattern V in the main section does not emerge due to high future benefits. Because of the complementarity, her incentive to invest a lot 'in advance' pays off stronger in the later period.

Although the complementarity version features a different core assumption, the main results hold. First, the principal's investments are lower the higher the existing level of Autonomy Support, as $\frac{\partial s^*_{a,2}}{\partial s^*_{a,1}} < 0$ and $\frac{\partial s^*_{a,2}}{\partial \bar{s}} < 0$ for $\delta > 0$. In a way, there is a certain degree of substitutability even under the assumption of complementarity. Second, whether the principal invests in each or only one period depends on the β-δ relationship. As in the main section, this relationship is displayed by $\beta \lesseqgtr \frac{1}{\delta} + \sqrt{\delta}$ reflecting the future benefits of investments.

Appendix C

The Dry Powder Paradox of Monetary Policy

Proof for Section 4

Fiscal Dominance: Discretionary Case

The constraint on real interest expenses is given by

$$(i_t - \pi_t)B_{t-1} \leq (1+\pi_t)[\Omega_t - (1+i_t)G_t] \tag{C.1}$$

Using the definitions

$$g_t = \frac{G_t - G}{Y}$$
$$b_t = \frac{B_t - B}{Y}$$
$$\omega_t = \frac{\Omega_t - \Omega}{Y}$$

we can rewrite this condition into

$$(i_t - \pi_t)\left[b_{t-1} + \frac{B}{Y}\right] \leq (1+\pi_t)\left[\omega_t + \frac{\Omega}{Y} - (1+i_t)\left(g_t + \frac{G}{Y}\right)\right] \tag{C.2}$$

Expanding and neglecting cross terms

$$(i_t - \pi_t)\frac{B}{Y} \leq \omega_t + (1+\pi_t)\frac{\Omega}{Y} - g_t - (1+\pi_t + i_t)\frac{G}{Y} \tag{C.3}$$

multiplying by Y and collecting terms

$$i_t\,[B+G] \leq (\omega_t - g_t)Y + \pi_t\,[B+\Omega-G] + \Omega - G \tag{C.4}$$

which yields

$$i_t \leq \frac{\Omega - G}{B+G} + \frac{Y}{B+G}(\omega_t - g_t) + \frac{B+\Omega-G}{B+G}\pi_t \tag{C.5}$$

Under discretion, the equilibrium has

$$i_t = \pi(\delta) + \phi_r r_t^e + \phi_u u_t$$
$$\pi_t = \pi(\delta) + \psi_u u_t + \psi_r r_t^e$$

where, as before, $\pi(\delta) = \frac{\kappa(\lambda-\delta\sigma)}{\alpha(1-\beta)+\kappa^2}$. The approximated rollover condition after some reformulations gets

$$\iota_c + \iota_r r_t^e + \iota_u u_t + \iota_g(\omega_t + g_t) \leq \pi(\delta)\left[\frac{B+\Omega-G}{B+G} - 1\right] \equiv \bar{S} \tag{C.6}$$

where $\iota_r, \iota_u, \iota_c, \iota_g$ are coefficients independent of δ. Since $\pi'(\delta) < 0$ it must hold that

$$\frac{\partial\bar{S}}{\partial\delta} > 0 \;\Leftrightarrow\; \frac{B+\Omega-G}{B+G} < 1 \;\Leftrightarrow\; G > \frac{\Omega}{2}$$

which implies that δ diminishes the chances for encountering a fiscal dominance scenario under discretion if the government sector is sufficiently large in comparison to the rollover capacity.

If the sign of $\pi'(\delta)$ is reversed (as in the commitment case), the opposite is true.

Bibliography

Adam, K. and Billi, R. M. (2007). Discretionary monetary policy and the zero lower bound on nominal interest rates. *Journal of Monetary Economics*, 54(3):728–752.

Akgün, A. E., Keskin, H., and Byrne, J. (2009). Organizational emotional capability, product and process innovation, and firm performance: An empirical analysis. *Journal of Engineering and Technology Management*, 26(3):103–130.

