Essays on Public Economics

Inauguraldissertation
zur
Erlangung des Doktorgrades
der
Wirtschafts- und Sozialwissenschaftlichen Fakultät
der
Universität zu Köln

2024

vorgelegt
von
Marco Wysietzki, M.Sc.
aus
Lindlar

Referent: Prof. Dr. Felix Bierbrauer

Korreferent: Prof. Dr. Andreas Schabert

Tag der Promotion: 17.05.2024

Acknowledgment

Working on this dissertation was a very unique experience. I have learned a lot and I developed on many levels. I would like to express my gratitude to the people who have supported me during this process.

First of all, I would like to thank my supervisors Felix Bierbrauer and Andreas Schabert. I consider myself very lucky to have two such smart and kind supervisors. It was always possible to talk to them and to get valuable feedback. All chapters of this dissertation have benefited a lot from their advice. I am especially glad that both were very caring during the pandemic, which coincided with the start of my PhD phase. They insisted on very regular virtual meetings and made sure that the other PhD students and I were doing well. I really appreciated the pleasant atmosphere during our discussions, which allowed us to talk about non-economic topics as well. I am also thankful for their encouraging manner.

Besides, I want to thank my colleagues and friends Florian Schuster and Jonas Zdrzalek. At the beginning of the pandemic, we connected virtually every day. They were two of the three people I spent by far the most time with during this phase. Fortunately, they are two such great people and we had a good and productive time. I have learned a lot from both of them and I am very happy that we became friends.

I would also like to thank my friend and co-author Anna Kremer. I enjoyed a lot to develop research questions in which both of our research fields overlapped. I have gained a lot of knowledge through working with her and I admire how she manages to take care of work and children at the same time while being remarkably relaxed.

I am very grateful for all the feedback I received during personal conversations or seminars at the Center for Macroeconomic Research. Thanks to Martin Barbie and Michael Krause. Special thanks to Erik Hornung for his kindness, his advice and expertise during my excursion into economic history for my second chapter. I would also like to thank Helge Braun, Michael Kramm and Peter Funk for many insightful conversations during the organization of our teaching. Thanks to Ina Dinstühler, Eva Karnagel-Meinhardt, Yvonne Havertz, Sylvia Hoffmeyer and Kristin Winnefeld for their support with administrative matters and for their exceptionally quick responses to all emails.

Being in the office would not have been half as nice without my great colleagues. Thanks to Tom Brinker, Fabian Knapp, Johannes Kochems, Marten Ritterrath, Fabian Roeben, Luisa Santiago Wolf and Niklas Taft for many good conversations and funny moments during cafeteria and coffee breaks.

Moreover, I would like to thank my family. The biggest thanks to my parents, who have always supported me not only throughout my doctoral studies but also in all the preceding years. You gave me useful advice whenever I needed it and helped me in so many ways. Thank you for all your love, your encouragement and for always giving me the feeling that you were proud of me.

Finally, I would like to thank Nicola Trenz, who has been by my side the whole time and has contributed a lot to the work with her wise thoughts. She endured every articulation of my thoughts on research topics, no matter the time, and seemed at least mostly interested. I cannot express in words how infinitely grateful I am for all her patience, support and love.

Contents

List of Figures

List of Tables

Introduction

This thesis consists of three research projects on different topics connected to taxation. More specifically, the first chapter is about the impact of heterogeneous beliefs on financial crises and how corrective taxation could increase welfare in this setting. The second chapter focuses on income tax reforms in Prussia. It can be argued that some of the used taxes were non-distortionary and can therefore yield new insights in efficient taxation. The third chapter provides an overview of different factors influencing people's willingness to pay taxes, with a focus on regional identity. This allows to evaluate if centralized or decentralized fiscal systems yield higher welfare. These three chapters therefore shed some light on corrective taxation, optimal income taxation and decentralization of fiscal systems.

Chapter 1, which is joint work with Florian Schuster and Jonas Zdrzalek, contributes to the literature on systemic risk. This topic has been studied a lot after the financial crisis. For example Dávila and Korinek (2018) or Jeanne and Korinek (2020) figure out optimal policy in a model with financial amplification. We set up a similar model by adding heterogeneity in beliefs across the population concerning the future state of the world.

Specifically, the model contains a price-dependent collateral constraint on borrowing. Individuals do not internalize that their decisions can have an impact on the collateral price and therefore influence their own borrowing capacities as well as the ones of others. This pecuniary externality gives rise to a financial amplification mechanism: A lower collateral price leads to tighter borrowing constraints. This forces people to sell the collateral, yielding, again, a decrease in collateral price. Agents can differ in their perceived probability distribution of future collateral values. This allows to distinguish relatively more optimistic from relatively more pessimistic individuals and to measure the contribution to systemic risk by different financial market participants.

A competitive equilibrium as well as a constrained social planner's allocation is described. We show the impact of belief heterogeneity in this setting. Furthermore, we show that welfare improvements are possible and we characterize the optimal corrective tax policy.

This analysis yields three main results. First, the likelihood and the extent of financial distress is higher in an economy with heterogeneous beliefs compared to one with homogeneous beliefs. Second, we find an asymmetry concerning individuals' contributions to financial distress. More optimistic agents have a stronger impact compared to more pessimistic ones. Third, we show that a constrained social planner is able to improve welfare using a non-linear Pigouvian tax system, even if the agent's beliefs are private information. In a numerical application, we can show that this tax system yields welfare gains and reduces the probability of financial distress.

Chapter 2 focuses on Prussian income tax reforms and traces the way to a modern comprehensive income tax. During these reforms, certain taxes, that had attractive features from a modern tax theory perspective, were used and abolished after some years. For example, Prussia used lump sum taxes. These taxes do not change individuals' work incentives and thus imply fewer distortions compared to a comprehensive income tax. Lump sum taxes are especially useful if they differ based on individual characteristics that cannot easily be change, but are correlated with ability. This concept is called tagging and allows redistribution without distortions.

In 19th century Prussia, a so-called class tax was introduced. People were taxed depending on which social class they belonged to. It is an example of tagging, as social classes can be assumed to be correlated with abilities. After some decades, several reforms replaced the class tax by a comprehensive income tax. Thus, Prussia did no longer use lump sum taxes and tagging in this context. But why should they abolish instruments that seem to be efficient from today's perspective?

Most of the Prussian reforms were intended to increase tax revenue. Chapter 2 focuses on this aim and checks if it can explain the different Prussian tax reforms and especially the final abolition of the class tax. A model based on Diamond (1998) is used to evaluate the potential tax revenue due to the different reforms. This setup also allows statements concerning the political feasibility of the reforms.

It is shown that Prussian tax reforms increased potential tax revenue even though some appealing features were given up. All reforms can be supported by a majority. In this context, the special Prussian three-class franchise is taken into account. This system connects voting rights to tax payments. Thus, a majority of population is not necessarily the same as a majority of voting power. However, the reforms can be feasible in both cases. Historical data confirm these results and show that most reforms are indeed revenue increasing and political feasible.

Concerning the abolition of tagging, there are hints that the correlation between social class and ability became weaker over time. Thus, chapter 2 discusses the link between this correlation and the gain of tagging. It is proofed formally that a diminishing correlation results in a decreasing benefit from tagging. At some point the negative aspects of tagging, like considerations of justice, can no longer be outweighed. Hence, its abolition can be a consequence.

Chapter 3, which is joint work with Anna Kremer, focuses on the connection of tax morale and the design of tax systems. Some countries use a centralized fiscal system while others use a decentralized one. It basically means that taxes are collected and allocated by a national or a subnational government, respectively.

There is a broad literature on the question of decentralization. A well-known result is the theorem of Oates (1972). It claims that a decentralized system better fits individuals' preferences, while a centralized one implies lower costs due to scale effects. Depending on the size of these effects, one or the other system should be preferred.

However, matching preferences better is not the only advantage of a decentralized system. For example, there is also a positive impact on tax morale (Güth et al., 2005). This mechanism is obviously relevant to make a decision in favor of a decentralized or a centralized system. Thus, the theorem of Oates (1972) is reevaluated in this chapter taking this channel into account.

Apart from the degree of centralization, various other factors can influence the tax morale of individuals. For example, their national identity plays a role (Konrad and Qari, 2012). This could as well be the case for a more regional identity, which in turn could interact with the impact of a decentralized system on tax morale. If a decentralized system is more in line with local preferences, people with a strong local identity may particularly appreciate it, leading to higher tax morale.

A stylized model is used to analyze these effects from a theoretical perspective. This allows to compare individuals' utilities and tax morale in a centralized and a decentralized system. Furthermore, it is possible to compare this framework to the one of Oates (1972) and identify parameter ranges for which the latter would suggest a centralized system, while a decentralized one would be better taking tax morale into account.

These theoretical considerations are followed by an empirical analysis. Using survey data of the European Value Study and the World Value Survey, the impact of decentralization and place identity on tax morale are quantified. For both, the impact is significantly positive. This is true no matter if local, regional or national identity is taken into account. However, local identity has numerically the largest impact. The interaction of place identity and decentralization is negative. Thus, the impact of place identity on tax morale is larger in a centralized system. This could be explained, among others, by scale effects.

Combining the theoretical and the empirical analysis, the model is calibrated for the example of Germany. It shows that a decentralized system yields a higher welfare in this country compared to a centralized one.

Contribution

Chapter 1 is joint work with Florian Schuster and Jonas Zdrzalek. We developed the research idea together. Afterwards, we set up the framework and did most parts of the analysis jointly. Apart from that, Florian Schuster wrote the first draft and provided most parts of the derivations for section 3.3 *Equilibrium effects of variations in beliefs*, Jonas Zdrzalek prepared section 2 *Related literature*, while I developed the *Numerical application* in section 4.4. We jointly refined everything.

Chapter 3 is joint work with Anna Kremer. We developed the research idea together. Especially the first steps of our work were done jointly. Afterwards, Anna Kremer concentrated on section 4 *Empirical analysis* while I focused on sections 3 *Theoretical considerations* and 5 *Decentralization of Germany*. However, we were in constant communication and refined everything jointly. We wrote the first draft together.

Chapter 1

How Heterogeneous Beliefs Trigger Financial Crises

By Florian Schuster, Marco Wysietzki and Jonas Zdrzalek

Abstract

We present a theoretical framework to characterize how financial market participants contribute to systemic risk, allowing us to derive optimal corrective policy interventions. To that end, we embed belief heterogeneity in a model of frictional financial markets. We document the asymmetry that, by their behavior, relatively more optimistic agents contribute more strongly to financial distress than more pessimistic agents do. We further show that financial distress is generally more likely in an economy whose agents hold heterogeneous rather than homogeneous beliefs. Based on these findings, we propose a system of non-linear Pigouvian taxes as the optimal corrective policy, which proves to generate considerable welfare gains over the linear policy advocated by former studies.

1.1 Introduction

Systemic risk has been well studied since the global financial crisis. An important question remains yet to be explored: How can individual financial market participants' contributions to system-wide financial distress be measured, and how can they be addressed accordingly by Pigouvian policies?

The literature has provided valuable insights into the matter of measuring systemic risk. In particular, various approaches to specifying the financial system's exposure to certain institutions' risk taking have been suggested (Acharya et al., 2012; Adrian and Brunnermeier, 2016; Acharya et al., 2017). However, they cannot causally attribute the extent to which individual financial decisions, and the resulting marginal effects on market prices, contribute to systemic distress. Corrective policies thus lack a basis to be designed on.

In this paper, we build a theoretical framework to analyze contributions to systemic risk and optimal corrective policy interventions. We augment established models of financial frictions by heterogeneity of beliefs across the population, giving rise to differentiated risk taking in financial decisions. The latter are observable, so we may characterize distress contributions explicitly.

We find that economic agents make asymmetric contributions to financial distress, with more optimistic agents making larger contributions than pessimistic agents. We further show that financial distress is generally more likely in an economy whose agents hold heterogeneous rather than homogeneous and rational beliefs. The optimal policy we propose is a system of non-linear Pigouvian taxes, which proves to generate considerable welfare gains over the linear policy advocated by previous studies.

To the best of our knowledge, we are the first to combine a model of frictional financial markets with belief heterogeneity embedded in a single framework. Specifically, our model incorporates a price-dependent collateral constraint on borrowing. It introduces a pecuniary externalitiy, as economic agents do not internalize that their decisions mutually affect their borrowing capacities, which, in turn, establishes a financial amplification mechanism. Agents may hold heterogeneous beliefs in the sense of perceiving differentiated probability distributions over the future state of the world. This setup allows us to distinguish relatively more optimistic from pessimistic individuals.

We use this model to analyze the interaction of the financial friction and belief heterogeneity. First, we characterize how the latter impacts the probability of distress in the competitive equilibrium, as well as the equilibrium allocation, collateral prices, and externalities. We then perform an efficiency analysis, showing how a constrained social planner can attain a welfare improvement compared to the competitive equilibrium. We characterize her optimal corrective policies numerically, and evaluate how they influence social welfare and the probability of financial distress.

The analysis produces three key results. First, we find an asymmetry between optimistic and pessimistic agents' contributions to financial distress, attributing a stronger impact to the former. Moreover, optimistic agents prove to put downward pressure on collateral prices, tightening financial constraints. Under reasonable assumptions on the distribution of beliefs across the population, we conclude that belief heterogeneity precipitates financial distress.

Second, we show that, compared to an economy where agents hold a homogeneous and rational belief, belief heterogeneity raises the likelihood of financial distress. The reason is that, for collateral constraints to be binding, no sharp exogeneous shock to aggregate investment or net worth is required. Such a shock is typically assumed in the literature. Instead, it suffices that some agents' beliefs deviate from the ex post state of the world.

Third, we prove that, even though beliefs are agents' private information, a constrained social planner is able to establish a welfare improvement by means of a non-linear Pigouvian tax system. Our policy proposal contrasts the linear Pigouvian taxation proposed by previous studies. We provide numerical applications suggesting that our non-linear approach produces welfare gains relative to linear policies, and reduces the probability of financial distress.

This paper makes several important contributions. It provides a formal framework which can be used for further analyses of financial amplification mechanisms in environments where economic agents do not have rational expectations, but potentially feature heterogeneous beliefs. This helps to explicitly characterize how different market participants contribute to financial crises. Moreover, this lays the ground for an optimal design of prudential policies. Policy proposals not accounting for belief divergences might have only limited success if beliefs vary widely across financial market participants, as they cannot account for their respective contributions to systemic risk. This is particularly relevant during different phases of the business cycle, as investors' beliefs prove to fluctuate and diverge largely between booms and busts (Aliber and Kindleberger, 2015; Minsky, 1986; Kaplan et al., 2020; Mian and Sufi, 2022).

The remainder of this paper is organized as follows. We review the related literature in section 1.2. Section 1.3 develops the baseline model, and analyzes the competitive equilibrium. In section 1.4, we describe the externalities present in our model, derive optimal corrective policies, and perform normative analyses numerically. We provide some final remarks in section 1.5.

1.2 Related Literature

Financial amplification and pecuniary externalities. Our model combines two strands of literature. First, it relates to the literature on financial amplification, including studies of pecuniary externalities in particular. This literature originates from Fisher (1933), and was extended by analyses of borrowing constraints and their effects on asset prices by Bernanke and Gertler (1990), Shleifer and Vishny (1992), Kiyotaki and Moore (1997), Brunnermeier and Pedersen (2009), and Acharya et al. (2011). Hart (1975) and Stiglitz (1982) moreover prove the presence of pecuniary externalities in incomplete markets.[1]

Welfare implications of pecuniary externalities are examined in Gromb and Vayanos (2002), Caballero and Krishnamurthy (2003), Lorenzoni (2008), and Caballero and Lorenzoni (2014). While these papers focus on externalities affecting borrowers' net worth, Jeanne and Korinek (2010), Bianchi (2011), Bianchi and Mendoza (2018), Dávila and Korinek (2018), and Jeanne and Korinek (2019) explore the collateral channel of financial amplification that can lead to financial crises. Since we are modeling externalities equivalently, we adopt their terminology and basic model structure.

Furthermore, we derive optimal corrective policies implemented by a a constrained social planner, referring to the early contributions of Stiglitz (1982) and Geanakoplos and Polemarchakis (1986). The policy maker in our model applies an ex-ante Pigouvian tax along the lines of Jeanne and Korinek (2010), Dávila and Korinek (2018), Jeanne and Korinek (2019), and Jeanne and Korinek (2020).[2]

In the domain of the aforementioned literature on financial market externalities, this paper links to articles that focus on defining measures of systemic risk. Notably, Adrian and Brunnermeier (2016) propose $\Delta CoVar$, a measure capturing the interdependences between specific financial institutions and the entire financial system. Furthermore, Acharya et al. (2012) and Acharya et al. (2017) model individual institutions' exposure to financial crises. For an overview of quantitative measures of systemic risk, see Bisias et al. (2012).

Macroeconomic perspectives on belief heterogeneity. The second strand of the literature relates to macroeconomic perspectives on belief heterogeneity. The idea of belief heterogeneity shaping market outcomes was pioneered by Keynes (1936), Minsky (1977), and Aliber and Kindleberger (2015). Since then, the literature has provided

[1] For survey articles, see Shleifer and Vishny (2011) and Brunnermeier and Oehmke (2013).

[2] The social planner in our model has an instrument at hand which could be interpreted as a financial transaction tax. So the interested reader is referred to the literature on financial transaction taxes initiated by Tobin (1978), and extended by Summers and Summers (1989) and Stiglitz (1989).

evidence that belief heterogeneity is relevant for asset prices and market volatility, in particular during the recent financial crisis (Harrison and Kreps, 1978; Scheinkman and Xiong, 2003; Reinhart and Rogoff, 2008; Simsek, 2013; Cheng et al., 2014; Gennaioli and Shleifer, 2018; Adam and Nagel, 2022).