Amabile, T. M., Schatzel, E. A., Moneta, G. B., and Kramer, S. J. (2004). Leader behaviors and the work environment for creativity: Perceived leader support. *The Leadership Quarterly*, 15(1):5–32.

Andreoni, J., Che, Y.-K., and Kim, J. (2007). Asymmetric information about rivals' types in standard auctions: An experiment. *Games and Economic Behavior*, 59(2):240–259.

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Armstrong, J. S. (2001). Combining forecasts. In *Principles of forecasting*, pages 417–439. Springer.

Arrow, K. J. and Kurz, M. (1970). *Public investment, the rate of return, and optimal fiscal policy*. Baltimore: Johns Hopkins Press.

Axtell, C., Holman, D., and Wall, T. (2006). Promoting innovation: A change study. *Journal of Occupational and Organizational Psychology*, 79(3):509–516.

Backhaus, K., Blechschmidt, B., and Eisenbeiß, M. (2006). Der Stichprobeneinfluss bei Kausalanalysen. *Die Betriebswirtschaft*, 66(6):711.

Bammens, Y. P. (2016). Employees' innovative behavior in social context: A closer examination of the role of organizational care. *Journal of Product Innovation Management*, 33(3):244–259.

Barberis, N. and Thaler, R. (2003). A survey of behavioral finance. *Handbook of the Economics of Finance*, 1:1053–1128.

Barhoumi, K., Cherif, R., and Rebei, N. (2016). Stochastic trends, debt sustainability and fiscal policy. IMF Working Paper 2016/059, International Monetary Fund.

Barro, R. J. and Gordon, D. B. (1983). Rules, discretion and reputation in a model of monetary policy. *Journal of Monetary Economics*, 12(1):101–121.

Bazerman, M. H. and Moore, D. A. (2012). *Judgment in managerial decision making*. John Wiley & Sons.

Benigno, P. and Woodford, M. (2003). Optimal monetary and fiscal policy: A linear-quadratic approach. *NBER macroeconomics annual*, 18:271–333.

Benita, M., Roth, G., and Deci, E. L. (2014). When are mastery goals more adaptive? It depends on experiences of autonomy support and autonomy. *Journal of Educational Psychology*, 106(1):258.

Benz, M. and Frey, B. S. (2008). Being independent is a great thing: Subjective evaluations of self-employment and hierarchy. *Economica*, 75(298):362–383.

Bernanke, B. S. and Woodford, M. (1997). Inflation forecasts and monetary policy. *Journal of Money, Credit, and Banking*, pages 653–684.

Bilsland, K. and Cumbers, A. (2018). Managerial control and the limits to employee participation in retail work spaces: Evidence from a UK IKEA store. *New Technology, Work and Employment*, 33(2):130–148.

Blanchard, O. (2004). Fiscal dominance and inflation targeting: Lessons from brazil. NBER Working Paper 10389, National Bureau of Economic Research, Inc.

Blanchard, O. and Galí, J. (2007). Real wage rigidities and the new keynesian model. *Journal of Money, Credit and Banking*, 39:35–65.

Blom, G. (1958). *Statistical estimates and transformed beta-variables*. PhD thesis, Almqvist & Wiksell.

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., and Roberts, J. (2013). Does management matter? Evidence from India. *The Quarterly Journal of Economics*, 128(1):1–51.

Bloom, N., Sadun, R., and Van Reenen, J. (2015). Do private equity owned firms have better management practices? *The American Economic Review*, 105(5):442–46.

Bloom, N. and Van Reenen, J. (2010). Why do management practices differ across firms and countries? *Journal of Economic Perspectives*, 24(1):203–24.

Bonaccio, S. and Dalal, R. S. (2006). Advice taking and decision-making: An integrative literature review, and implications for the organizational sciences. *Organizational Behavior and Human Decision Processes*, 101(2):127–151.

Bordalo, P., Coffman, K., Gennaioli, N., and Shleifer, A. (2016). Stereotypes. *The Quarterly Journal of Economics*, 131(4):1753–1794.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2017). Memory, Attention, and Choice. *The Quarterly Journal of Economics*.