As in all normative studies involving heterogeneity of beliefs, we face the challenge of how to aggregate welfare properly. Several approaches have been suggested, such as the welfare criteria put forth by Gilboa et al. (2014), Gayer et al. (2014), Brunnermeier et al. (2014), Blume et al. (2018), and Kim and Kim (2021).

Prior research has already combined the two strands of literature, particularly in the context of heterogeneous beliefs and leveraged speculation.[3] Geanakoplos (1996) was the first to model a general equilibrium with endogenous collateral constraints and heterogeneous beliefs, which was further developed in subsequent studies (Geanakoplos, 2003, 2010), showing that heterogeneity of beliefs fosters credit and leverage cycles. Simsek (2013) generalizes the framework, and focuses on various degrees of heterogeneity.

Belief heterogeneity and business cycles. Furthermore, our analysis is associated with the literature on business cycles. In particular, we refer to the role of beliefs in booms and busts. Minsky (1977, 1986) and Aliber and Kindleberger (2015) show how asset price booms are linked to increasing optimism. Rising asset prices create states of 'mania' in which investors are overly optimistic and hold the asset as they strongly believe that prices will continue to rise. Adam et al. (2017) argue that shifts in investors beliefs about future capital gains are highly relevant in explaining cyclical asset price fluctuations. Kaplan et al. (2020) show that such shifts were the driving force of the house-price boom prior to the global financial crisis. Moreover, Mian and Sufi (2022) elaborate on how important increasing divergence of beliefs was in the build-up of the house price boom prior to the 2007-2008 financial crisis.

Methodological approach. Lastly, our investigation of comparative statics with respect to the economy's belief structure closely relates to Dávila and Walther (2023), who study optimal leverage policies in response to changing beliefs. We follow their approach of applying methods from the calculus of variation to equilibrium variables under belief heterogeneity.

[3] Xiong (2013) and Simsek (2021) review the literature on asset trading driven by heterogeneous beliefs in great detail.

1.3 Model

The aim of this chapter is to explore financial amplification mechanisms in an environment where agents hold heterogeneous beliefs about the future. To that end, we set up a model featuring frictional financial markets, and enrich it by belief heterogeneity across agents. We derive the competitive equilibrium of this economy, and study how it is impacted by variations in beliefs. The framework allows us to distinguish the respective contributions of optimistic and pessimistic agents to financial amplification, and to evaluate the probability of distress in economies with different belief structures. Our results lay the ground for the study of optimal Pigouvian policies in the next section.

1.3.1 Setup

We develop a model of a small open economy with three periods $t = 0, 1, 2$, and two classes of agents, called lenders and investors. Lenders trade debt securities with investors, or save in a zero return storage technology. The interest rate is exogeneous and normalized to zero for simplicity, and lenders are assumed to be risk-neutral.

Investors are divided into J groups indexed by $j \in \{1, ...J\}$, each of which consists of a continuum of investors. Each group has a population share s^j, that is common knowledge, and derives utility from a single consumption good c_t^j according to a concave and strictly increasing utility function $u\left(c_t^j\right)$. Population shares are collected in the vector $s = \{s^j\}_{j\in\{1,...J\}}$.

In $t = 0$, investors receive an endowment $e > 0$, as well as an initial amount of assets $\bar{a} > 0$. They can borrow or save d_0^j to finance consumption, and to further invest into a_0^j units of the asset.[4] The asset is traded at a price q_0, and exists in fixed supply.

In $t = 1$, financial investment pays off an a priori uncertain dividend $R \in \left[\underline{R}, \overline{R}\right]$, which different groups of investors hold specific beliefs about. After all uncertainty has been resolved at the beginning of the period, investors repay former debt d_0^j, issue new debt d_1^j, and trade again, purchasing or liquidating l_1^j claims on the asset at price q_1.

Debt issuance in $t = 1$ is restricted by a borrowing constraint

$$d_1^j \leq \phi q_1 \left(a_0^j - l_1^j\right).$$

The constraint implies that investors borrow against their asset position at the end of the period.[5]

[4]Lenders' endowment is assumed to make the supply of debt securities perfectly elastic to demand. That is, all investors' borrowing preferences can be satisfied by assumption. This includes the possibility of savings $d_0^j < 0$.

[5]To rationalize this constraint, we adopt the mechanism suggested by Jeanne and Korinek (2019).

In $t = 2$, net of claims $a_0^j - l_1^j$ materializes and debt d_1^j must be repaid, determining final consumption c_2^j.

Our model features two important components. First, financial markets exhibit a friction, captured by the the borrowing constraint. It incorporates a financial amplification mechanism within our framework, and results in a pecuniary externality. Second, we allow investors to hold different beliefs about the asset pay-off R.

Definition 1. *Let $F(R)$ be the true cumulative distribution function (cdf) of R, and $F^j(R)$ be the cdf perceived by type-j investors. We refer to heterogeneous beliefs if each type of investors j perceives an idiosyncratic distribution of R, i.e. $F^i(R) \neq F^j(R)$ for all $i \neq j$. We refer to homogeneous beliefs if all types of investors have rational expectations, i.e. $F^j(R) = F(R)$ for all j.*

The vector $\mathcal{F} = \{F^j(R)\}_{j \in \{1,\ldots J\}}$ characterizes the complete set of beliefs existing in the economy, which is publicly known. Beliefs are distributed discretely across types, so each cdf $F^j(R)$ appears with frequency s^j.

1.3.2 Competitive Equilibrium

To derive the competitive equilibrium, we first conduct individual optimization backwards from $t = 2$ to $t = 0$. We distinguish between state variables of type-j individuals, i.e. $\left\{a_0^j, d_0^j\right\}$, and aggregate state variables of group j, denoted by $\left\{\tilde{a}_0^j, \tilde{d}_0^j\right\}$.

Optimization in $\mathbf{t = 1, 2}$. The optimization problem of type-j investors in $t = 1$ reads

$$V_1^j\left(a_0^j, d_0^j | \tilde{a}_0, \tilde{d}_0\right) = \max_{c_1^j, c_2^j, d_1^j, l_1^j \leq a_0^j} u\left(c_1^j\right) + u\left(c_2^j\right) \quad \text{s.t.}$$

$$\left(\lambda_1^j\right) \quad c_1^j = R a_0^j + q_1 l_1^j + d_1^j - d_0^j \tag{1.1}$$

$$\left(\lambda_2^j\right) \quad c_2^j = a_0^j - l_1^j - d_1^j \tag{1.2}$$

$$\left(\eta_1^j\right) \quad d_1^j \leq \phi q_1 \left(a_0^j - l_1^j\right), \tag{1.3}$$

The constraint bases on the presumption that investors lack commitment to repay. When investors renegotiate debt obligations, they make a take-it-or-leave-it offer in order to lower the amount of outstanding debt. If lenders reject the offer, they may seize a fraction ϕ of investors' assets and sell it at the prevailing market price. Lenders will hence accept the offer provided the repayment exceeds the current market value of seizable positions. This being said, we may assume without loss of generality that default and renegotiations never occur in equilibrium. One could further consider a similar restriction of debt issuance in period $t = 0$, which we neglect on the grounds that there is no role for macroprudential interventions in that period. Binding borrowing constraints in $t = 0$ would limit the set of cases when the period-1 constraint is binding, however without altering the results of our analysis, which focuses on situations within this set.

where investors take group-wide aggregate states $\tilde{a}_0 = \left\{\tilde{a}_0^j\right\}_{j\in\{1,\ldots J\}}$ and $\tilde{d}_0 = \left\{\tilde{d}_0^j\right\}_{j\in\{1,\ldots J\}}$ as given because they affect the equilibrium asset price q_1. Let λ_1^j and λ_2^j be the Lagrange multipliers for the budget constraints (1.1) and (1.2), respectively, and η_1^j for the borrowing constraint (1.3).

This problem produces the following Euler equations for each j:

$$u'\left(c_1^j\right) - \eta_1^j = u'\left(c_2^j\right) \tag{1.4}$$

$$q_1 u'\left(c_1^j\right) - \eta_1^j \phi q_1 = u'\left(c_2^j\right), \tag{1.5}$$

jointly yielding equilibrium price equations

$$q_1 = \frac{u'\left(c_2^j\right)}{(1-\phi)u'\left(c_1^j\right) + \phi u'\left(c_2^j\right)} \tag{1.6}$$

for each j.

Optimization in t = 0. In $t = 0$, the optimization of a type-j investor is

$$\max_{c_0^j, a_0^j \geq 0, d_0^j} \; u\left(c_0^j\right) + E^j\left[V_1^j\left(a_0^j, d_0^j | \tilde{a}_0, \tilde{d}_0\right)\right] \quad \text{s.t.}$$

$$\left(\lambda_0^j\right) \quad c_0^j = e + d_0^j + q_0\left(\bar{a} - a_0^j\right), \tag{1.7}$$

where the expectation operator is indexed by j, capturing potentially differing beliefs, and λ_0^j denotes the Lagrange multiplier for the period-0 budget constraint. Eliminating Lagrange multipliers, we obtain the following two optimality conditions:

$$q_0 u'\left(c_0^j\right) = E^j\left[Ru'\left(c_1^j\right) + u'\left(c_2^j\right) + \eta_1^j \phi q_1\right] \tag{1.8}$$

$$u'\left(c_0^j\right) = E^j\left[u'\left(c_1^j\right)\right]. \tag{1.9}$$

Equilibrium. In equilibrium, the asset market is cleared in both periods $t = 0$ and $t = 1$, formalized by the conditions

$$\sum_{j=1}^{J} s^j a_0^j = \bar{a} \tag{1.10}$$

and

$$\sum_{j=1}^{J} s^j l_1^j = 0. \tag{1.11}$$

Complementing the optimality conditions derived thus far, they complete the set of equilibrium conditions. In a symmetric equilibrium, investors are identical within each group j, i.e. $x_t^j = \tilde{x}_t^j$ for all j with $x \in \{c, a, d, l, \lambda, \eta\}$. We may thus define the symmetric competitive equilibrium as follows.

Definition 2. *A competitive equilibrium consists of an allocation* $\left\{\tilde{c}_0^j, \tilde{c}_1^j, \tilde{c}_2^j, \tilde{a}_0^j, \tilde{d}_0^j, \tilde{d}_1^j, \tilde{l}_1^j\right\}_{j\in\{1,\ldots J\}}$*, a sequence of multipliers* $\tilde{\eta}_1 = \left\{\tilde{\eta}_1^j\right\}_{j\in\{1,\ldots J\}}$*, and prices* $\{q_0, q_1\}$*, satisfying equations* (1.1)*,* (1.2)*,* (1.4)*,* (1.5)*,* (1.7)*,* (1.8)*,* (1.9)*, and a complementary slackness condition for all* j*, as well as the market clearing conditions* (1.10) *and* (1.11)*, given population shares* s *and beliefs* $\mathcal{F}$*.*

The competitive equilibrium reflects the two main components of our model: the financial friction and potential belief disagreements. The financial friction introduces a wedge between market prices of the asset as well as debt, and investors' marginal rates of substitution across periods. The wedge is formally represented by the multiplier $\tilde{\eta}_1^j$ that appears in equations (1.4), (1.5), and (1.8). In the latter two equations, the term $\tilde{\eta}_1^j \phi q_1$ captures the collateral premium of the asset, as each additional unit of $\tilde{a}_0^j$ and $\tilde{l}_1^j$ relaxes the constraint.

To highlight the impact of belief heterogeneity, we compare the competitive equilibrium under heterogeneous and homogeneous beliefs. If investors have heterogeneous expectations of the return R, they evaluate expected marginal benefits of investment and borrowing differently. Formally, group-specific expectation operators E^j apply in the Euler equations (1.8) and (1.9), resulting in group-specific values of $\tilde{a}_0^j$, $\tilde{d}_0^j$, and of the shadow price of borrowing $\tilde{\eta}_1^j$.

If, in contrast, investors hold a homogeneous belief, their marginal rates of substitution are identical, as is the shadow value of borrowing. Importantly, intertemporal substitution in this case is only possible through debt or savings $\tilde{d}_t^j$. The reason is that investors do not trade in excess of the initial asset endowment neither in $t = 0$ nor in $t = 1$, i.e. $\tilde{a}_0^j = \bar{a}$ and $\tilde{l}_1^j = 0$ for all j.

In the following, we restrict the set of equilibria taken into account in the analysis. Since we are only interested in situations when financial distress occurs, the model parameters, comprising risk aversion A, beliefs $\mathcal{F}$, the realized return $\hat{R}$, as well as the margin requirement ϕ, must satisfy that, in equilibrium, the asset is traded and constraints are binding ($\tilde{\eta}_1^j > 0$).[6]

Period-1 equilibrium price. Given its impact on the borrowing constraint, the equilibrium collateral price q_1 is a key variable in our model. We show its existence and

[6] We make parameter restrictions explicit in the derivations of our results, provided in the appendix.

uniqueness, and how it interacts with the multiplier of the borrowing constraint.

Proposition 1.

(i) The equilibrium price q_1 exists.

(ii) If at least one type of investors j receives a return as expected or higher, i.e. $E^j[R] \leq \hat{R}$ for at least one j and any realization $\hat{R}$ of R, the equilibrium price is unique, satisfying $q_1 \leq 1$, and the following two equivalences hold:

(1) $q_1 = 1$ iff $\tilde{\eta}_1^j = 0$ for all j

(2) $q_1 < 1$ iff $\tilde{\eta}_1^j > 0$ for at least one j.

Proposition 1 first states that the equilibrium exists. Second, assuming that there is positive demand because at least one type makes a profit from investment, it claims that the equilibrium price is unique, and characterizes its relation with the borrowing constraint.[7] The constraint is binding at a price smaller than 1, but slack if $q_1 = 1$. At this price, investors are indifferent between purchasing or selling claims.

The two equivalences in part *(ii)* of Proposition 1 formalize this indifference property. They imply that either all or none of the investors are constrained by the borrowing limit. It is sufficient that only one group of investors is forced to liquidate claims on the market, i.e. $\tilde{l}_1^j > 0$, to reduce the price q_1 to a level below one. This deflation either constrains other investors via a tighter borrowing limit, or it gives them a pecuniary incentive to issue as much debt as possible. They do so to purchase additional claims, i.e. $\tilde{l}_1^j < 0$. To see this, recall the budget constraints (1.1) and (1.2), and note that, provided $q_1 < 1$, every purchased unit of claims offers a positive return $1 - q_1 > 0$ in the final period. Hence, in order to transfer funds to $t = 2$, solvent investors prefer additional investment $\tilde{l}_1^j < 0$ over savings $\tilde{d}_1^j < 0$. For a price $q_1 = 1$, they are indifferent between both ways of intertemporal substitution.[8]

1.3.3 Equilibrium Effects of Variations in Beliefs

In this section, we analyze how variations in beliefs affect the allocation and prices in the competitive equilibrium. We show how the two main ingredients of our model,

[7] However, the equilibrium price exists even if demand is zero, as this scenario corresponds to all investors being bankrupt, and infinitely many prices satisfy the Walrasian equilibrium definition. Abstracting from this case, we focus on equilibria with positive demand, which turn out to be uniquely determined.

[8] Formally, one of the Euler equations (1.4) and (1.5) is redundant in the unconstrained case, i.e. if $q_1 = 1$ and $\tilde{\eta}_1^j = 0$ for all j. Intuitively, investors are indifferent between the instruments $\tilde{l}_1^j$ and $\tilde{d}_1^j$, given that both promise a zero net return. We assume without loss of generality that there is no trade in the unconstrained economy, i.e. $\tilde{l}_1^j = 0$ for all j.

the financial friction and heterogeneity of beliefs, interact. The results of this comparative statics exercise allow us to specify how different types contribute to financial amplification, and how belief heterogeneity affects the overall probability of financial distress.

To keep the model tractable, we henceforth impose the following assumption without further mention.

Assumption 1. *Investors have exponential preferences of the form* $u\left(c_t^j\right) = -\exp\left(-Ac_t^j\right)$*, where absolute risk aversion* $A = -\frac{u''\left(c_t^j\right)}{u'\left(c_t^j\right)}$ *is constant (CARA).*[9]

The assumption that absolute risk aversion is constant is useful to simplify the comparative statics analysis below.

We start out by examining the effect of changes in period-0 variables on the equilibrium price in $t = 1$, before analyzing how belief variations impact the equilibrium values of investment and borrowing in period $t = 0$. Note that the period-1 equilibrium price q_1 is no direct function of beliefs $\mathcal{F}$, but only through period-0 choices $\tilde{a}_0(\mathcal{F})$ and $\tilde{d}_0(\mathcal{F})$, i.e. $q_1 = q_1\left(\tilde{a}_0(\mathcal{F}), \tilde{d}_0(\mathcal{F})\right)$. Thus, this two-step procedure allows us to elaborate the relationship between the set of beliefs in the economy and the equilibrium price q_1, which defines the tightness of the borrowing constraint, and measures the extent of financial distress.

Period-0 allocation and the equilibrium price. Proposition 2 states how the equilibrium price q_1 is linked to period-0 levels of investment and debt.

Proposition 2.

(i) If investors hold heterogeneous beliefs $\mathcal{F}$*, the period-1 equilibrium price* q_1 *is decreasing with period-0 investment and borrowing, i.e., for all* j*,*

$$\frac{\partial q_1}{\partial \tilde{a}_0^j} < 0 \text{ and } \frac{\partial q_1}{\partial \tilde{d}_0^j} < 0.$$

(ii) If investors hold the homogeneous belief $F(R)$*, the period-1 equilibrium price* q_1 *is decreasing with period-0 borrowing, i.e.*

$$\frac{\partial q_1}{\partial \tilde{d}_0} < 0.$$

[9] For expositional reasons, we continue using the general notation $u\left(c_t^j\right)$.

Proposition 2 states that more investment and borrowing in period $t = 0$ have a diminishing effect on the future equilibrium asset price. While the former is irrelevant in the homogeneous case, where trade does not occur, the negative effect of borrowing persists.

The two effects work through different channels, illustrated by the budget constraints (1.1) and (1.2). First, investment in $\tilde{a}_0^j$ increases period-2 consumption $\tilde{c}_2^j$ one-to-one, while $\tilde{c}_1^j$ rises with factor $\hat{R}$. Thus, in a sufficiently adverse state, satisfying $\hat{R} < 1$, consumption in the last period $\tilde{c}_2^j$ increases by more in response to investment than $\tilde{c}_1^j$. To smooth consumption, investors redistribute resources from $t = 2$ to $t = 1$ by liquidating $\tilde{l}_1^j$ units of their asset position (or purchasing less additional units). Second, higher indebtedness $\tilde{d}_0^j$ reduces the initial period-1 wealth $\hat{R}\tilde{a}_0^j - \tilde{d}_0^j$, raising the risk of being constrained and forced to liquidate a fraction of the portfolio. Both channels result in a higher supply (and a lower demand) of claims, which, in turn, reduce the equilibrium price q_1.

Beliefs and the period-0 allocation. We now turn to the relationship between investment $\tilde{a}_0(\mathcal{F})$ and borrowing $\tilde{d}_0(\mathcal{F})$ and investors' beliefs $\mathcal{F}$. To that end, we employ methods from the calculus of variation. We adopt the following procedure, that was first applied to heterogeneous belief environments by Dávila and Walther (2023). Recall that type-j investors' beliefs are characterized by the perceived distribution of R with cdf $F^j(R)$. Consider a perturbation to beliefs of the form $F^j(R) + \epsilon G^j(R)$, where $\epsilon > 0$ is an arbitrary number, and $G^j(R)$ captures the direction of the perturbation. $F^j(R) + \epsilon G^j(R)$ is required to be a valid cdf for small enough ϵ, so we assume it is continuous and differentiable, satisfies $G(\underline{R}) = G\left(\overline{R}\right) = 0$, and $\partial\left(F^j(R) + \epsilon G^j(R)\right)/\partial R \geq 0$ for sufficiently small ϵ.

This setup allows us to specify the concepts of optimism and pessimism. These terms are defined relative to each other in the sense of first-order stochastic dominance. A perturbation $G^j(R)$ makes type-j investors more optimistic if and only if it satisfies $F^j(R) + \epsilon G^j(R) \leq F^j(R)$ for all R. It is easy to see that a more optimistic belief requires the perturbation to have a non-positive direction, i.e. $G^j(R) \leq 0$ for all R. Analogously, investors of type j are made more pessimistic through a perturbation with direction $G^j(R) \geq 0$ for all R. Intuitively, investors are more optimistic if they assign lower probabilities than pessimists to low returns, so their cdf is shifted downwards.[10]

Using this technique, we show how a variation of a type's belief alters its individual

[10] In the case of investors holding homogeneous beliefs, a perturbation implies a variation of the true distribution $F(R)$.

choices of investment and debt issuance. The corresponding functional derivatives are

$$\frac{\delta \tilde{a}_0^j}{\delta F^j} \cdot G^j \text{ and } \frac{\delta \tilde{d}_0^j}{\delta F^j} \cdot G^j,$$

where δ denotes the operator for functional derivatives. Proposition 3 summarizes the results.

Proposition 3.

(i) Let investors hold heterogeneous beliefs $\mathcal{F}$, and let $G^j(R)$ be the direction of a perturbation of type-j investors' belief $F^j(R)$. More optimistic (pessimistic) investors invest and borrow more (less), i.e.

$$\frac{\delta \tilde{a}_0^j}{\delta F^j} \cdot G^j \begin{cases} \geq 0, & G^j(R) \leq 0 \\ < 0, & G^j(R) \geq 0 \end{cases} \quad \textit{and} \quad \frac{\delta \tilde{d}_0^j}{\delta F^j} \cdot G^j \begin{cases} \geq 0, & G^j(R) \leq 0 \\ < 0, & G^j(R) \geq 0 \end{cases}.$$

(ii) Let investors hold the homogeneous belief $F(R)$, and let $G(R)$ be the direction of a perturbation. The more optimistic (pessimistic) the homogeneous belief is, the more (less) investors borrow, i.e.

$$\frac{\delta \tilde{d}_0}{\delta F} \cdot G \begin{cases} \geq 0, & G(R) \leq 0 \\ < 0, & G(R) \geq 0 \end{cases}.$$

The essential insight from Proposition 3 is that investment and borrowing are monotonic functions of beliefs. The more optimistic a group of investors is, the more it invests into the asset, and the more debt it issues. The opposite holds true for more pessimistic groups. If investors are homogeneous, only borrowing responds to variations in beliefs, while the asset is not traded.

Beliefs and the equilibrium price. Combining the results from Propositions 2 and 3, we describe how behavioral responses of investors to changes in beliefs $\mathcal{F}$ impact the period-1 equilibrium price $q_1\left(\tilde{a}_0(\mathcal{F}), \tilde{d}_0(\mathcal{F})\right)$ in Theorem 1.

Theorem 1.

(i) Let investors hold heterogeneous beliefs $\mathcal{F}$.

(1) Let further $G^j(R)$ be the direction of a perturbation of type-j investors' belief $F^j(R)$, and beliefs $F^i(R)$ be constant for all $i \neq j$. If the perturbation makes

investors of type j more optimistic (pessimistic), the period-1 equilibrium price q_1 is lower (higher), i.e.

$$\frac{\delta q_1}{\delta F^j} \cdot G^j \begin{cases} \leq 0, & G^j(R) \leq 0 \\ > 0, & G^j(R) \geq 0 \end{cases}.$$

(2) Let further $G^j(R) < 0 < G^i(R)$ with $|G^j(R)| = |G^i(R)|$ for all R be the directions of two perturbations that make investors of type j more optimistic, and investors of type i more pessimistic by the same magnitude. The behavioral responses to the perturbation with direction $G^j(R)$ have a stronger impact on the period-1 equilibrium price q_1 than those of the perturbation with direction $G^i(R)$, i.e.

$$\left|\frac{\delta q_1}{\delta F^j} \cdot G^j\right| \geq \left|\frac{\delta q_1}{\delta F^i} \cdot G^i\right|.$$

(ii) Let investors hold the homogeneous belief $F(R)$, and let $G(R)$ be the direction of a perturbation. If the perturbation makes investors more optimistic (pessimistic), the period-1 equilibrium price q_1 is lower (higher), i.e.

$$\frac{\delta q_1}{\delta F} \cdot G \begin{cases} \leq 0, & G(R) \leq 0 \\ > 0, & G(R) \geq 0 \end{cases}.$$

Theorem 1 comprises the first set of key results of this paper. Part (i) characterizes the relationship of q_1 and heterogeneous beliefs. The more optimistic investors are, the lower the collateral price is in equilibrium. Conversely, if investors hold more pessimistic beliefs, the equilibrium price is higher. This result originates from the two monotonicities we have established in Propositions 2 and 3: q_1 responds monotonically to period-0 investment and borrowing, which, in turn, are monotonically driven by beliefs.

However, according to statement (2), the equilibrium price responds asymmetrically to symmetric variations of beliefs. Consider the thought experiment of two distinct perturbations, one making investors of type j more optimistic, the other making investors of type i more pessimistic, both to the very same extent. Formally, this is equivalent to decreasing type j's and increasing type i's probability mass for each realization $\hat{R}$ by the same amount. The statement argues that the perturbation to j dominates the perturbation to i. Thus, the equilibrium price turns out to be lower. More precisely, the perturbation to the optimistic type j exerts a downward effect that outweighs the upward effect from the perturbation to the pessimistic type i, resulting in a lower equilibrium price. The asymmetry between optimistic and pessimistic investors' influence on q_1 is the main result of Theorem 1, which we will use to derive optimal corrective policies in the following section.

Key to understand the asymmetry is the collateral constraint. By the two perturbations, type-j investors become more optimistic, willing to invest and borrow more, while type-i investors become more pessimistic, willing to invest less and save more. Importantly, both types have the incentive to invest into the asset as collateral in $t = 1$. In $t = 0$, this incentive amplifies type j's willingness to extend investment, but it counteracts type i's willingness to reduce investment. Accordingly, it induces type j to increase period-0 borrowing by more than type i increases period-0 savings. Therefore, when the constraint is binding in the following period $t = 1$, type-j investors' supply of liquidated claims will relatively exceed type-i investors' demand, which can only be equated for a lower equilibrium price q_1.

Part (*ii*) of Theorem 1 states that the former result holds true in the case of a homogeneous belief as well. A lower equilibrium price will arise if the uniform belief is more optimistic, and q_1 will be higher if it is more pessimistic.

Probability of financial distress. While Theorem 1 specifies how different types of investors contribute to financial amplification, we finally evaluate how heterogeneity affects the overall probability of financial distress. We apply the method proposed by Dávila and Walther (2023) to prove that financial distress is more likely under heterogeneous beliefs. The probability of financial distress is determined by the lowest possible realization of R such that the constraints are slack.

Definition 3. *Let* $\hat{R}^*_{het} \equiv \min\left\{\hat{R} \mid \tilde{\eta}^j_1 = 0 \text{ for all } j\right\}$ *and* $\hat{R}^*_{hom} \equiv \min\left\{\hat{R} \mid \tilde{\eta}_1 = 0\right\}$ *be the lowest possible realizations of* R *such that the borrowing constraints are slack in the competitive equilibrium if investors hold heterogeneous beliefs* $\mathcal{F}$ *or the homogeneous belief* F*, respectively.*

Definition 3 translates into the mappings $\hat{R} \mapsto q_1\left(\hat{R}\right)$ as q_1 serves as a measure of financial distress, formally written as

$$q_1 \begin{cases} = 1 & \hat{R} \geq \hat{R}^*_{het} \\ < 1 & \hat{R} < \hat{R}^*_{het} \end{cases} \text{or } q_1 \begin{cases} = 1 & \hat{R} \geq \hat{R}^*_{hom} \\ < 1 & \hat{R} < \hat{R}^*_{hom} \end{cases}.$$

in the heterogeneous and the homogeneous case, respectively. Figure 1.1 portrays an illustration of the two mappings.[11]

[11] Figure 1.1 is based on the numerical application provided in section 1.4.4.

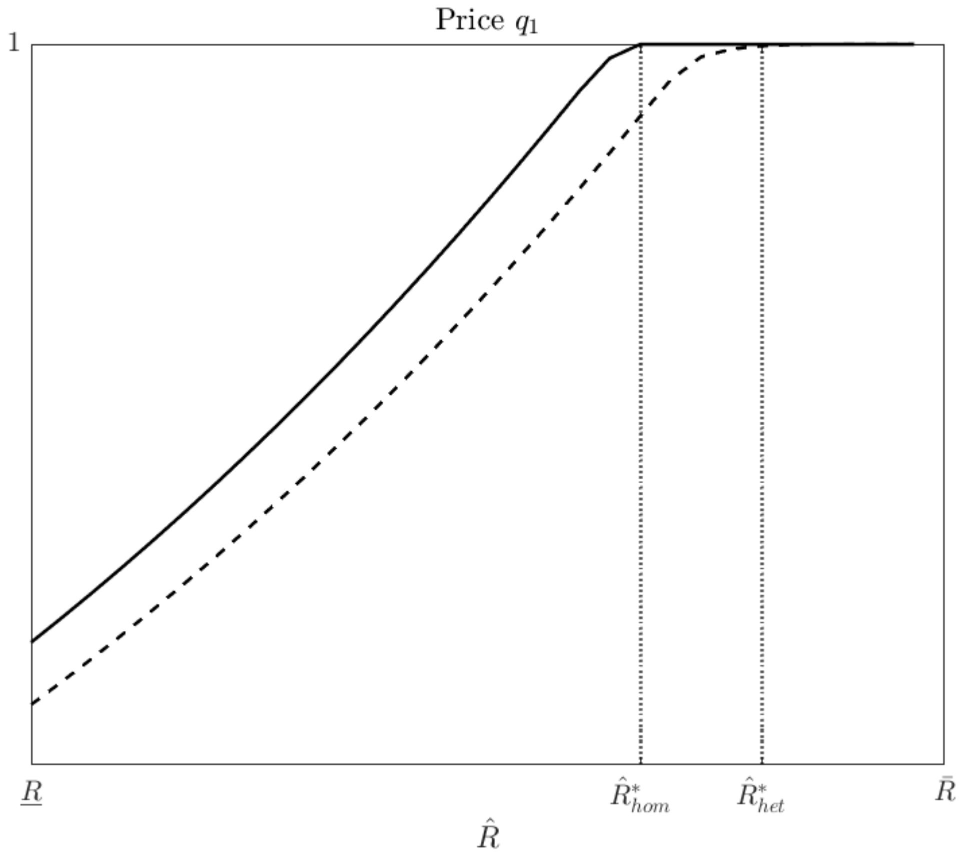


Figure 1.1: Mapping from $\hat{R}$ to q_1

Note: This figure shows the mapping from $\hat{R}$ to q_1 for the two cases when investors hold the homogeneous belief $F(R)$ or heterogeneous beliefs $\mathcal{F}$, respectively. The solid line refers to the homogeneous case, and the dashed line refers to the heterogeneous case. $\hat{R}^*_{hom}$ and $\hat{R}^*_{het}$ are thresholds as defined in Definition 3. The assumptions underlying this simulation are given in section 1.4.4.

We show that the threshold is lower if investors hold a homogeneous belief, compared to a setting of heterogeneous beliefs varying around it.

Theorem 2. *Consider two distinct populations with investors holding heterogeneous beliefs $\mathcal{F}$ in one, and the homogeneous belief $F(R)$ in the other. If the homogeneous belief is not more optimistic than any other belief in the heterogeneous case, i.e. $F^j(R) < F(R)$ for all R and at least one j, the probability of financial distress in the competitive equilibrium is higher under heterogeneity than under homogeneity, which is equivalent to*

$$\hat{R}^*_{het} > \hat{R}^*_{hom}.$$

Theorem 2 constitutes the second key result of our analysis. In an environment of heterogeneous beliefs, it is more likely that financial distress occurs. In general, it occurs whenever the realized return $\hat{R}$ is insufficient so that each investor could comply with her repayment obligations. If investors share a homogeneous belief, each $\hat{R} < \hat{R}^*_{hom}$ will constrain *all* investors. However, if beliefs are heterogeneous, it is enough that $\hat{R}$ is too low for *one* group to make everyone's borrowing constraint binding. In fact, under heterogeneity, the threshold $\hat{R}^*_{het}$ corresponds to the most optimistic type reaching the constraint, as it has built up the highest exposure to low returns.

We find that the most optimistic type – and all other types with it – is financially distressed even for higher returns compared to if they held a homogeneous belief. Consequently, under heterogeneity, financial distress occurs in even more favorable states of the world (as depicted in Figure 1.1), and is hence more likely. It rests on the presumption that the most optimistic belief is sufficiently off the ex post realization.

Hence, Theorem 2 highlights that financial distress may have an additional source. As is well known from the literature, a spiral of financial amplification can be initiated by adverse shocks sufficiently strong to drive excessively borrowing agents towards the constraint. Beyond that, we document that the dispersion of beliefs lays the ground for another source of distress, namely that some agents' beliefs deviate sufficiently from the true shock distribution.

1.3.4 Discussion

In the previous section, we have shown that belief heterogeneity increases the probability of financial distress, and how it affects the equilibrium collateral price. This price, in turn, is the main determinant of the financial friction, as it governs the tightness or slackness of the borrowing limit. Theorems 1 and 2 thus allow us to characterize the interaction of the collateral constraint and belief divergence, and to specify how different types of agents contribute to financial amplification. The mechanism emerging from this interaction has two features.

The first property is that heterogeneity of beliefs raises the likelihood of financial distress relative to the homogeneous benchmark. If investors have diverging expectations of future returns ex ante, some of these will differ from ex post realizations, which is sufficient to constrain all investors' borrowing. In contrast, under the homogeneous benchmark, when investors have rational expectations, the constraint binds only if the ex post realization is starkly adverse for all. Therefore, we conclude that belief disagreements facilitate the triggering of financial amplification.

The second feature refers to different investors' contributions to the financial amplification mechanism. Principally, during financial distress, optimistic and pessimistic investors drive collateral prices in opposing directions, as the former tend to sell, and the latter tend to purchase. However, we find an asymmetry of their contributions, attributing a larger impact to optimistic behavior. Hence, to distinguish the behavior of borrowing constraints in the presence of heterogeneous beliefs from the homogeneous benchmark, we must take into account how beliefs are distributed over the population.

We find that financial frictions are more severe under heterogeneity rather than homogeneity if the mean belief coincides with, or is more optimistic than the homogeneous belief. This implies that, so long as the belief distribution is symmetric around the homogeneous belief, or skewed towards more optimistic beliefs, heterogeneity exacerbates the financial amplification mechanism. The reason is that optimistic investors' (negative) contribution more than outweighs pessimistic investors' (positive) contribution.[12]

Our approach to financial amplification goes beyond the existing literature. These studies, presuming rational expectations, establish mechanisms where financial constraints bind in response to exogenous reductions of aggregate investment or aggregate net worth (Bianchi, 2011; Dávila and Korinek, 2018; Jeanne and Korinek, 2020). We extend this approach, and show that belief differences are sufficient to make such constraints binding. We may further quantify how market participants contribute to their tightness on the micro level.

In the following section, we turn to the welfare implications of the interaction mechanism between heterogeneous beliefs and financial frictions, which we have established hitherto.