Borrill, C. S., Carletta, J., Carter, A., Dawson, J. F., Garrod, S., Rees, A., Richards, A., Shapiro, D., and West, M. A. (2000). *The effectiveness of health care teams in the National Health Service*. University of Aston.

Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). *Time Series Analysis: Forecasting and Control*. John Wiley & Sons.

Brunnermeier, M. K. and Koby, Y. (2016). The reversal interest rate: An effective lower bound on monetary policy. *unpublished paper, Princeton University*.

Brzeski, C. (2020). ECB: Keeping its powder dry. Retrieved last on 2021/02/19 from https://think.ing.com/snaps/ecb-apr-rate-decision.

Bullard, J. and Mitra, K. (2002). Learning about monetary policy rules. *Journal of Monetary Economics*, 49(6):1105–1129.

Bursztyn, L., González, A. L., and Yanagizawa-Drott, D. (2018). Misperceived Social Norms: Female Labor Force Participation in Saudi Arabia. Working Paper 2018-042, Human Capital and Economic Opportunity Working Group.

Cassar, L. and Meier, S. (2018). Nonmonetary incentives and the implications of work as a source of meaning. *Journal of Economic Perspectives*, 32(3):215–38.

Cassotti, M., Habib, M., Poirel, N., Aïte, A., Houdé, O., and Moutier, S. (2012). Positive emotional context eliminates the framing effect in decision-making. *Emotion*, 12(5):926.

Charness, G. and Grieco, D. (2018). Creativity and incentives. *Journal of the European Economic Association*, 17(2):454–496.

Clarida, R., Galí, J., and Gertler, M. (1999). The Science of Monetary Policy: a New Keynesian perspective. *Journal of Economic Literature*, 37(4):1661–1707.

Corgnet, B. and Hernán González, R. (2013). Don't ask me if you will not listen: The dilemma of consultative participation. *Management Science*, 60(3):560–585.

Cox, J. C., Smith, V. L., and Walker, J. M. (1988). Theory and individual behavior of first-price auctions. *Journal of Risk and Uncertainty*, 1(1):61–99.

Crosetto, P., Filippin, A., Katuščák, P., and Smith, J. (2020). Central tendency bias in belief elicitation. *Journal of Economic Psychology*, 78:102273.

De Resende, C. and Rebei, N. (2008). The welfare implications of fiscal dominance. Staff Working Paper 08-28, Bank of Canada.

Dorenbosch, L., Engen, M. L. v., and Verhagen, M. (2005). On-the-job innovation: The impact of job design and human resource management through production ownership. *Creativity and Innovation Management*, 14(2):129–141.

Dougherty, M. R., Gettys, C. F., and Ogden, E. E. (1999). Minerva-dm: A memory processes model for judgments of likelihood. *Psychological Review*, 106(1):180.

Dufrénot, G., Jawadi, F., and Khayat, G. A. (2018). A Model of Fiscal Dominance under the "Reinhart Conjecture". *Journal of Economic Dynamics and Control*, 93:332–345.

Early, T. (2012). How to implement Kaizen. Retrieved last on 2019/09/10 from http://leanmanufacturingtools.org/625/planning-and-running-kaizen-events.

Ederer, F. and Manso, G. (2013). Is pay for performance detrimental to innovation? *Management Science*, 59(7):1496–1513.

Edgar, F. and Geare, A. (2005). HRM practice and employee attitudes: Different measures–different results. *Personnel Review*, 34(5):534–549.

Eggertsson, G. B. and Woodford, M. (2003). Optimal monetary policy in a liquidity trap. Working paper, National Bureau of Economic Research.

Ellenberg, J. (2014). *How Not to Be Wrong: The Hidden Maths of Everyday Life*. Penguin UK.

Erat, S. and Gneezy, U. (2016). Incentives for creativity. *Experimental Economics*, 19(2):269–280.

Fahrmeir, L., Heumann, C., Künstler, R., Pigeot, I., and Tutz, G. (2016). *Statistik: Der Weg zur Datenanalyse*. Springer.

Filiz-Ozbay, E. and Ozbay, E. Y. (2007). Auctions with anticipated regret: Theory and experiment. *The American Economic Review*, 97(4):1407–1418.

Fischer, S. (1995). Central-bank independence revisited. *The American Economic Review*, 85(2):201–206.

Fischer, S. (2016). Monetary policy, financial stability, and the zero lower bound. *The American Economic Review*, 106(5):39–42.

Folkes, V. S. (1988). The availability heuristic and perceived risk. *Journal of Consumer Research*, 15(1):13–23.

Gagné, M. and Deci, E. L. (2005). Self-determination theory and work motivation. *Journal of Organizational Behavior*, 26(4):331–362.

Galí, J. (2015). *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications*. Princeton University Press.

Galí, J. (2018). The state of new keynesian economics: a partial assessment. *Journal of Economic Perspectives*, 32(3):87–112.

Galton, F. (1886). Regression towards mediocrity in hereditary stature. *The Journal of the Anthropological Institute of Great Britain and Ireland*, 15:246–263.

Geisser, S. (1993). *Predictive Inference: An Introduction*. Chapman and Hall, London.

Gennaioli, N. and Shleifer, A. (2010). What comes to mind. *The Quarterly Journal of Economics*, 125(4):1399–1433.

Gharib, S. (2019). Novartis CEO: It's time to "pivot towards transformational innovation". Retrived last on 2019/09/10 from https://fortune.com/2019/03/20/novartis-ceo-vasant-narasimhan-on-leading.

Gibbs, M., Neckermann, S., and Siemroth, C. (2017). A field experiment in motivating employee ideas. *Review of Economics and Statistics*, 99(4):577–590.

Gilovich, T., Griffin, D., and Kahneman, D. (2002). *Heuristics and Biases: The Psychology of Intuitive Judgment*. Cambridge university press.

Gilovich, T., Vallone, R., and Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive psychology*, 17(3):295–314.

Gino, F. and Schweitzer, M. E. (2008). Blinded by anger or feeling the love: how emotions influence advice taking. *Journal of Applied Psychology*, 93(5):1165.

Goodman, D. (2019). Bank of england keeps powder dry as brexit moves into next stage. Retrieved last on 2021/02/19 from https://www.bloomberg.com/news/articles/2019-12-19/boe-keeps-powder-dry-as-brexit-moves-into-next-stages.

Gronau, A. I. (2020). *Autonomy in the Workplace: Conceptualization and Application of a Social Psychological Construct in Organizational Economics*. PhD thesis, Universitätsbibliothek der Ludwig-Maximilians-Universität.

Hagger, M. S., Chatzisarantis, N. L., Hein, V., Pihu, M., Soós, I., and Karsai, I. (2007). The perceived autonomy support scale for exercise settings (passes): Development, validity, and cross-cultural invariance in young people. *Psychology of Sport and Exercise*, 8(5):632–653.

Hagger, M. S., Koch, S., and Chatzisarantis, N. L. (2015). The effect of causality orientations and positive competence-enhancing feedback on intrinsic motivation: A test of additive and interactive effects. *Personality and Individual Differences*, 72:107–111.

Hahn, G. J. (1977). A prediction interval on the difference between two future sample means and its application to a claim of product superiority. *Technometrics*, 19(2):131–134.

Harrison, G. W. (1989). Theory and misbehavior of first-price auctions. *The American Economic Review*, pages 749–762.

Harvey, N. and Fischer, I. (1997). Taking advice: Accepting help, improving judgment, and sharing responsibility. *Organizational Behavior and Human Decision Processes*, 70(2):117–133.

Helper, S. and Henderson, R. (2014). Management practices, relational contracts, and the decline of General Motors. *Journal of Economic Perspectives*, 28(1):49–72.

Hermalin, B. E. (1998). Toward an economic theory of leadership: Leading by example. *The American Economic Review*, 88(5):1188–1206.

Ichniowski, C. and Shaw, K. (1999). The effects of human resource management systems on economic performance: An international comparison of US and Japanese plants. *Management Science*, 45(5):704–721.