1.4 Efficiency Analysis

We proceed by exploring the efficiency properties of our baseline economy. Given that the borrowing constraint is price-dependent, investors are subject to a pecuniary externality, as they do not internalize how their decisions affect other agents' individual welfare. We characterize these uninternalized welfare effects and their interplay with belief heterogeneity in the following section. Subsequently, we derive the constrained-efficient allocation as a welfare benchmark to contrast the competitive equilibrium, and develop optimal corrective policies that allow to implement it. Lastly, we quantify the

[12] Belief heterogeneity may mitigate financial amplification compared to the homogeneous benchmark, on the contrary, provided that the distribution is sufficiently skewed towards more pessimistic beliefs. The skewness would have to be large enough to reverse the relation of optimistic and pessimistic investors' influence on the collateral price. However, we argue that the presumption of a symmetric distribution is likely to prevail in financial markets. A range of studies provides both empirical and theoretical evidence that financial market participants' beliefs are distributed symmetrically, if not (close to) normally (Söderlind, 2009; Cvitanic and Malamud, 2011; Atmaz, 2014; Atmaz and Basak, 2016). Under this premise, extreme beliefs are either sufficiently improbable or counteracted by an equiprobable set of contrasting beliefs.

welfare impact of such policy interventions numerically.

1.4.1 Uninternalized Welfare Effects

The collateral price q_1 links individual choices and utilities across investors in two ways. First, it changes the value of investors' budgets in $t = 1$. Second, it determines the tightness of the borrowing constraints. Investors do not internalize these price effects. We use the terminology of Dávila and Korinek (2018) of *distributive* and *collateral* externalities.

Definition 4. *The uninternalized effects of changes in any type j's aggregate state variables* $\left\{\tilde{a}_0^j, \tilde{d}_0^j\right\}$ *on any i's individual welfare in periods* $t = 1, 2$ *can be written as*

$$\frac{\partial V_1^i}{\partial \tilde{a}_0^j} = \tilde{\lambda}_1^i D_{\tilde{a}_0^j}^i + \tilde{\eta}_1^i C_{\tilde{a}_0^j}^i$$
$$\frac{\partial V_1^i}{\partial \tilde{d}_0^j} = \tilde{\lambda}_1^i D_{\tilde{d}_0^j}^i + \eta_1^i C_{\tilde{d}_0^j}^i,$$

where $D_{\tilde{a}_0^j}^i$ *and* $D_{\tilde{d}_0^j}^i$ *are referred to as distributive externalities, and* $C_{\tilde{a}_0^j}^i$ *and* $C_{\tilde{d}_0^j}^i$ *are referred to as collateral externalities.*

(i) If investors hold heterogeneous beliefs $\mathcal{F}$*, distributive externalities are given by*

$$D_{\tilde{a}_0^j}^i = \frac{\partial q_1}{\partial \tilde{a}_0^j} \tilde{l}_1^i,$$
$$D_{\tilde{d}_0^j}^i = \frac{\partial q_1}{\partial \tilde{d}_0^j} \tilde{l}_1^i,$$

and collateral externalities are given by

$$C_{\tilde{a}_0^j}^i = \phi \frac{\partial q_1}{\partial \tilde{a}_0^j} \left(\tilde{a}_0^i - \tilde{l}_1^i\right),$$
$$C_{\tilde{d}_0^j}^i = \phi \frac{\partial q_1}{\partial \tilde{d}_0^j} \left(\tilde{a}_0^i - \tilde{l}_1^i\right).$$

(ii) If investors hold the homogeneous belief $F(R)$*, distributive externalities are zero, and collateral externalities are given by*

$$C_{\tilde{a}_0} = \phi \frac{\partial q_1}{\partial \tilde{a}_0} \bar{a},$$
$$C_{\tilde{d}_0} = \phi \frac{\partial q_1}{\partial \tilde{d}_0} \bar{a}.$$

Distributive effects describe the price-induced redistribution between trading agents, altering their marginal rates of substitution. Collateral effects measure the price-induced change in an agent's capacity to borrow.

In an environment of heterogeneous beliefs, it turns out that, the more optimistic investors are, the more likely it is that they will sell claims on the asset in $t = 1$ ($\tilde{l}_1^j \geq 0$). Accordingly, more pessimistic investors will more probably enter the market as buyers ($\tilde{l}_1^j < 0$). The reason is that a group's exposure to adverse states, reflected by its position $\tilde{a}_0^j$, is a monotonic function of beliefs (see Proposition 3). We use this fact, as well as Proposition 2, to characterize the direction of distributive and collateral externalities.

Proposition 4.

(i) If investors hold heterogeneous beliefs $\mathcal{F}$, distributive externalities have a non-positive sign for period-1-sellers, i.e. $D^i_{\tilde{a}_0^j} \leq 0$ and $D^i_{\tilde{d}_0^j} \leq 0$ if $\tilde{l}_1^i \geq 0$, and a non-negative sign for period-1-buyers, i.e. $D^i_{\tilde{a}_0^j} \geq 0$ and $D^i_{\tilde{d}_0^j} \geq 0$ if $\tilde{l}_1^i \leq 0$. If investors hold the homogeneous belief $F(R)$, distributive externalities are zero.

(ii) Collateral externalities have a non-positive sign for any type i of investors, and irrespective of beliefs, i.e. $C^i_{\tilde{a}_0^j} \leq 0$ and $C^i_{\tilde{d}_0^j} \leq 0$ for each i.

Distributive externalities are signed reflective on the fact that a decline of the equilibrium price q_1 benefits buyers and harms sellers in $t = 1$. Collateral externalities, in turn, are unambiguously adverse to each type of agent, as more investment and borrowing reduce the collateral value, cutting any investor's borrowing capacity.

Ultimately, we evaluate the welfare implications of the interaction mechanism between beliefs and the equilibrium price q_1, which we have established in Theorem 1.

Proposition 5.

(i) Let investors hold heterogeneous beliefs $\mathcal{F}$.

(1) Let further $G^j(R)$ be the direction of a perturbation of type-j investors' belief $F^j(R)$, and beliefs $F^i(R)$ be constant for all $i \neq j$. If the perturbation makes investors of type j more optimistic (pessimistic), both distributive and collateral externalities of any type-i investor are larger (smaller) in absolute value, i.e., for each $i \neq j$ and $x \in \{a, d\}$,

$$\left| \frac{\delta D^i_{\tilde{x}_0^j}}{\delta F^j} \cdot G^j \right| \begin{cases} \geq 0, & G^j(R) \leq 0 \\ \leq 0, & G^j(R) \geq 0 \end{cases} \quad \text{and} \quad \frac{\delta C^i_{\tilde{x}_0^j}}{\delta F^j} \cdot G^j \begin{cases} \leq 0, & G^j(R) \leq 0 \\ \geq 0, & G^j(R) \geq 0 \end{cases}.$$

(2) Let further $G^j(R) < 0 < G^k(R)$ with $|G^j(R)| = \left|G^k(R)\right|$ for all R be the directions of two perturbations that make investors of type j more optimistic, and investors of type k more pessimistic by the same magnitude. The uninternalized welfare effects under the perturbation with direction $G^j(R)$ are stronger than those under the perturbation with direction $G^k(R)$, i.e., for each $i \neq j, k$ and $x \in \{a, d\}$,

$$\left|\frac{\delta D^i_{\tilde{x}^j_0}}{\delta F^j} \cdot G^j\right| \geq \left|\frac{\delta D^i_{\tilde{x}^k_0}}{\delta F^k} \cdot G^k\right| \quad \text{and} \quad \left|\frac{\delta C^i_{\tilde{x}^j_0}}{\delta F^j} \cdot G^j\right| \geq \left|\frac{\delta C^i_{\tilde{x}^k_0}}{\delta F^k} \cdot G^k\right|.$$

(ii) Let investors hold the homogeneous belief $F(R)$, and let $G(R)$ be the direction of a perturbation. If the perturbation makes investors more optimistic (pessimistic), collateral externalities are larger (smaller) in absolute value, i.e., for $x \in \{a, d\}$

$$\frac{\delta C_{\tilde{x}_0}}{\delta F} \cdot G \begin{cases} \leq 0, & G(R) \leq 0 \\ \geq 0, & G(R) \geq 0 \end{cases}.$$

Proposition 5 describes the welfare effects associated with the interaction of beliefs and the equilibrium price q_1. It states that more optimistic types exerting downward pressure on the collateral price, due to large investment and borrowing, impose more intense negative distributive externalities on sellers ($\tilde{l}^i_1 > 0$), and more intense positive ones on buyers ($\tilde{l}^i_1 < 0$). In contrast, more pessimistic types' choices have an increasing impact on the collateral price, by this causing the reverse response of distributive externalities.

By the same logic, collateral externalities, being non-positive in general, turn out to be more or less pronounced in the case of more optimistic or pessimistic groups, respectively. This result holds true analogously in the homogeneous case.

Importantly, the asymmetry between optimistic and pessimistic investors' influence on q_1 translates into asymmetric welfare effects, as we formalize in statement (2) of part (i). Since the price responds more markedly to optimistic than pessimistic behavior, the former further dominates in welfare terms. If the two groups j's and k's beliefs are made more optimistic and pessimistic to the same extent, respectively, any further type i's group-wide welfare losses from j's high investment and borrowing exceed the gains from k's precaution.

1.4.2 Constrained Efficiency

Investors do not internalize the distributive or collateral side effects of their behavior which materialize through the collateral price q_1. These externalities render the competitive equilibrium allocation inefficient. To evaluate its welfare properties, we employ the concept of constrained efficiency.

The constrained-efficient allocation solves the problem of a constrained social planner who chooses investment and borrowing in period $t = 0$, while leaving all later choices to private agents. Specifically, she maximizes social welfare subject to all resource constraints, technological constraints, market clearing conditions, and financial frictions, respecting the competitive equilibrium price formation (see equation (1.6)).

Social welfare is evaluated by aggregating investors' expected lifetime utilities, and applying arbitrary Pareto weights $\omega = \{\omega^j\}_{j\in\{1,\dots,J\}}$. A relevant question in this setting is the planner's belief (Blume et al., 2018; Dávila, 2023; Kim and Kim, 2021). If we assigned a specific belief to the planner, she would naturally disagree with investors upon their beliefs. Abstracting from this trivial motive of correction, we aim at isolating ex ante corrective policies related to the financial friction, and, thus, make the following assumption.

Assumption 2. *The constrained social planner has no superior information, and respects individual beliefs for each type j.*

We solve the following social planner problem.

$$\begin{aligned}
\max_{\{\tilde{c}_0^j,\tilde{a}_0^j,\tilde{d}_0^j\}_{j\in\{1,\dots J\}}} \quad & \sum_{j=1}^{J} \omega^j s^j \left[u\left(\tilde{c}_0^j\right) + E^j \left[V_1^j \left(\tilde{a}_0^j, \tilde{d}_0^j | \tilde{a}_0, \tilde{d}_0 \right) \right] \right] \quad \text{s.t.} \\
(\tilde{\lambda}_0) \quad & \sum_{j=1}^{J} s^j \tilde{c}_0^j = \sum_{j=1}^{J} s^j \left[e + \tilde{d}_0^j \right] \\
(\tilde{\psi}) \quad & \sum_{j=1}^{J} s^j \tilde{a}_0^j = \bar{a}.
\end{aligned} \tag{1.12}$$

With the first order conditions for consumption, $\tilde{\lambda}_0 = \omega^j u'\left(\tilde{c}_0^j\right)$, the planner's optimality conditions for each j are

$$0 = E^j \left[R u'\left(\tilde{c}_1^j\right) + u'\left(\tilde{c}_2^j\right) + \tilde{\eta}_1^j \phi q_1 \right] - \frac{\tilde{\psi}}{\omega^j} + \sum_{i=1}^{J} \frac{\omega^i}{\omega^j} \frac{s^i}{s^j} E^i \left[D^i_{\tilde{a}_0^j} u'\left(\tilde{c}_1^i\right) + \tilde{\eta}_1^i C^i_{\tilde{a}_0^j} \right] \tag{1.13}$$

$$0 = u'\left(\tilde{c}_0^j\right) - E^j \left[u'\left(\tilde{c}_1^j\right) \right] + \sum_{i=1}^{J} \frac{\omega^i}{\omega^j} \frac{s^i}{s^j} E^i \left[D^i_{\tilde{d}_0^j} u'\left(\tilde{c}_1^i\right) + \tilde{\eta}_1^i C^i_{\tilde{d}_0^j} \right]. \tag{1.14}$$

We can now define the constrained-efficient allocation.

Definition 5. *The period-0 allocation* $\left\{\tilde{c}_0^j, \tilde{a}_0^j, \tilde{d}_0^j\right\}_{j\in\{1,\dots J\}}$ *is constrained-efficient if and only if there are shadow prices* $\tilde{\lambda}_0$, $\tilde{\psi}$, $\{\tilde{\eta}_1^j\}_{j\in\{1,\dots J\}}$, *and a set of Pareto weights* $\{\omega^j\}_{j\in\{1,\dots J\}}$ *such that it satisfies the price relation* (1.6) *for each* j*, the market clearing condition* (1.10)*, and the resource constraint* (1.12)*, as well as the equations* (1.13)*,* (1.14)*, and* $\tilde{\lambda}_0 = \omega^j u'\left(\tilde{c}_0^j\right)$ *for each* j*, given population shares* s *and beliefs* $\mathcal{F}$*.*

Equations (1.13) and (1.14) differ from the competitive equilibrium conditions (1.8) and (1.9) through the aggregate terms of externalities on the right-hand side. They indicate formally that the competitive allocation is not constrained-efficient, whereas the social planner takes distributive and collateral externalities into account. Furthermore, she accounts for market clearing in $t = 0$, represented by the multiplier $\tilde{\psi}$.

1.4.3 Optimal Corrective Policies

The constrained-efficient allocation can be achieved in a decentralized market using a set of adequate policy instruments. We start out by characterizing optimal corrective taxes under both heterogeneous and homogeneous beliefs. We contrast a system of non-linear taxes under heterogeneity with a simple linear Pigouvian tax. The latter allows us to quantify welfare differences between our approach and previous policy proposals in the following section.

Decentralization. To decentralize the constrained-efficient allocation, we provide the social planner with access to Pigouvian taxes, available to manipulate agents' investment and borrowing decisions, and lump sum transfers. These instruments satisfy the conditions stated in the following Proposition.

Proposition 6.

(i) If investors hold heterogeneous beliefs $\mathcal{F}$*, the social planner can implement the constrained-efficient allocation by taxing investment and borrowing, satisfying*

$$\tau_a^j = \operatorname{sgn}\left(\bar{a} - \tilde{a}_0^j\right)\left(s^j q_0 \tilde{\lambda}_0\right)^{-1} \sum_{i=1}^{J} \omega^i s^i E^i \left[D^i_{\tilde{a}_0^j} u'\left(\tilde{c}_1^i\right) + \tilde{\eta}_1^i C^i_{\tilde{a}_0^j}\right] \tag{1.15}$$

$$\tau_d^j = -\operatorname{sgn}\left(\tilde{d}_0^j\right)\left(s^j \tilde{\lambda}_0\right)^{-1} \sum_{i=1}^{J} \omega^i s^i E^i \left[D^i_{\tilde{d}_0^j} u'\left(\tilde{c}_1^i\right) + \tilde{\eta}_1^i C^i_{\tilde{d}_0^j}\right] \tag{1.16}$$

for each j, and rebating revenues through type-specific lump sum transfers $T^j = \tau_a^j \operatorname{sgn}\left(\bar{a} - \tilde{a}_0^j\right) q_0 \left(\bar{a} - \tilde{a}_0^j\right) + \tau_d^j \operatorname{sgn}\left(\tilde{d}_0^j\right) \tilde{d}_0^j$.[13]

(ii) If investors hold the homogeneous belief $F(R)$, the social planner can implement the constrained-efficient allocation by taxing borrowing, satisfying

$$\tau_d = -\tilde{\lambda}_0^{-1} E\left[\tilde{\eta}_1 C_{\tilde{d}_0}\right], \tag{1.17}$$

and rebating revenues through lump sum transfers $T = \tau_d \tilde{d}_0$, while the tax on investment is arbitrary.

In the heterogeneous case, our optimal Pigouvian taxes are characterized by a range of sufficient statistics related to distributive and collateral externalities, aggregated in the squared brackets in equations (1.15) and (1.16).[14]

Three components determine distributive effects. First, when price movements induce a redistribution of funds between period-1-buyers and -sellers, this affects their marginal rates of substitution. Second, price movements themselves measure the intensity of redistribution. Third, the direction of redistribution depends on whether an investor is a seller ($\tilde{l}_1^j > 0$) or a buyer ($\tilde{l}_1^j < 0$) in $t = 1$. The latter two components are captured by the distributive externalities $D^i_{\tilde{a}_0^j}$ and $D^i_{\tilde{d}_0^j}$, given in Definition 4.

Collateral effects are driven by another three components. First, the multiplier $\tilde{\eta}_1^j$ measures the welfare gain (loss) when the constraint is relaxed (tightened) by one unit. Second, price movements describe the change in an investor's borrowing capacity per unit of collateral, whose total magnitude available matters third. The last two elements are incorporated in the collateral externalities $C^i_{\tilde{a}_0^j}$ and $C^i_{\tilde{d}_0^j}$ from Definition 4.

If, however, investors hold the homogeneous and rational belief, these sufficient statistics turn out to be vastly simplified. Since investors do not trade the asset under homogeneity, the social planner cannot manipulate investment decisions. The resulting tax on investment is arbitrary. Moreover, for the very same reason, distributive externalities are zero, rendering the tax on borrowing responsive solely to collateral externalities (see equation (1.17)).

Notably, the instruments derived in Proposition 6 may well be subsidies instead of taxes, depending on the extent of externalities induced by type j, and its specific choices of investment and borrowing. Taxes/subsidies turn out to be zero only if *all* investors expect their collateral constraints to be slack. To put it another way, it suffices that one group of investors expects to be constrained to let taxes/subsidies take on either

[13] We use a sign operator for an easier interpretation of taxes and subsidies, given the fact that investors can take short and long positions in the asset, as well as borrow and save.

[14] For a more detailed description of sufficient statistics, see Dávila and Korinek (2018).

sign for the entire population. We will return to the signing of policy instruments in the next section.