Ichniowski, C. and Shaw, K. (2003). Beyond incentive pay: Insiders' estimates of the value of complementary human resource management practices. *Journal of Economic Perspectives*, 17(1):155–180.

Ichniowski, C., Shaw, K., and Prennushi, G. (1995). The effects of human resource management practices on productivity. Working Paper 5333, National Bureau of Economic Research.

Illing, G. (1997). *Theorie der Geldpolitik: Eine spieltheoretische Einführung*. Springer.

Imbens, G. W. and Rubin, D. B. (2015). *Causal Inference in Statistics, Social, and Biomedical Sciences.* Cambridge University Press.

Ivanova-Stenzel, R. and Sonsino, D. (2004). Comparative study of one-bid versus two-bid auctions. *Journal of Economic Behavior & Organization*, 54(4):561–583.

Janssen, O., Van de Vliert, E., and West, M. (2004). The bright and dark sides of individual and group innovation: A special issue introduction. *Journal of Organizational Behavior*, 25(2):129–145.

Jayachandran, T. (1983). A note on a Prediction Interval for a first-order gauss markov process. *Naval Research Logistics Quarterly*, 30(3):505–508.

Jensen, J. L. W. V. (1906). Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta Mathematica*, 30:175–193.

Jonsson, G. and Klein, P. (1996). Stochastic fiscal policy and the swedish business cycle. *Journal of Monetary Economics*, 38(2):245–268.

Jung, T., Teranishi, Y., and Watanabe, T. (2005). Optimal monetary policy at the zero-interest-rate bound. *Journal of Money, Credit, and Banking*, 37(5):813–835.

Kadous, K., Leiby, J., and Peecher, M. E. (2013). How do auditors weight informal contrary advice? the joint influence of advisor social bond and advice justifiability. *The Accounting Review*, 88(6):2061–2087.

Kaesler, M., Welsh, M., and Semmler, C. (2016). Predicting overprecision in range estimation. In *Proceedings of the 38th Annual Meeting of the Cognitive Science Society*, pages 502–507. Austin, TX: Cognitive Science Society.

Kahana, M. J. (2014). *Foundations of Human Memory.* Oxford University Press.

Kahneman, D. (2011). *Thinking, Fast and Slow.* Farrar, Straus and Giroux, New York.

Kahneman, D., Slovic, S. P., Slovic, P., and Tversky, A. (1982). *Judgment under Uncertainty: Heuristics and Biases.* Cambridge university press.

Kareev, Y., Arnon, S., and Horwitz-Zeliger, R. (2002). On the misperception of variability. *Journal of Experimental Psychology: General*, 131(2):287.

Kark, R. and Carmeli, A. (2009). Alive and creating: The mediating role of vitality and aliveness in the relationship between psychological safety and creative work involvement. *Journal of Organizational Behavior*, 30(6):785–804.

Kasberger, B. (2020). An equilibrium model of the first-price auction with strategic uncertainty: Theory and empirics. Working paper, University of Oxford.

Keynes, J. M. (1938). Letter to Roy Harrod, 4 July. *The General Theory and After: Part II. Defence and Development. The Collected Writings of John Maynard Keynes*, 14.

King, R. G. and Kerr, W. (1996). Limits on interest rate rules in the is model. *FRB Richmond Economic Quarterly*, 82(2):47–75.

Kirchkamp, O. and Reiss, J. P. (2004). The overbidding-myth and the underbidding-bias in first-price auctions. Working paper, Sonderforschungsbereich 504, Universität Mannheim.

Konovalova, E. and Le Mens, G. (2018). Learning variability from experience. In *Proceedings of the 40th Annual Conference of the Cognitive Science Society*, pages 1942–1947. Austin, TX: Cognitive Science Society.

Krishna, V. (2009). *Auction Theory*. Academic press.

Kumhof, M., Nunes, R., and Yakadina, I. (2010). Simple monetary rules under fiscal dominance. *Journal of Money, Credit and Banking*, 42(1):63–92.

Kydland, F. E. and Prescott, E. C. (1977). Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy*, 85(3):473–491.