Incentive compatibility. In an environment of heterogeneous agents, whose type is their private information, corrective policies may not be incentive-compatible. The instruments we have derived in Proposition 6 are type-specific, raising the question of knowledge required by the social planner to impose taxes in an incentive-compatible way.

Importantly, the optimal non-linear taxes in equations (1.15) and (1.16) incorporate no more than publicly known objects. To be precise, to set group-specific taxes, the social planner must be informed about the set of beliefs $\mathcal{F}$ in the economy, each type's respective population share s^j, as well as investment and borrowing choices $\tilde{a}_0$ and $\tilde{d}_0$, which are publicly observable in the market. Since the latter are monotonic functions of beliefs, as we have shown in Proposition 3, they perfectly reveal any investor's belief.

Therefore, the constrained-efficient allocation can be implemented by means of the following system of non-linear Pigouvian taxes.

Theorem 3. *If investors hold heterogeneous beliefs $\mathcal{F}$, the social planner can implement the constrained-efficient allocation by taxing investment and borrowing according to the tax system $(\tilde{\tau}_a, \tilde{\tau}_d)$, satisfying*

$$\tilde{\tau}_a : \quad \tilde{a}_0^k \mapsto \tilde{\tau}_a(\tilde{a}_0^k) \; s.t. \; \tilde{\tau}_a(\tilde{a}_0^k) = \begin{cases} RHS \text{ of } (1.15) & \text{if } \tilde{a}_0^k = \tilde{a}_0^j \text{ for any } j \text{ with } \tilde{a}_0^j \in \tilde{a}_0 \\ \infty & \text{if } \tilde{a}_0^k \notin \tilde{a}_0 \end{cases} \tag{1.18}$$

$$\tilde{\tau}_d : \quad \tilde{d}_0^k \mapsto \tilde{\tau}_d(\tilde{d}_0^k) \; s.t. \; \tilde{\tau}_d(\tilde{d}_0^k) = \begin{cases} RHS \text{ of } (1.16) & \text{if } \tilde{d}_0^k = \tilde{d}_0^j \text{ for any } j \text{ with } \tilde{d}_0^j \in \tilde{d}_0 \\ \infty & \text{if } \tilde{d}_0^k \notin \tilde{d}_0, \end{cases} \tag{1.19}$$

and corresponding lump sum transfers.

The essential point of Theorem 3 is that the social planner does not rely on knowledge of individual beliefs. The peculiar nature of our optimal Pigouvian taxes ensures that the constrained-efficient allocation is indeed decentralizable, even in a setting of heterogeneous beliefs.

Our results on optimal corrective policies give rise to several issues linked to the welfare implications of the interplay between belief heterogeneity and the financial friction. First, analyzing the responses of group-specific taxes/subsidies to variations of beliefs is informative on different types' contributions to changes in social welfare. Second, we seek to compare the efficiency properties of our economy under homogeneity and heterogeneity of beliefs. Third, it is enlightening to evaluate how the probability of financial distress is altered through a planner intervention of the kind sketched above.

Moreover, we aim at quantifying the welfare impact of the non-linear tax instruments we propose in contrast to a linear Pigouvian tax on borrowing. The latter is a standard macroprudential instrument which has gained much attention in the literature (Bianchi, 2011; Dávila and Korinek, 2018; Jeanne and Korinek, 2019, 2020). In our model, it corresponds to equation (1.17), being a tax on borrowing calibrated to the case of homogeneous and rational expectations.

Examining these questions is analytically intractable. The clear signing of tax instruments depends on the specific belief distribution, which we have kept general thus far. To gain insights into the welfare implications of our policy proposals, we provide a numerical application of our model in the following.

1.4.4 Numerical Application

The numerical analysis requires a simplified version of our model. In this section, we first describe the simplifications applied to make the baseline model numerically tractable, and briefly characterize the resulting equilibrium allocations, prices, and, importantly, optimal corrective policies for different levels of belief heterogeneity. Subsequently, we quantify the welfare implications of such policies, and assess how these interventions impact the probability of financial distress.

Simplifications. Suppose the economy is populated by two groups of investors, called optimists and pessimists, indexed by o and p. We let both groups be of equal mass, i.e. $s^o = s^p = 1$, and differ in terms of their return expectations, i.e. $E^o[R] > E^p[R]$. Furthermore, there are only two states of the world. To be precise, R may take on either a *good* or a *bad* value, denoted by $R^g > R^b$.

We choose parameters in line with the assumptions underlying our theoretical analysis, simulating equilibria with significant trade volumes and binding financial constraints. Table 1.1 summarizes the parameter values chosen in the application.

The parameter ϕ, capturing the margin requirement for borrowing, is selected following Mendoza (2002) and Bianchi (2011), who suggest that debt is required to not exceed a fraction of 30 to 40 percent of tradable assets. Averaging these values, we set $\phi = 0.35$. The two states R^g and R^b are chosen with the aim to make trading incentives strong enough, which, in turn, ensures a significant trade volume. This condition is

Parameter		**Value**
Margin requirement	ϕ	0.35
Good state	R^g	2
Bad state	R^b	0
Initial endowment of consumption goods	e	1
Initial asset endowment	$\bar{a}$	2
Risk aversion	A	0.5
Heterogeneity step	μ	0.025
Initial belief	π^g	0.5

Table 1.1: Parameter Values

met for $R^g = 2$ and $R^b = 0$. For the same argument, we set initial endowments of consumption goods e and assets $\bar{a}$ to $e = 1$ and $\bar{a} = 2$, and choose a moderate degress of risk aversion $A = 0.5$.

Heterogeneity itself is defined as the linear distance between the probabilities that the two types assign to the good state, i.e. $\pi^{j,g} = 1 - \pi^{j,b}$. We increase this distance symmetrically by N steps of size $\mu = 0.025$ (see Simsek (2013) for comparison). The multiples N thus serve as a measure of belief heterogeneity. The benchmark case is a population with homogeneous beliefs, where $\pi^{o,g} = \pi^{p,g} \equiv \pi^g$, which we set to $\pi^g = 0.5$. Finally, the two types' beliefs at any given level of heterogeneity N are given by

$$E^o[R] = (\pi^g + N\mu)R^g + (\pi^b - N\mu)R^b$$
$$E^p[R] = (\pi^g - N\mu)R^g + (\pi^b + N\mu)R^b.$$

Notably, we let the social planner apply Pareto weights ω such that the constrained-efficient allocation replicates the unconstrained competitive allocation, i.e. when the collateral constraints are slack. This choice ensures that the simulated corrective interventions by the planner are solely related to inefficiencies from the financial friction, but not to differences in the aggregation of social welfare.

Allocations, prices, and corrective policies. Figure 1.2 displays the responses of key variables to different levels of heterogeneity. Specifically, it shows the equilibrium values of period-0 investment and borrowing, the period-1 price q_1 – the main determinant of the collateral constraint – as well as taxes and the externalities therein. The two beliefs diverge increasingly the further one follows the x-axis. The blue and red lines refer to the optimists and pessimists, respectively. Solid lines refer to variables from the competitive equilibrium, while dotted lines refer to the constrained-efficient equilibrium.

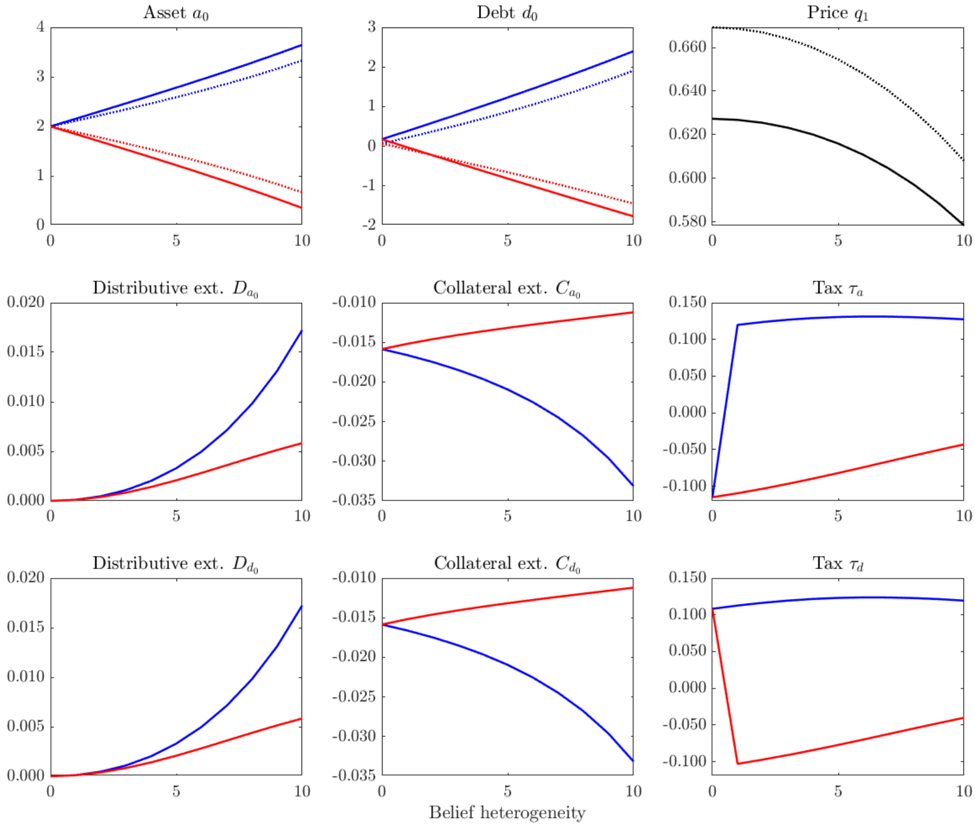


Figure 1.2: Equilibrium Allocations, Prices, and Optimal Corrective Policies

Note: The three upper panels show period-0 choices of investment, and borrowing, as well as the period-1 asset price. The three middle panels show optimal taxes on investment, and aggregate distributive and collateral externalities therein. The three middle panels show optimal taxes on borrowing, and aggregate distributive and collateral externalities therein. The blue and red lines refer to the optimistic and the pessimistic type, respectively. Solid lines refer to variables from the competitive equilibrium, while dotted lines refer to the constrained-efficient equilibrium. Each number on the x-axis relates to the N-th heterogeneity step, where $N = 0$ stands for the benchmark case of homogeneous beliefs.

The top-left and top-central panels illustrate the monotonicity of period-0 investment and borrowing in beliefs. Starting from a no-trade equilibrium under homogeneous beliefs, where investors keep their initial asset position constant, investment and borrowing increase (decrease) the more optimistic (pessimistic) they become. Contrasting the competitive allocation, the social planner induces agents to trade, borrow, and save less. Importantly, the planner reduces optimists' borrowing by more than pessimists' saving, reflecting the asymmetry between optimistic and pessimistic types' contributions to financial distress, formalized in Theorem 1.

In the top-right panel, this asymmetry becomes evident in the response of the equilibrium price q_1 to increasing belief heterogeneity. Given that the influence of optimistic behavior is dominant, the equilibrium price declines even though we have not altered the economy's mean belief, but made the two types more heterogeneous in a symmetric manner. Furthermore, the social planner improves on the competitive allocation by sustaining a higher price, alleviating the tightness of the financial friction.

The panels in the second row of Figure 1.2 depict the aggregate distributive and collateral externalities associated with each type's investment, and the corresponding corrective policies, formalized in equation (1.15). To achieve constrained efficiency, the planner taxes investment by optimists ($\tau_a^o > 0$), and subsidizes asset purchases by pessimists ($\tau_a^p < 0$).

The interplay of aggregate distributive and collateral externalities determine the signs of the instruments. The tax on optimists' investment is driven by negative collateral externalities clearly outweighing positive distributive externalities. The latter arise because pessimists, buying claims in $t = 1$, benefit from the price decline induced by optimists' behavior. However, as the collateral price continues falling with increasing heterogeneity, optimists pass over more intense collateral externalities to pessimists.

Pessimists, in contrast, are subsidized because their cautious investment decisions tend to mitigate the price decline, benefiting optimists' budget in $t = 1$, and reducing collateral externalities. Since they behave with more precaution the more pessimistic they become, the social planner is less inclined to correct their behavior, and the subsidy reverts to zero.

The lower panels of Figure 1.2 refer to aggregate externalities associated with borrowing and saving, and the respective policy instruments, captured by equation (1.16). By the same mechanisms as for the correction of investment, borrowing by optimists is increasingly taxed ($\tau_d^o > 0$), and borrowing by pessimists is subsidized ($\tau_d^p < 0$).[15] If the two types of investors hold the homogeneous beliefs, their borrowing is slightly taxed.

Welfare effects. Thus far, we have qualified both the direction and the extent of corrective taxes. In the following, we turn to the normative question of how the Pigouvian correction translates into social welfare. We are particularly interested in measuring welfare gains from the non-linear tax policy as opposed to a linear Pigouvian tax system, which is the most frequently proposed instrument in the literature on pecuniary externalities and prudential policy responses, (Bianchi, 2011; Dávila and Korinek, 2018; Jeanne and Korinek, 2019, 2020). This literature typically presumes rational expecta-

[15] Aggregate distributive and collateral externalities from borrowing turn out to be equal to those from investment in this example due to our assumption $R^b = 0$. In this case, price effects are identical, and so are type-specific externalities (see Definition 4).

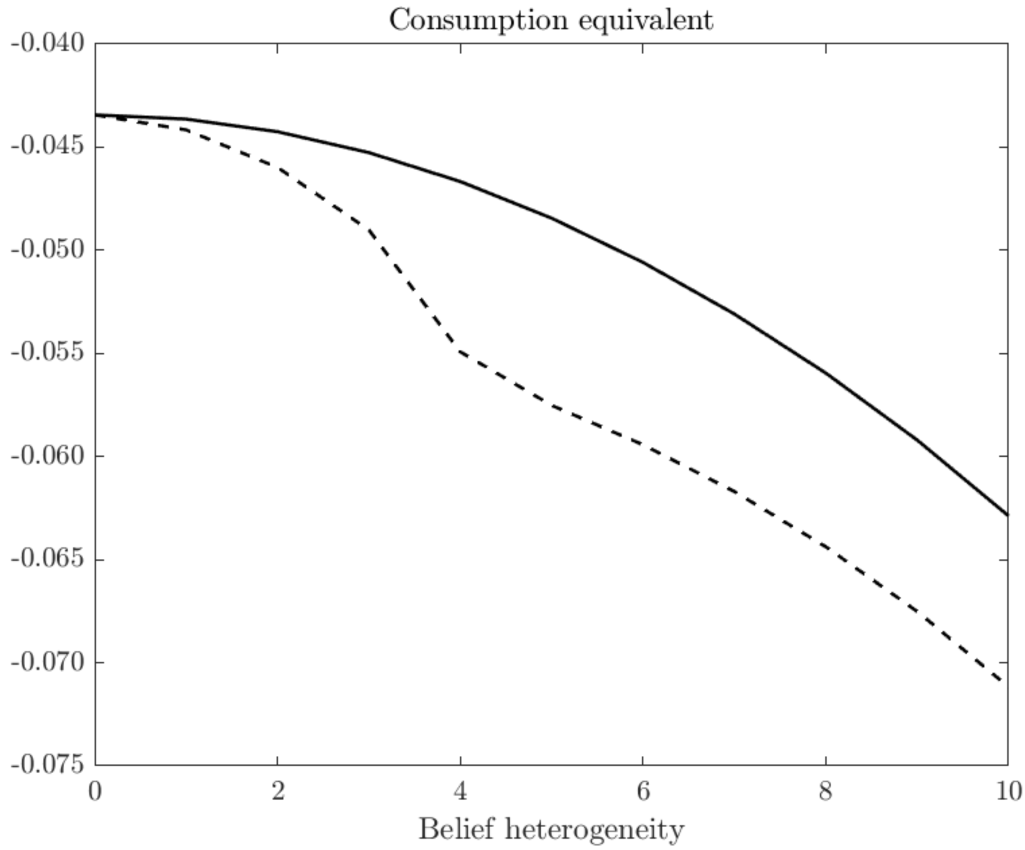


Figure 1.3: Welfare Effects of Linear and Non-Linear Corrective Taxes

Note: This figure shows the consumption equivalents of two types of allocations relative to the unconstrained competitive allocation. The solid line refers to constrained-efficient allocations, which are implemented by means of the system of non-linear taxes proposed in Theorem 3. The dotted line refers to allocations implemented by means of the system of linear taxes proposed in part (*ii*) of Proposition 6. Each number on the x-axis relates to the N-th heterogeneity step, where $N = 0$ stands for the benchmark case of homogeneous beliefs.

tions.

In our model, this policy corresponds to the system of linear corrective taxes in the case of homogeneous beliefs (see part (*ii*) of Proposition 6). This is when investors feature rational expectations, and the social planner optimally taxes borrowing, while any correction of investment decisions is ineffective. Figure 1.3 displays the welfare effects of this policy in comparison to the non-linear tax system from above.

We employ consumption equivalents relative to the unconstrained competitive allocation, which is when no policy intervention is required, as an ex ante social welfare measure. In Figure 1.3, the solid line depicts consumption equivalents of allocations with non-linear corrective taxes, while the dotted line refers to allocations with linear corrective taxes. Each point on the x-axis indicates a specific belief distribution, with beliefs becoming increasingly heterogeneous along the axis.

We find significant welfare gains from non-linear policies over linear Pigouvian taxes. The planner's intervention contains welfare losses at a level of about four to six percent relative to the unconstrained economy. However, if linear taxes are applied to a heterogeneous population, welfare is well below. Corresponding allocations result in welfare losses which are by up to 14 percent larger than compared to allocations with non-linear policies.

Probability of financial distress. The last numerical exercise we provide is related to the above evaluation how probable financial distress is in the competitive equilibrium. We have found that belief disagreements across investors do indeed raise the probability that financial distress occurs, relative to the case of rational and homogeneous beliefs. We repeat the simulation from above, but further account for the constrained-efficient allocation. To that end, we first define the lowest possible realization of R such that collateral constraints in the constrained-efficient allocation are slack.

Definition 6. *Let* $\hat{R}^{**}_{het} \equiv \min\left\{\hat{R} \mid \tilde{\eta}^j_1 = 0 \text{ for all } j\right\}$ *and* $\hat{R}^{**}_{hom} \equiv \min\left\{\hat{R} \mid \tilde{\eta}_1 = 0\right\}$ *be the lowest possible realizations of* R *such that the borrowing constraints are slack in the constrained-efficient equilibrium if investors hold heterogeneous beliefs* $\mathcal{F}$ *or the homogeneous belief* F*, respectively.*

Figure 1.4 illustrates the mapping from the realization $\hat{R}$ to q_1 for both the competitive and the constrained-efficient equilibrium. The probability of financial distress is indeed lower under constrained efficiency than in the competitive equilibrium. By manipulating investors' behavior through non-linear taxes, the social planner manages to reduce the thresholds of $\hat{R}$, implying that financial distress in the constrained-efficient equilibrium would only arise in markedly unfavorable states. Our previous finding that financial distress is generally less likely under the homogeneous belief than under heterogeneity is further robust to the planner intervention.

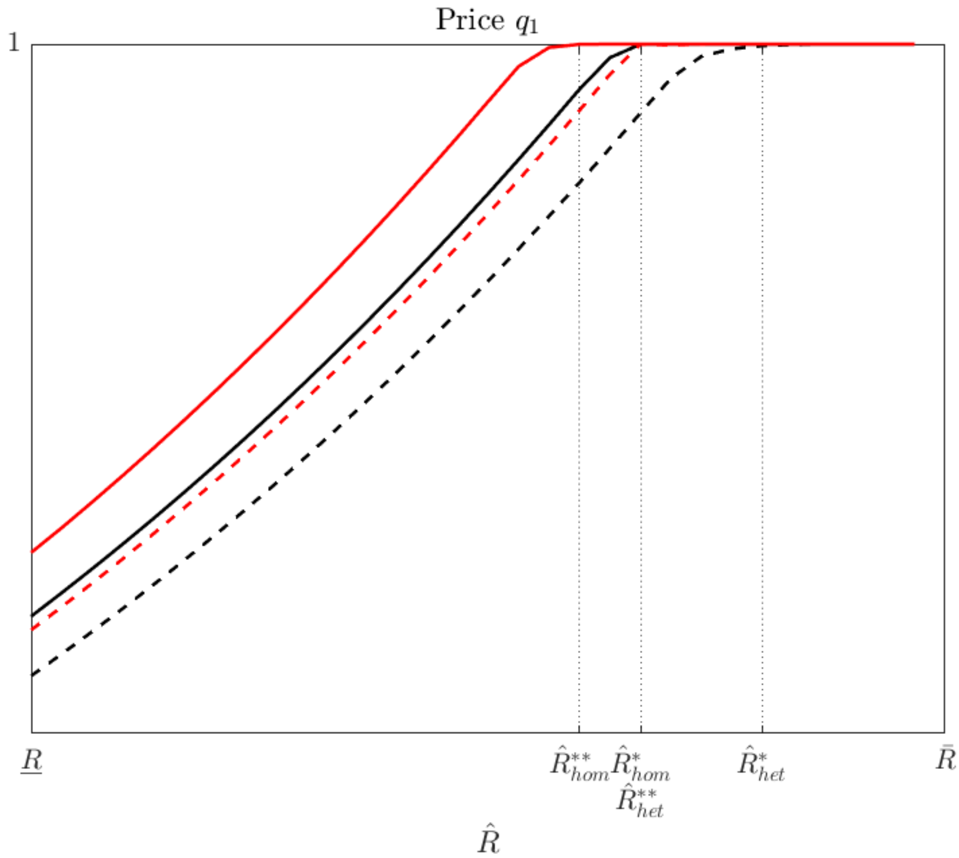


Figure 1.4: Mapping from $\hat{R}$ to q_1

Note: This figure shows the mapping from $\hat{R}$ to q_1 for the two cases when investors hold the homogeneous belief $F(R)$ or heterogeneous beliefs $\mathcal{F}$, respectively. Solid lines refer to the homogeneous case, and dashed lines refer to the heterogeneous case. Black lines refer to the competitive equilibrium, and red lines refer to the constrained-efficient equilibrium. $\hat{R}^{*}_{hom}$, $\hat{R}^{*}_{het}$, $\hat{R}^{**}_{hom}$, and $\hat{R}^{**}_{het}$ are thresholds as defined in Definitions 3 and 6.

1.5 Final Remarks

This paper presents a theoretical framework to study the contributions of economic agents to financial distress, being the basis on which optimal Pigouvian policies are designed. We build on a model incorporating financial frictions, and enrich it by the heterogeneity of beliefs across economic agents. The framework is employed to analyze the competitive equilibrium, its sensitivity to changes in the underlying set of beliefs, as well as its efficiency properties. We derive optimal corrective policies, which are furthermore quantified in a numerical application.

Our analysis puts forward three key findings. First, we show that, conditional on their beliefs, investors make differentiated contributions to financial distress, where relatively more optimistic agents have an overproportional and decreasing impact on the collateral price. Second, it turns out that financial distress is generally more likely in an economy populated by agents with heterogeneous beliefs, compared to the homogeneous case. Third, we find that a constrained-efficient allocation can be implemented through a system of non-linear Pigouvian taxes, which proves to generate considerable welfare gains over the linear policy advocated by previous articles.

These results add to the literature on financial crises in several ways. We characterize explicitly how financial market participants contribute to distress states. Moreover, in our setting, financial constraints may be binding through ex ante return expectations sufficiently off the ex post realization. This differs from former studies, focusing on financial distress in response to aggregate shocks to investment or net worth. Hence, our framework formalizes a further source of financial distress. Ultimately, our policy proposal improves on linear Pigouvian taxes in an economy featuring heterogeneity of beliefs. The latter point is especially relevant when studying optimal financial regulation in booms and busts, which typically go along with high belief divergence and fluctuations.

Our work lays the ground for further research. Whereas we study optimal ex ante policies in a prudential sense, it may be worthwhile examining optimal ex post policies, such as central bank liquidity injections, under belief heterogeneity. In addition, several types of financial frictions are considered in the literature on prudential policies. The collateral constraints used in this paper link debt issuance to market-valued collateral. However, pecuniary externalities and corrective policies have further been studied in environments with flow constraints, relating to household income or firm cash flows. Their interaction with belief disagreements must still be examined. Ultimately, our three period model may be extended to a dynamic framework, allowing for a more profound quantitative exploration of the effects documented in this paper.

1.A Appendix

Proof of Proposition 1

Models with price-dependent collateral constraints like ours bear the risk that equilibrium prices do not exist. The reason is that these models face downward-sloping supply functions. Constraint agents must sell more if the collateral price is low, but less if it high, and the constraint is less tight.

Existence. We first prove the existence of the equilibrium price. Let

$$S\left(q_1\right)=\sum_{j=1}^{J}\mathbb{1}_{\left\{\tilde{l}_1^j(q_1)>0\right\}}s^j\tilde{l}_1^j\left(q_1\right) \tag{1.20}$$

denote the supply of claims as a function of q_1. Analogously, define demand as

$$D\left(q_1\right)=-\sum_{j=1}^{J}\mathbb{1}_{\left\{\tilde{l}_1^j(q_1)<0\right\}}s^j\tilde{l}_1^j\left(q_1\right). \tag{1.21}$$

Let $D\left(q_1\right)$ and $S\left(q_1\right)$ be continuous and differentiable functions on the interval $(0,1]$. Note that $S\left(q_1\right)$ is bounded from above for any q_1. This follows from the fact that investors cannot sell more claims than they possess, i.e. $\tilde{l}_1^j\leq\tilde{a}_0^j$, and, hence, for any q_1

$$S\left(q_1\right)=\sum_{j=1}^{J}\mathbb{1}_{\left\{\tilde{l}_1^j(q_1)>0\right\}}s^j\tilde{l}_1^j\left(q_1\right)\leq\sum_{j=1}^{J}s^j\tilde{a}_0^j=\bar{a}.$$

Specifically, it follows that $\lim\limits_{q_1\to0}S\left(q_1\right)\leq\bar{a}$.

We consider two cases when characterizing the demand curve. First, if demand is zero, there is still excess supply. According to the Walrasian equilibrium definition, all prices q_1 are equilibrium prices.

Second, if demand is positive, we ensure the existence of an equilibrium price q_1 by showing that demand is infinite as the price approaches zero, i.e. $\lim\limits_{q_1\to0}D\left(q_1\right)=\infty$. First, note that buyers will exhaust their entire borrowing limit as they trade, i.e. $\tilde{d}_1^j=\phi q_1\left(\tilde{a}_0^j-\tilde{l}_1^j\right)$, because any price $q_1<1$ grants them a pecuniary benefit. From the period-2 budget constraint (1.2), we obtain

$$\tilde{c}_2^j=\left(1-\phi q_1\right)\left(\tilde{a}_0^j-\tilde{l}_1^j\right). \tag{1.22}$$

Suppose the price approaches its lower limit of zero, i.e. $q_1\to0$. From the price equation (1.6), it follows that either the numerator tends to zero, i.e. $u'\left(\tilde{c}_2^j\right)\to0$, or the denominator tends to infinity, i.e. $(1-\phi)u'\left(\tilde{c}_1^j\right)+\phi u'\left(\tilde{c}_2^j\right)\to\infty$, or both.

If the numerator tends to zero, the concavity of $u\left(\tilde{c}_t^j\right)$ implies that $\tilde{c}_2^j$ becomes infinitely large, i.e. $\tilde{c}_2^j\to\infty$, and, by (1.22), so does the demand for claims, i.e. $\tilde{l}_1^j\to-\infty$.

If, in contrast, the denominator tends to infinity, this can be caused by consumption in $t=1$ and $t=2$ approaching zero, i.e. either $\tilde{c}_1^j\to 0$ or $\tilde{c}_2^j\to 0$. In the first case, all consumption is shifted to the final period, i.e. $\tilde{c}_2^j\to\infty$, from which an infinite demand for claims, i.e. $\tilde{l}_1^j\to-\infty$, follows again. In the second case, both numerator and denominator of the pricing equation (1.6) would tend to infinity, yet the numerator at a faster pace as $\phi<1$, and, consequently, the assumption $q_1\to 0$ would be violated.

Thus, at the minimum price of $q_1\to 0$, period-2 consumption $\tilde{c}_2^j$ will tend to infinity and $\tilde{l}_1^j$ will tend to minus infinity for all j with $\tilde{l}_1^j<0$. We conclude that overall demand for claims becomes infinitely large, i.e. $\lim\limits_{q_1\to 0}D\left(q_1\right)=\infty$.

All in all, for $q_1\to 0$, we obtain a bounded supply and an infinitely high demand. It is only required to ensure that this demand exists. We ensure a positive mass of $D(0)$ through assuming that at least one type of investors has had correct expectations ex-post, receiving a return that is as high as expected or higher. Formally, $E^j[R]\leq\hat{R}$ for at least one j and all realizations $\hat{R}$ of R ensures that there is at least one group that has sufficient funds available in period $t=1$ to demand claims on the asset.

There are different possibilities how supply and demand can intersect. Either $D\left(q_1\right)$ and $S\left(q_1\right)$ intersect on $(0,1]$ at (possibly multiple) price(s). Then, all prices in this set are equilibrium prices. Or they do not have an intersection on the interval. We have shown that, in this case, demand is permanently larger than supply, i.e. $D\left(q_1\right)>S\left(q_1\right)$ for any $q_1\in(0,1]$ as $D(0)>S(0)$ and there is not intersection on $(0,1]$. Hence, $q_1=1$ is the equilibrium price since, for this price, buying investors are indifferent between all levels of feasible demand, and the bounded supply $S(1)<D(1)$ can be fully met. In conclusion, we have shown that the equilibrium price exists.

Uniqueness. Second, we prove that the equilibrium price is unique and satisfies $q_1\leq 1$ in the case of positive demand. Uniqueness is ensured if, first, $\lim\limits_{q_1\to 0}D\left(q_1\right)=\infty$, second, $D\left(1\right)=S\left(1\right)=0$, and third, if $D\left(q_1\right)$ and $S\left(q_1\right)$ are monotonically decreasing functions on $(0,1]$ with $\frac{\partial D(1)}{\partial q_1}=\frac{\partial S(1)}{\partial q_1}=0$. We continue assuming their continuity and differentiability.

Regarding the first two conditions, we have shown $\lim\limits_{q_1\to 0}D\left(q_1\right)=\infty$ in the previous part, and $D\left(1\right)=S\left(1\right)=0$ follows from our assumption $\tilde{l}_1^j(1)=0$ for all j.

Next, we prove that both supply and demand are monotonic functions on $(0,1]$.

Specifically, we determine the signs of

$$\frac{\partial S(q_1)}{\partial q_1} = \sum_{j=1}^{J} \mathbb{1}_{\left\{\tilde{l}_1^j(q_1)>0\right\}} s^j \frac{\partial \tilde{l}_1^j}{\partial q_1} \tag{1.23}$$

$$\frac{\partial D(q_1)}{\partial q_1} = -\sum_{j=1}^{J} \mathbb{1}_{\left\{\tilde{l}_1^j(q_1)<0\right\}} s^j \frac{\partial \tilde{l}_1^j}{\partial q_1}. \tag{1.24}$$

Using the period-1 equilibrium conditions (1.1), (1.2), (1.3), and (1.9), and applying the implicit function theorem to (1.5), we obtain

$$\frac{\partial \tilde{l}_1^j}{\partial q_1} = \frac{1}{1+(1-2\phi)\, q_1}\left[\frac{1}{(1-\phi q_1)\, A q_1} - 2\phi \tilde{a}_0^j + (2\phi-1)\tilde{l}_1^j\right] \tag{1.25}$$

Inserting (1.25) into (1.23) and (1.24) yields

$$\frac{\partial S(q_1)}{\partial q_1} = \frac{1}{1+(1-2\phi)\, q_1}\left[\frac{J^S}{(1-\phi q_1)\, A q_1} - 2\phi \sum_{j=1}^{J} \mathbb{1}_{\left\{\tilde{l}_1^j(q_1)>0\right\}} s^j \tilde{a}_0^j + (2\phi-1) S(q_1)\right] \tag{1.26}$$

$$\frac{\partial D(q_1)}{\partial q_1} = -\frac{1}{1+(1-2\phi)\, q_1}\left[\frac{J^D}{(1-\phi q_1)\, A q_1} - 2\phi \sum_{j=1}^{J} \mathbb{1}_{\left\{\tilde{l}_1^j(q_1)<0\right\}} s^j \tilde{a}_0^j + (2\phi-1) D(q_1)\right], \tag{1.27}$$

where J^S and J^D are the number of types that are on the supply and the demand side of the market, respectively. We assume that the margin requirement is sufficiently tight, i.e. $\phi < 1/2$.

We first show that the supply curve is a weakly decreasing function of q_1. Recall that $S(q_1)$ is continuous on $(0,1]$, $\lim\limits_{q_1 \to 1} S(q_1) = 0$ and an equilibrium with positive demand $D(q_1) > 0$ requires that there is a q_1 such that $S(q_1) > 0$. Hence, there must further be a $q_1^* \equiv \min\left\{q_1 \mid \frac{\partial S(q_1)}{\partial q_1} < 0 \text{ for all } q_1 > q_1^*\right\}$.

Now we distinguish two cases. If $\frac{\partial S\left(q_1^*\right)}{\partial q_1} \neq 0$, there is no $q_1 < q_1^*$ such that $\frac{\partial S(q_1)}{\partial q_1} > 0$, and it follows $\frac{\partial S(q_1)}{\partial q_1} \leq 0$ for all $q_1 \in (0,1]$, making the supply curve monotonically decreasing.

If, however, $\frac{\partial S\left(q_1^*\right)}{\partial q_1} = 0$, this is equivalent to $S(q_1^*) = \frac{1}{2\phi-1}\left[2\phi \sum_{j=1}^{J} \mathbb{1}_{\left\{\tilde{l}_1^j\left(q_1^*\right)\geq 0\right\}} s^j \tilde{a}_0^j - \frac{J^S}{\left(1-\phi q_1^*\right) A q_1^*}\right]$. For $q_1 < q_1^*$, we prove by contradiction that supply is constant.

First suppose that $\frac{\partial S(q_1)}{\partial q_1} > 0$. From (1.26), it follows that $S(q_1) > S(q_1^*)$ in this case, which would imply $\frac{\partial S(q_1)}{\partial q_1} < 0$, violating the assumption. Now suppose that $\frac{\partial S(q_1)}{\partial q_1} < 0$. From (1.26), it follows that $S(q_1) < S(q_1^*)$ in this case, which would imply $\frac{\partial S(q_1)}{\partial q_1} > 0$, violating the assumption.