Laske, K. and Schröder, M. (2017). Quantity, quality and originality: The effects of incentives on creativity. Annual Conference 2017 (Vienna): Alternative Structures for Money and Banking 168151, Verein für Socialpolitik / German Economic Association.

Leith, C. and Wren-Lewis, S. (2000). Interactions between monetary and fiscal policy rules. *The Economic Journal*, 110(462):C93–C108.

Leith, C. and Wren-Lewis, S. (2013). Fiscal sustainability in a new keynesian model. *Journal of Money, Credit and Banking*, 45(8):1477–1516.

Levitt, S. D. and Neckermann, S. (2014). What field experiments have and have not taught us about managing workers. *Oxford Review of Economic Policy*, 30(4):639–657.

Logg, J. M., Haran, U., and Moore, D. A. (2018). Is overconfidence a motivated bias? experimental evidence. *Journal of Experimental Psychology: General*, 147(10):1445.

Mageau, G. A., Ranger, F., Joussemet, M., Koestner, R., Moreau, E., and Forest, J. (2015). Validation of the perceived parental autonomy support scale (p-pass). *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 47(3):251.

Malmendier, U. and Tate, G. (2008). Who makes acquisitions? ceo overconfidence and the market's reaction. *Journal of Financial Economics*, 89(1):20–43.

Malthus, T. R. (1798). *An Essay on the Principle of Population.* History of Economic Thought Books. McMaster University Archive for the History of Economic Thought.

Mannes, A. E. (2009). Are we wise about the wisdom of crowds? the use of group judgments in belief revision. *Management Science*, 55(8):1267–1279.

Manso, G. (2011). Motivating innovation. *The Journal of Finance*, 66(5):1823–1860.

McCallum, B. T. and Nelson, E. (1997). An optimizing is-lm specification for monetary policy and business cycle analysis. Working paper, National bureau of economic research.

McLeay, M. and Tenreyro, S. (2020). Optimal inflation and the identification of the phillips curve. *NBER Macroeconomics Annual*, 34(1):199–255.

Milgrom, P. R. and Roberts, J. D. (1992). *Economics, Organization and Management.* Englewood Cliffs, N.J: Prentice-Hall.

Moore, D. A. and Schatz, D. (2017). The three faces of overconfidence. *Social and Personality Psychology Compass*, 11(8):e12331.

Moore, D. A., Tenney, E. R., and Haran, U. (2015). Overprecision in judgment. *The Wiley Blackwell Handbook of Judgment and Decision Making*, 2:182–209.

Mullainathan, S. (2002). A memory-based model of bounded rationality. *The Quarterly Journal of Economics*, 117(3):735–774.

Mumford, M. D., Scott, G. M., Gaddis, B., and Strange, J. M. (2002). Leading creative people: Orchestrating expertise and relationships. *The Leadership Quarterly*, 13(6):705–750.

Nakata, T. and Schmidt, S. (2019). Gradualism and liquidity traps. *Review of Economic Dynamics*, 31:182–199.

Nakov, A. (2008). Optimal and simple monetary policy rules with zero floor on the nominal interest rate. *International Journal of Central Banking*.

Nilsson, H., Olsson, H., and Juslin, P. (2005). The cognitive substrate of subjective probability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(4):600.

Odean, T. (1999). Do investors trade too much? *The American Economic Review*, 89(5):1279–1298.

Ohly, S., Sonnentag, S., and Pluntke, F. (2006). Routinization, work characteristics and their relationships with creative and proactive behaviors. *Journal of Organizational Behavior*, 27(3):257–279.

Oldham, G. R. and Cummings, A. (1996). Employee creativity: Personal and contextual factors at work. *The Academy of Management Journal*, 39(3):607–634.

Orphanides, A. (2020). The fiscal-monetary policy mix in the euro area: challenges at the zero lower bound. *Economic Policy*.

Ortoleva, P. and Snowberg, E. (2015). Overconfidence in political behavior. *The American Economic Review*, 105(2):504–35.

Paese, P. W., Bieser, M., and Tubbs, M. E. (1993). Framing effects and choice shifts in group decision making. *Organizational Behavior and Human Decision Processes*, 56(1):149–165.

Persson, T. and Tabellini, G. (1993). Designing institutions for monetary stability. *Carnegie-Rochester Conference Series on Public Policy*, 39 (1):53–84.

Pisano, G. P. (2019). The hard truth about innovative cultures. Retrieved last on 2019/09/10 from https://hbr.org/2019/01/the-hard-truth-about-innovative-cultures.

Poulson, O., Holst, E., Christensen, J., et al. (1997). Calculation and application of coverage intervals for biological reference values. *Pure Appl. Chem*, 69(7):1601–1612.

Quality-One International (2015). How to implement Kaizen. Retrieved last on 2019/09/10 from https://quality-one.com/kaizen/#How.

Rabin, M. (2002). Inference by believers in the law of small numbers. *The Quarterly Journal of Economics*, 117(3):775–816.

Rabin, M. and Vayanos, D. (2010). The gambler's and hot-hand fallacies: Theory and applications. *The Review of Economic Studies*, 77(2):730–778.

Rammer, C., Czarnitzki, D., and Spielkamp, A. (2009). Innovation success of non-r&d-performers: Substituting technology by management in SMEs. *Small Business Economics*, 33(1):35–58.

Rao, C. R. (2009). *Linear Statistical Inference and its Applications, volume 22*. John Wiley & Sons.

Rapoport, A. and Budescu, D. V. (1997). Randomization in individual choice behavior. *Psychological Review*, 104(3):603.

Reis, R. (2013). Central bank design. *Journal of Economic Perspectives*, 27(4):17–44.

Rogoff, K. (1985). The optimal degree of commitment to an intermediate monetary target. *The Quarterly Journal of Economics*, 100(4):1169–1189.

Rostagno, M., Altavilla, C., Carboni, G., Lemke, W., Motto, R., Saint Guilhem, A., and Yiangou, J. (2019). A tale of two decades: The ecb's monetary policy at 20. ECB working paper 2346, European Central Bank.

Rotemberg, J. J. and Woodford, M. (1997). An optimization-based econometric framework for the evaluation of monetary policy. *NBER macroeconomics annual*, 12:297–346.

Rubinstein, A. (2012). *Lecture Notes in Microeconomic Theory: The Economic Agent*. Princeton University Press.

Shalley, C. E., Gilson, L. L., and Blum, T. C. (2009). Interactive effects of growth need strength, work context, and job complexity on self-reported creative performance. *Academy of Management Journal*, 52(3):489–505.

Shiller, R. J. (2019). Narratives about technology-induced job degradations then and now. Working Paper 25536, National Bureau of Economic Research.

Shipton, H., West, M. A., Dawson, J., Birdi, K., and Patterson, M. (2006). HRM as a predictor of innovation. *Human Resource Management Journal*, 16(1):3–27.

Sieck, W. R. and Yates, J. F. (2001). Overconfidence effects in category learning: A comparison of connectionist and exemplar memory models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(4):1003.

Siegel, J. J. (2015). *Stocks for the long run: the definitive guide to financial market returns and long-term investment strategies*. McGraw-Hill, New Delhi.

Silver, N. (2012). *The Signal and the Noise: Why So Many Predictions Fail–but Some Don't*. Penguin.

Simon, M. and Houghton, S. M. (2003). The relationship between overconfidence and the introduction of risky products: Evidence from a field study. *Academy of Management journal*, 46(2):139–149.

Simon, M. and Kim, J. (2017). Two sources of overconfidence: Incorporating disconfirming feedback in an entrepreneurial context. *Journal of Small Business Strategy*, 27(3):9–24.

Singh, J. and Singh, H. (2009). Kaizen philosophy: A review of literature. *IUP Journal of Operations Management*, 8(2):51.

Sloof, R. and von Siemens, F. A. (2019). Effective leadership and the allocation and exercise of power in organizations. *The Leadership Quarterly*, In Press, Corrected Proof.

Spiller, S. A. and Belogolova, L. (2017). On consumer beliefs about quality and taste. *Journal of Consumer Research*, 43(6):970–991.