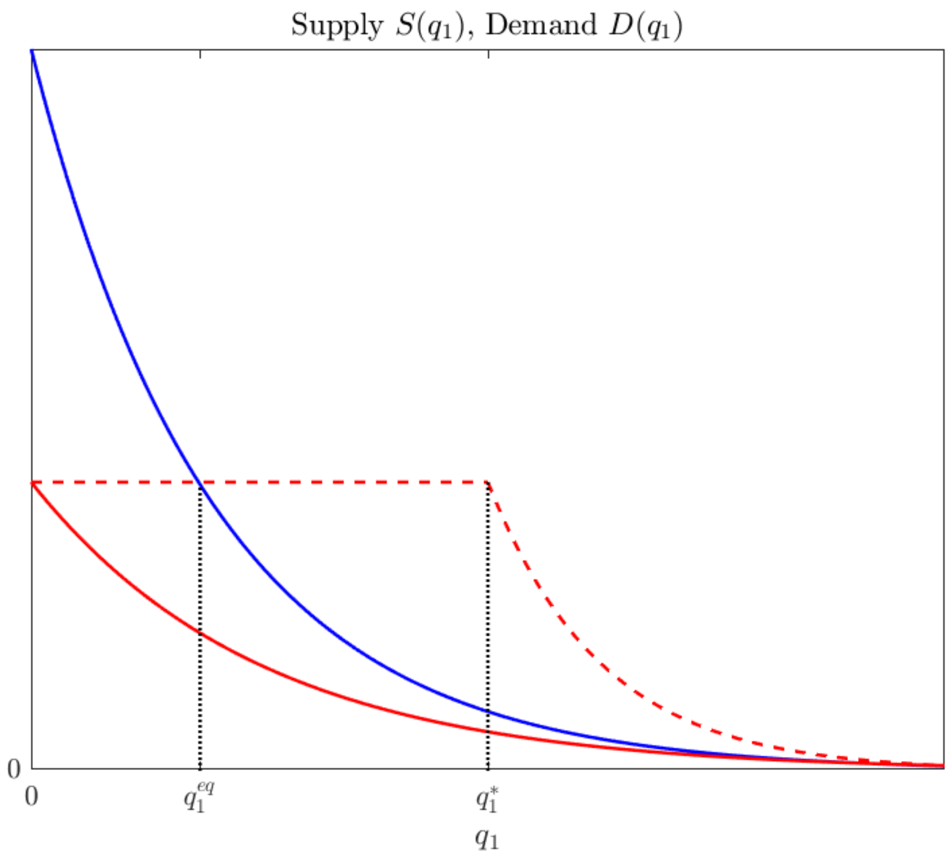


Figure 1.5: Supply and Demand in $t = 1$

This figure sketches two possible supply curves and a demand curve in period $t = 1$. Supply curves are depicted in red, while the demand curve is depicted in blue. q_1^{eq} is the equilibrium price, and q_1^* is defined as in the proof of Proposition 1.

Therefore, we obtain $\frac{\partial S(q_1)}{\partial q_1} = 0$ for all $q_1 < q_1^*$. The constancy of supply for low collateral prices reflects the fact that supply is bounded from above by the amount invested in $t = 0$. q_1^* is thus the price below which distressed investors are willing to liquidate their entire position.

The slope of the demand curve, i.e. the sign of the left-hand side of equation (1.27), is determined by the term in brackets. Under the assumption of $\phi < 1/2$, and restricting the initial endowment to $\bar{a} \leq 2$, the term in brackets is positive, yielding $\frac{\partial D(q_1)}{\partial q_1} < 0$ for any $q_1 \in (0, 1]$.

Lastly, equations (1.26) and (1.27) reveal that $\frac{\partial D(1)}{\partial q_1} = \frac{\partial S(1)}{\partial q_1} = 0$ because $J^S = J^D = \mathbb{1}_{\{\tilde{l}_1^j(1)>0\}} = \mathbb{1}_{\{\tilde{l}_1^j(1)<0\}} = S(1) = D(1) = 0$ at $q_1 = 1$.

Since all the conditions for uniqueness are satisfied, we deduce that the equilibrium price is unique (see Figure 1.5 for illustration).

Equivalences. Third, we show the two equivalences in part (ii). For part (i), suppose $q_1 = 1$. Combining equations (1.4) and (1.5) yields $\tilde{\eta}_1^j = \tilde{\eta}_1^j \phi$. The only solution for the latter condition is $\tilde{\eta}_1^j = 0$. Now, suppose $\tilde{\eta}_1^j = 0$. Equation (1.4) then becomes $u'\left(\tilde{c}_1^j\right) = u'\left(\tilde{c}_2^j\right)$. Substituting out $u'\left(\tilde{c}_2^j\right)$ in equation (1.5) yields $q_1 = 1$.

For part (ii), the equivalence is shown formally:

$$\begin{aligned} & q_1 = \frac{u'\left(\tilde{c}_2^j\right)}{(1-\phi)u'\left(\tilde{c}_1^j\right) + \phi u'\left(\tilde{c}_2^j\right)} < 1 \\ \iff & (1-\phi)u'\left(\tilde{c}_2^j\right) < (1-\phi)u'\left(\tilde{c}_1^j\right) \\ \iff & 0 < u'\left(\tilde{c}_1^j\right) - u'\left(\tilde{c}_2^j\right) = \tilde{\eta}_1^j. \end{aligned}$$

□

Proof of Proposition 2

For the proof of part (i), recall that the period-1 equilibrium price satisfies equation (1.6), where $\tilde{c}_1^j$ and $\tilde{c}_2^j$ are given by equations (1.1) and (1.2) for all j. Since the equilibrium price equals one if $\tilde{\eta}_1^j = 0$, we restrict ourselves to price effects in the case of $\tilde{\eta}_1^j > 0$. For the borrowing constraint to be binding, assume that the realization $\hat{R}$ is sufficiently adverse, satisfying $\hat{R} < 1$. Using CARA $A = -\frac{u''\left(\tilde{c}_t^j\right)}{u'\left(\tilde{c}_t^j\right)}$ for all j and t, we obtain the following equilibrium price derivatives:

$$\frac{\partial q_1}{\partial \tilde{a}_0^j} = \frac{(1-\phi)(1-R)\left(q_1\right)^2}{\frac{u'\left(\tilde{c}_2^j\right)}{u''\left(\tilde{c}_1^j\right)} + (1-\phi)\left(q_1\right)^2 \tilde{l}_1^j} \tag{1.28}$$

$$\frac{\partial q_1}{\partial \tilde{d}_0^j} = \frac{(1-\phi)\left(q_1\right)^2}{\frac{u'\left(\tilde{c}_2^j\right)}{u''\left(\tilde{c}_1^j\right)} + (1-\phi)\left(q_1\right)^2 \tilde{l}_1^j}. \tag{1.29}$$

The numerators of equations (1.28) and (1.29) are positive, and the denominator is negative. To see this, note that $\frac{\partial q_1}{\partial \tilde{c}_1^j} = -(1-\phi)\frac{u''\left(\tilde{c}_1^j\right)}{u'\left(\tilde{c}_2^j\right)}\left(q_1\right)^2 > 0$. For the denominator, it follows

$$\begin{aligned} & \frac{u'\left(\tilde{c}_2^j\right)}{u''\left(\tilde{c}_1^j\right)} + (1-\phi)\left(q_1\right)^2 \tilde{l}_1^j \leq 0 \qquad (1.30) \\ \iff & 1 \geq \frac{\partial q_1}{\partial \tilde{c}_1^j} \tilde{l}_1^j, \end{aligned}$$

which is always satisfied. If $\tilde{l}_1^j \leq 0$, the left-hand side of (1.30) is negative. But it is exceeded by one even if $\tilde{l}_1^j > 0$. The reason is that $1 \geq \frac{\partial q_1}{\partial \tilde{c}_1^j}\tilde{l}_1^j$ is the condition for finite consumption $\tilde{c}_1^j$. Consider the period-1 budget constraint $\tilde{c}_1^j = R\tilde{a}_0^j + q_1\tilde{l}_1^j + \tilde{d}_1^j - \tilde{d}_0^j$. Increasing the budget by one unit of the consumption good has two effects. First, it directly increases consumption by one unit. Second, it raises q_1, and further increases consumption by $\frac{\partial q_1}{\partial \tilde{c}_1^j}\tilde{l}_1^j$. Suppose $1 < \frac{\partial q_1}{\partial \tilde{c}_1^j}\tilde{l}_1^j$. In this case, the latter effect via q_1 dominates the direct effect, and the initial stimulus initiated an upward loop towards infinite consumption $\tilde{c}_1^j$. Hence, a finite solution requires $1 \geq \frac{\partial q_1}{\partial \tilde{c}_1^j}\tilde{l}_1^j$, concluding the proof of part (i).

Turning to part (ii), for the equilibrium price derivative with respect to borrowing under a homogeneous belief, we obtain

$$\frac{\partial q_1}{\partial \tilde{d}_0} = \frac{(1-\phi)\,(q_1)^2}{\frac{u'(\tilde{c}_2)}{u''(\tilde{c}_1)}}, \tag{1.31}$$

which is negative for a concave utility function. □

Proof of Proposition 3

For the proof of part (i), let investors hold heterogeneous beliefs $\mathcal{F}$. The individual type-j decisions for investment and borrowing are governed by equations (1.8) and (1.9), that we rewrite as functions of its belief $F^j(R)$ in the following way:

$$\begin{aligned} q_0 u'\left(\tilde{c}_0^j\left(\tilde{a}_0^j\left(F^j(R)\right),\tilde{d}_0^j\left(F^j(R)\right)\right)\right) &= \int_{\underline{R}}^{\overline{R}} R u'\left(\tilde{c}_1^j\left(\tilde{a}_0^j\left(F^j(R)\right),\tilde{d}_0^j\left(F^j(R)\right)\right)\right) \ldots \\ \ldots + u'\left(\tilde{c}_2^j\left(\tilde{a}_0^j\left(F^j(R)\right)\right)\right) &+ \tilde{\eta}_1^j\left(\tilde{a}_0^j\left(F^j(R)\right),\tilde{d}_0^j\left(F^j(R)\right)\right)\phi q_1 dF^j(R) \quad (1.32) \\ u'\left(\tilde{c}_0^j\left(\tilde{a}_0^j\left(F^j(R)\right),\tilde{d}_0^j\left(F^j(R)\right)\right)\right) &= \int_{\underline{R}}^{\overline{R}} u'\left(\tilde{c}_1^j\left(\tilde{a}_0^j\left(F^j(R)\right),\tilde{d}_0^j\left(F^j(R)\right)\right)\right) dF^j(R). \end{aligned}$$

(1.33)

Notably, period-0 choices $\tilde{a}_0^j\left(F^j(R)\right)$ and $\tilde{d}_0^j\left(F^j(R)\right)$ are direct functions of type j's belief, while period-1 and period-2 variables are both indirect functions of $F^j(R)$ via $\tilde{a}_0^j\left(F^j(R)\right)$ and $\tilde{d}_0^j\left(F^j(R)\right)$ direct functions of it through the expectation operator.

In the following, we apply the calculus of variation, as explained in the main text. Consider a perturbation to beliefs of the form $F^j(R) + \epsilon G^j(R)$, where $\epsilon > 0$ is an arbitrary number, and $G^j(R)$ captures the direction of the perturbation. $F^j(R)+\epsilon G^j(R)$ is required to be a valid cdf for small enough ϵ, so we assume it is continuous and differentiable, it satisfies $G(\underline{R}) = G(\overline{R}) = 0$, and $\partial\left(F^j(R) + \epsilon G^j(R)\right)/\partial R \geq 0$ for sufficiently small ϵ. Lastly, let δ denote the operator for functional derivatives.

We characterize the variational derivatives of investment and borrowing choices when beliefs $F^j(R)$ are perturbed with direction $G^j(R)$, i.e. $\frac{\delta \tilde{a}_0^j}{\delta F^j} \cdot G^j$ and $\frac{\delta \tilde{d}_0^j}{\delta F^j} \cdot G^j$. Optimism and pessimism are measured relative to each other in the sense of first order stochastic dominance. A perturbation $G^j(R)$ makes type-j investors more optimistic if and only if it satisfies $F^j(R) + \epsilon G^j(R) \leq F^j(R)$ for all R. It is easy to see that more optimism requires the perturbation to have a negative direction, i.e. $G^j(R) \leq 0$ for all R. Analogously, investors of type j are made more pessimistic through a perturbation with direction $G^j(R) \geq 0$ for all R.

Applying the implicit function theorem to (1.32) and (1.33), and combining the resulting expressions yield

$$\frac{\delta \tilde{a}_0^j}{\delta F^j} \cdot G^j = \frac{\int_{\underline{R}}^{\overline{R}} u''\left(\tilde{c}_1^j\right) \tilde{a}_0^j G^j(R) dR \cdot \left(\int_{\underline{R}}^{\overline{R}} (1+\phi) u''\left(\tilde{c}_1^j\right) dF^j(R) + q_0 u''\left(\tilde{c}_0^j\right)\right)}{\left(\int_{\underline{R}}^{\overline{R}} R u''\left(\tilde{c}_1^j\right) dF^j(R) + q_0 u''\left(\tilde{c}_0^j\right)\right) \cdot \left(\int_{\underline{R}}^{\overline{R}} (1+\phi) u''\left(\tilde{c}_1^j\right) dF^j(R) + q_0 u''\left(\tilde{c}_0^j\right)\right)} \cdots$$
$$\cdots \frac{-\int_{\underline{R}}^{\overline{R}} \left(u'\left(\tilde{c}_1^j\right) + (R+\phi q_1) u''\left(\tilde{c}_1^j\right) \tilde{a}_0^j\right) G^j(R) dR \cdot \left(\int_{\underline{R}}^{\overline{R}} u''\left(\tilde{c}_1^j\right) dF^j(R) + u''\left(\tilde{c}_0^j\right)\right)}{-\left(\int_{\underline{R}}^{\overline{R}} (R+\phi q_1) R u''\left(\tilde{c}_1^j\right) + (1-\phi q_1) u''\left(\tilde{c}_2^j\right) dF^j(R)\right) \cdot \left(\int_{\underline{R}}^{\overline{R}} u''\left(\tilde{c}_1^j\right) dF^j(R) + u''\left(\tilde{c}_0^j\right)\right)} \tag{1.34}$$

$$\frac{\delta \tilde{d}_0^j}{\delta F^j} \cdot G^j = \frac{-\int_{\underline{R}}^{\overline{R}} u''\left(\tilde{c}_1^j\right) \tilde{a}_0^j G^j(R) dR}{u''\left(\tilde{c}_0^j\right) + \int_{\underline{R}}^{\overline{R}} u''\left(\tilde{c}_1^j\right) dF^j(R)} + \frac{\int_{\underline{R}}^{\overline{R}} R u''\left(\tilde{c}_1^j\right) dF^j(R) + q_0 u''\left(\tilde{c}_0^j\right)}{u''\left(\tilde{c}_0^j\right) + \int_{\underline{R}}^{\overline{R}} u''\left(\tilde{c}_1^j\right) dF^j(R)} \cdot \frac{\delta \tilde{a}_0^j}{\delta F^j} \cdot G^j. \tag{1.35}$$

First, we further investigate equation (1.34). Assuming that the choice of parameters ensures a non-zero trading volume, i.e. $A < 1$ and beliefs $\mathcal{F}$ sufficiently divergent such that $\bar{a} - \tilde{a}_0^j \neq 0$ for some j, and that the borrowing constraints bind in response to the adverse shock, i.e. $\hat{R} < 1$ and $\phi < \frac{1}{2}$ such that $\tilde{\eta}_1^j > 0$ for all j, the numerator is negative for $G^j(R) \leq 0$, and positive for $G^j(R) \geq 0$. The denominator is always negative. Hence, the functional derivative $\frac{\delta \tilde{a}_0^j}{\delta F^j} \cdot G^j$ is positive for $G^j(R) \leq 0$ and negative for $G^j(R) \geq 0$.

Given the signs of the components in (1.35), it follows that $\frac{\delta \tilde{a}_0^j}{\delta F^j} \cdot G^j$ and $\frac{\delta \tilde{d}_0^j}{\delta F^j} \cdot G^j$ have the same sign for each $G^j(R)$. Consequently, the two variational derivatives in (1.34) and (1.35) turn out to be positive if investors are more optimistic ($G^j(R) \leq 0$), and negative if they are more pessimistic ($G^j(R) \geq 0$).

Proving part (*ii*), we employ the identical procedure as above. Let investors hold the homogeneous belief $F(R)$. Let further $G(R)$ be the direction of a perturbation of the homogeneous belief. We obtain as the functional derivative of borrowing

$$\frac{\delta \tilde{d}_0}{\delta F} \cdot G = \frac{-\int_{\underline{R}}^{\overline{R}} u''\left(\tilde{c}_1\right) \bar{a} G(R) dR}{u''\left(\tilde{c}_0\right) + \int_{\underline{R}}^{\overline{R}} u''\left(\tilde{c}_1\right) dF(R)}, \tag{1.36}$$

which is as well positive for more optimistic investors ($G^j(R) \leq 0$), and negative for more pessimistic investors ($G^j(R) \geq 0$). □

Proof of Theorem 1

With regard to part (i), let investors hold heterogeneous beliefs $\mathcal{F}$. Let further $G^j(R)$ be the direction of a perturbation of type-j investors' belief $F^j(R)$, and beliefs $F^i(R)$ be constant for all $i \neq j$.

Recall that the functional derivative $\frac{\delta}{\delta F^j} \cdot G^j$ describes a gradient, so it is identical to a partial derivative if the functional argument is one-dimensional. We write the period-1 equilibrium price as a function of beliefs, i.e. $q_1 = q_1\left(\tilde{a}_0(\mathcal{F}), \tilde{d}_0(\mathcal{F})\right)$. It follows

$$\frac{\delta q_1}{\delta F^j} \cdot G^j = \frac{\delta q_1}{\delta \tilde{a}_0^j} \cdot \frac{\delta \tilde{a}_0^j}{\delta F^j} \cdot G^j + \frac{\delta q_1}{\delta \tilde{d}_0^j} \cdot \frac{\delta \tilde{d}_0^j}{\delta F^j} \cdot G^j = \frac{\partial q_1}{\partial \tilde{a}_0^j} \cdot \frac{\delta \tilde{a}_0^j}{\delta F^j} \cdot G^j + \frac{\partial q_1}{\partial \tilde{d}_0^j} \cdot \frac{\delta \tilde{d}_0^j}{\delta F^j} \cdot G^j. \quad (1.37)$$

Using Propositions 2 and 3, we obtain statement (1) of part (i).