Stansberry, G. (2011). 8 legendary business predictions that missed the mark. Retrived last on 2021/03/06 from https://www.americanexpress.com/en-us/business/trends-and-insights/articles/8-legendary-business-predictions-that-missed-the-mark/.

Stone, D. N., Deci, E. L., and Ryan, R. M. (2009). Beyond talk: Creating autonomous motivation through self-determination theory. *Journal of General Management*, 34(3):75–91.

Surowiecki, J. (2005). *The Wisdom of Crowds*. Anchor.

Svensson, L. E. (1997). Inflation forecast targeting: Implementing and monitoring inflation targets. *European Economic Review*, 41(6):1111–1146.

Thaler, R. H. and Ganser, L. (2015). *Misbehaving: The Making of Behavioral Economics*. WW Norton New York.

The Economist (2014). Keeping its powder dry. Retrieved last on 2019/02/19 from https://www.economist.com/free-exchange/2014/02/06/keeping-its-powder-dry.

Thomas, L. (2019). Walmart will save $ 30 million by putting new step stools in its warehouses. Retrieved last on 2019/09/10 from https://www.cnbc.com/2019/06/07/walmart-will-save-30-million-by-putting-new-step-stools-in-warehouses.html.

Tierney, P. and Farmer, S. M. (2002). Creative self-efficacy: Its potential antecedents and relationship to creative performance. *Academy of Management Journal*, 45(6):1137–1148.

Tost, L. P., Gino, F., and Larrick, R. P. (2012). Power, Competitiveness, and Advice Taking: Why the powerful don't listen. *Organizational Behavior and Human Decision Processes*, 117(1):53–65.

Traum, N. and Yang, S.-C. S. (2011). Monetary and fiscal policy interactions in the post-war us. *European Economic Review*, 55(1):140–164.

Tversky, A. and Kahneman, D. (1971). Belief in the law of small numbers. *Psychological bulletin*, 76(2):105.

Tversky, A. and Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive psychology*, 5(2):207–232.

Tversky, A. and Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, 185(4157):1124–1131.

Tversky, A. and Kahneman, D. (1989). Rational choice and the framing of decisions. In *Multiple criteria decision making and risk analysis using microcomputers*, pages 81–126. Springer.

Van de Ven, A. H. (1986). Central problems in the management of innovation. *Management Science*, 32(5):590–607.

Van Swol, L. M. and Sniezek, J. A. (2005). Factors affecting the acceptance of expert advice. *British Journal of Social Psychology*, 44(3):443–461.

Vardeman, S. B. (1992). What about the other intervals? *The American Statistician*, 46(3):193–197.

von Auer, L. (2005). *Ökonometrie: Eine Einführung, 3. überarbeitete Auflage*. Springer.

Walsh, C. (2003). Speed limit policies: the output gap and optimal monetary policy. *The American Economic Review*, 93(1):265–278.

Walsh, C. E. (1995). Optimal contracts for central bankers. *The American Economic Review*, 85(1):150–167.

West, M. A. (2002). Sparkling fountains or stagnant ponds: An integrative model of creativity and innovation implementation in work groups. *Applied Psychology*, 51(3):355–387.

Woodford, M. (1998). Control of the public debt: a requirement for price stability? In *The Debt Burden and Its Consequences for Monetary Policy*, pages 117–158. Springer.

Woodford, M. (2003). Optimal interest-rate smoothing. *The Review of Economic Studies*, 70(4):861–886.

Yoshida, D. T., Sendjaya, S., Hirst, G., and Cooper, B. (2014). Does servant leadership foster creativity and innovation? A multi-level mediation study of identification and prototypicality. *Journal of Business Research*, 67(7):1395–1404.

Zhou, J. and George, J. M. (2001). When job dissatisfaction leads to creativity: Encouraging the expression of voice. *Academy of Management Journal*, 44(4):682–696.

Zhou, J. and George, J. M. (2003). Awakening employee creativity: The role of leader emotional intelligence. *The Leadership Quarterly*, 14(4-5):545–568.