For statement (2), let $G^j(R) < 0 < G^i(R)$ with $|G^j(R)| = |G^i(R)|$ for all R be the directions of two perturbations that make investors of type j more optimistic, and investors of type i more pessimistic by the same magnitude. We investigate each factor in the two summands on the right-hand side of equation (1.37) separately. First, note that equations (1.34) and (1.35) imply that

$$\left| \frac{\delta \tilde{a}_0^j}{\delta F^j} \cdot G^j \right| = \left| \frac{\delta \tilde{a}_0^i}{\delta F^i} \cdot G^i \right| \text{ and } \left| \frac{\delta \tilde{d}_0^j}{\delta F^j} \cdot G^j \right| = \left| \frac{\delta \tilde{d}_0^i}{\delta F^i} \cdot G^i \right|.$$

Second, taking the derivatives of equations (1.28) and (1.29) shows that q_1 is a (decreasing and) concave function of investment and borrowing, i.e. $\frac{\partial^2 q_1}{\partial^2 \tilde{a}_0^j} \leq 0$ and $\frac{\partial^2 q_1}{\partial^2 \tilde{d}_0^j} \leq 0$. As for any concave function, it follows that

$$\left| \frac{\delta q_1}{\delta \tilde{a}_0^j} \right| > \left| \frac{\delta q_1}{\delta \tilde{a}_0^i} \right| \text{ and } \left| \frac{\delta q_1}{\delta \tilde{d}_0^j} \right| > \left| \frac{\delta q_1}{\delta \tilde{d}_0^i} \right|.$$

Inserting the two former results in equation (1.37) yields statement (2).

To prove part (ii), let investors hold the homogeneous belief $F(R)$. Let further $G(R)$ be the direction of a perturbation of the homogeneous belief. Equation (1.37) simplifies to

$$\frac{\delta q_1}{\delta F} \cdot G = \frac{\partial q_1}{\partial \tilde{d}_0} \cdot \frac{\delta \tilde{d}_0}{\delta F} \cdot G, \quad (1.38)$$

which is negative for $G(R) \leq 0$ and positive for $G(R) \geq 0$ by the same arguments as in statement (1) of part (i). □

Proof of Theorem 2

We start out by proving that $\hat{R}^*_{het} > \hat{R}^*_{hom}$, where $\hat{R}^*_{het}$ and $\hat{R}^*_{hom}$ are defined in Definition

3.

Consider a population with investors holding heterogeneous beliefs $\mathcal{F}$. Let $\hat{R}^{*j}_{het}$ denote the lowest possible realization $\hat{R}$ such that the collateral constraint of type-j investors is slack, i.e. $\tilde{\eta}^j_1 = 0$ and $q_1 = 1$, which are equivalent to $\tilde{c}^j_1 = \tilde{c}^j_2$. At this point, the borrowing constraint yields $\tilde{d}^j_1 = \phi\tilde{a}^j_0$. Using this, and equating the budget constraints (1.1) and (1.2), one obtains $\hat{R}^{*j}_{het} = 1 - 2\phi + \frac{\tilde{d}^j_0}{\tilde{a}^j_0}$.

Given the result from Proposition 1, it suffices that one type of investors is constrained to make all investors constrained. We refer to this situation as financial distress, and it follows that $\hat{R}^*_{het} = \max\left\{\hat{R}^{*j}_{het}\right\}_{j\in\{1,\dots,J\}}$. Assuming without loss of generality that investors are ordered from more to less optimistic types, i.e. $F^1(R) < \dots < F^J(R)$ for all R, we obtain $\hat{R}^*_{het} = \hat{R}^{*1}_{het}$. For the homogeneous case, we derive $\hat{R}^*_{hom} = 1 - 2\phi + \frac{\tilde{d}_0}{\bar{a}}$ equivalently.

To show that $\hat{R}^*_{het} > \hat{R}^*_{hom}$, it is sufficient to prove that $\frac{\tilde{d}^1_0}{\tilde{a}^1_0} > \frac{\tilde{d}_0}{\bar{a}}$. Since type $j = 1$ is the most optimistic type, we know that $\tilde{a}^1_0 > \bar{a}$ and $\tilde{d}^1_0 > \tilde{d}_0$. To prove that $\frac{\tilde{d}^1_0}{\tilde{a}^1_0} > \frac{\tilde{d}_0}{\bar{a}}$, we show that $\tilde{d}^1_0 - \tilde{d}_0 > \tilde{a}^1_0 - \bar{a}$.

The latter statement would follow if a perturbation, making a specific belief more optimistic, i.e. $G^1(R) < 0$ for all R, always increased borrowing by more than investment, i.e. $\frac{\delta\tilde{d}^1_0}{\delta F^1} \cdot G^1 > \frac{\delta\tilde{a}^1_0}{\delta F^1} \cdot G^1$. We deduce from equation (1.35) that this condition is satisfied provided that

$$\frac{\int_{\underline{R}}^{\overline{R}} Ru''\left(\tilde{c}^1_1\right) dF^1 + q_0 u''\left(\tilde{c}^1_0\right)}{u''\left(\tilde{c}^1_0\right) + \int_{\underline{R}}^{\overline{R}} u''\left(\tilde{c}^j_1\right) dF^1} > 1. \tag{1.39}$$

Under the presumption made in Theorem 2, requiring the homogeneous belief $F(R)$ to be less optimistic than at least one type's belief in the heterogeneous case, implying $F^1(R) < F(R)$ for all R, inequality (1.39) is satisfied for any type-1 belief F^1 sufficiently optimistic. Hence, under this assumption, we obtain $\hat{R}^*_{het} > \hat{R}^*_{hom}$.

Ultimately, we derive the corresponding probabilities of financial distress. In our setting, it is for the heterogeneous and the homogeneous case, respectively

$$\Pr\left(\tilde{\eta}^1_1 > 0\right) = \Pr\left(R \le \hat{R}^*_{het}\right) = F\left(\hat{R}^*_{het}\right)$$
$$\Pr\left(\eta_1 > 0\right) = \Pr\left(R \le \hat{R}^*_{hom}\right) = F\left(\hat{R}^*_{hom}\right).$$

Given $\hat{R}^*_{het} > \hat{R}^*_{hom}$ and the strict monotonicity of the cdf F, it follows that $F\left(\hat{R}^*_{het}\right) > F\left(\hat{R}^*_{hom}\right)$. $\square$

Proof of Proposition 4

Proposition 4 follows from Definition 4 and Proposition 2. $\square$

Proof of Proposition 5

With regard to part (i), let investors hold heterogeneous beliefs $\mathcal{F}$. Let further $G^j(R)$ be the direction of a perturbation of type-j investors' belief $F^j(R)$, and beliefs $F^i(R)$ be constant for all $i \neq j$. We calculate the functional derivatives of distributive and collateral externalities with respect to beliefs in the following way:

$$\frac{\delta D^i_{\tilde{a}^j_0}}{\delta F^j} \cdot G^j = \frac{\delta \left(\frac{q_1}{\partial \tilde{a}^j_0}\right)}{\delta F^j} \cdot G^j \cdot \tilde{l}^j_1 = \left(\frac{\partial^2 q_1}{\partial \tilde{a}^j_0 \partial \tilde{a}^j_0} \frac{\delta \tilde{a}^j_0}{\delta F^j} \cdot G^j + \frac{\partial^2 q_1}{\partial \tilde{a}^j_0 \partial \tilde{d}^j_0} \frac{\delta \tilde{d}^j_0}{\delta F^j} \cdot G^j\right) \tilde{l}^j_1,$$

and analogously for $D^i_{\tilde{d}^j_0}$, $C^i_{\tilde{a}^j_0}$, and $C^i_{\tilde{d}^j_0}$. Since q_1 is strictly decreasing and concave in both $\tilde{a}^j_0$ and $\tilde{d}^j_0$, and using our results from above on the sign of the functional derivatives $\frac{\delta \tilde{a}^j_0}{\delta F^j} \cdot G^j$ and $\frac{\delta \tilde{d}^j_0}{\delta F^j} \cdot G^j$, it follows that the term in brackets is unambiguously negative for $G^j(R) < 0$, and positive for $G^j(R) > 0$. This proves the first statement of part (i).

Statement (2) of part (i), as well as part (ii), follow from the same arguments as those used in the proof of Theorem 1. $\square$

Proof of Proposition 6

First, we derive the tax formulas in part (i). Consider the period-0 optimization problem of a type-j agent with taxes:

$$\max_{c^j_0, a^j_0 \geq 0, d^j_0} \quad u\left(c^j_0\right) + E^j\left[V^j_1\left(a^j_0, d^j_0 | \tilde{a}_0, \tilde{d}_0\right)\right] \quad \text{s.t.}$$

$$\left(\lambda^j_0\right) \quad c^j_0 = e + \left(1 - \tau^j_d \operatorname{sgn}\left(\tilde{d}^j_0\right)\right) d^j_0 + \left(1 - \tau^j_a \operatorname{sgn}\left(\bar{a} - \tilde{a}^j_0\right)\right) q_0 \left(\bar{a} - a^j_0\right) + T^j. \tag{1.40}$$

This problem gives rise to the following optimality conditions:

$$\left(1 - \tau^j_a \operatorname{sgn}\left(\bar{a} - \tilde{a}^j_0\right)\right) q_0 u'\left(c^j_0\right) = E^j\left[R u'\left(c^j_1\right) + u'\left(c^j_2\right) + \eta^j_1 \phi q_1\right] \tag{1.41}$$

$$\left(1 - \tau^j_d \operatorname{sgn}\left(\tilde{d}^j_0\right)\right) u'\left(c^j_0\right) = E^j\left[u'\left(c^j_1\right)\right]. \tag{1.42}$$

In a symmetric equilibrium, it will always be the case that $c^j_0 = \tilde{c}^j_0$, $a^j_0 = \tilde{a}^j_0$ and $d^j_0 = \tilde{d}^j_0$ for each j. Combining the latter two conditions with their counterparts from the social planner problem, i.e. equations (1.13) and (1.14), respectively, using the planner's pricing relation $\tilde{\psi} = q_0 \omega^j u'\left(\tilde{c}^j_0\right)$, and solving for the taxes yields the tax formulas (1.15) and (1.16).

Second, it follows that, using these taxes, the competitive allocation is constrained-efficient. Specifically, substituting (1.15) and (1.16) into the optimality conditions of the competitive allocation with taxes, i.e. (1.41), and (1.42), replicates the planner's optimality conditions (1.13) and (1.14), as well as $\tilde{\lambda}_0 = \omega^j u'\left(\tilde{c}_0^j\right)$ for each j. Moreover, rebating revenues through T^j for all j ensures that individual period-0 budget constraints are satisfied, and the same holds for the resource constraint in consequence. To summarize, the competitive allocation with taxes satisfies the identical set of conditions, so it turns out to be constrained-efficient.

By the same arguments, we derive the homogeneous tax formula 1.17 in part (ii). □

Proof of Theorem 3

Theorem 3 follows from Propositions 3 and 6. □

Chapter 2

The Evolution of Income Taxation in Prussia

By Marco Wysietzki

Abstract

In the 19th century, Prussia used a system of class-specific differentiated lump sum taxes. According to the second welfare theorem such taxes can be used to reach distributive objectives without having to sacrifice first best efficiency. Still, in the course of the 19th century Prussia replaced these differentiated lump sum taxes by a combination of differentiated lump sum and income taxes. This system in turn was replaced by a comprehensive income tax. Thus, there no longer was a possibility to tailor income taxes to class-specific features such as behavioral responses or within class income redistribution. In this paper, we study these reforms in detail through the lens of a Mirrleesian income tax model. As a main result, we show that the reforms increased potential government revenue drastically and in a politically feasible way, even though various appealing features of the prevailing tax system were given up, and even though the political economy of Prussia, with its three-class franchise, had its "median voter" in the top ten percent.

2.1 Introduction

This paper examines the path to modern comprehensive income taxation for the example of Prussia. It applies modern public finance theory to evaluate various tax reforms, proving them to be reasonable concerning their tax revenue potential and their majority support. This is particularly interesting as the reforms abolished some taxes that had attractive features from today's perspective. For example, taxing people with a lump sum tax does not change their incentives to work and therefore leads to fewer distortions compared to a comprehensive income tax. Lump sum taxes are especially useful if they are based on a characteristic that people cannot easily change and that is correlated with ability. This so-called tagging thus allows redistribution without distortions.

At the beginning of the 19th century, the existing poll tax in Prussia was replaced by a class tax. This means that people had to pay lump sum taxes depending on their social class. Since social classes can be assumed to be correlated with ability, the class tax was an example of tagging. In the middle of the 19th century, this tax was restricted in several steps gradually before it was replaced by a comprehensive income tax. Thus, Prussia abolished lump sum taxes and tagging in particular.

Most of these reforms were initiated because of the need for higher government revenue, which is referred to as an increase of fiscal state capacity in the literature. At first glance, however, this goal makes the reforms puzzling. During the reforms Prussia gave up some of the tools that make it possible to raise revenue without creating much of a distortion. Why should the government do so if it was seeking to increase revenue? At the same time, the political feasibility of these reforms seems questionable. During the reforms taxes for the rich were increased although the Prussian voting system favored them. Thus, the expansion of fiscal state capacity required the willingness of the rich to tax themselves more heavily. So how could these reforms be realized?

To solve these puzzles, we introduce a model based on Diamond (1998) that allows to evaluate the reforms concerning state capacity and political feasibility. The abolition of the class tax meant that it was no longer possible to differentiate between classes when setting taxes. However, it became possible to tax higher incomes within the classes more heavily. Depending on the distribution of skills and classes within the population, this could lead to an increase of tax revenue. In our model framework, we assume that tax revenue is paid back to individuals in per capita transfers. In terms of political economy, this can lead to some rich people being taxed more heavily but still benefiting due to the reforms. The higher tax payments are overcompensated by even higher per capita transfers, which are financed by raising taxes on the even richer.

So our main result here is that all reforms can lead to an increase in state capacity and that they can make a majority better off. These results are in line with historical data. Based on this data, we show that most reforms are not only supported by a majority of the population, but also by a majority of voting power taking the special Prussian voting system into account.

Furthermore, we discuss limitations of lump sum taxation and tagging in more detail. There are hints that the correlation between social class and ability became weaker over time. We use a perturbation approach to describe how the additional benefit of tagging depends on the strength of this correlation. We find that the convergence of ability distributions in the different classes leads to the benefits of tagging converging to zero. The Prussian reforms balanced the diminishing, albeit still positive, benefits of tagging with the advantages of an income tax. Although a positive impact on tax revenue was lost with the abolition of tagging, tax revenue nevertheless increased.

This paper connects three strands of literature. First, it joins several works on Prussian tax reforms. Many of them are descriptive, such as Hill (1892) or Schremmer (2013). Tilly (2010) uses historical data to determine the evolution of income inequality in Prussia. Spoerer (2015) focuses on changes of tax incidence. However, to the best of our knowledge, none of these papers has examined the political feasibility of the reforms.

Second, it relates to the literature on tagging. As mentioned above, Akerlof (1978) was one of the first to model and to evaluate the idea of tagging. Boadway and Pestieau (2006) extend his model and investigate optimal redistribution of tax systems depending on their level of tagging. Some papers propose specific characteristics as possible instruments for tagging and show how their use could theoretically improve welfare or increase tax revenue. Cremer et al. (2010) for example focus on gender, and Mankiw and Weinzierl (2010) on height. Weinzierl (2012) gives a reason why tagging is rarely used in practice. We contribute to this literature by exploring the class tax as an example for tagging and by formally showing limitations of tagging. We show that the theoretical advantages of tagging did not play out in reality and provide an explanation for its abolition: The reforms combined elements that increased tax revenue with the abolition of tagging.

Third, this paper adds to a growing literature on state capacity. With respect to its fiscal dimension, most notable is the work of Besley and Persson (2009). They assume that the state's capabilities of taxation and regulation depend on past investments in fiscal capacity. They discuss determinants of these investments. Besley and Persson (2010) add the risk of internal and external conflicts to this setting. Acemoglu et al. (2011) address the efficiency of collecting taxes. They show the conditions for an inefficient state and how this is related to income inequality. In this paper, we refer to the concept of state capacity by using it as a motivation for reforms.

The remainder of this paper is organized as follows. In section 2.2 we provide an overview of Prussian tax reforms related to incomes in the nineteenth century. Section 2.3 introduces our model. In section 2.4 we connect our model to historical data and examine the reforms' political feasibility before concluding in section 2.5.

2.2 Historical Background

In this section, we give an overview of the evolution of Prussia's taxation of incomes. We concentrate on four reforms: The introduction of a class tax in 1820, the combination of income and class taxes in 1851, some adjustments in 1873 and the turn to a progressive income tax in 1891. See Table 2.1 for an overview.

The Introduction of a Class Tax in 1820

In the beginning of the nineteenth century, taxation in Prussia differed a lot between regions. This was especially true as new regions were added to Prussia after the Congress of Vienna. The aim to standardize the system led to the introduction of a poll tax in most parts of the state in 1811 (Hill, 1892). This system however only lasted for a few years. Earlier attempts to follow the example of England and to introduce an income tax were not successful due to resistance among the population (Teschemacher, 1912). Thus, in 1820 it was not a statewide income tax to be implemented, but a class tax for most parts of the state. It led to people being taxed depending on the social class they belonged to. In bigger cities, however, class taxation was not used, but a grist and slaughter tax was enforced (Schwarz and Strutz, 1902).

1811	Poll Tax.
1820	Introduction of Class Tax.
1848	Revolution in Prussia.
1849	Introduction of Three-class Franchise.
1851	Class and Income Tax.
1873	Smaller Reforms towards Income Tax.
1891	Income Tax.

Table 2.1: Timeline of Reforms in Prussia

After some adjustments in 1821, there existed four different classes: Especially affluent and rich inhabitants in the first class, affluent residents in the second, lower bourgeoisie and peasants in the third and wage laborers, servants and day laborers in the fourth class. Each class consisted of three subclasses to take wealth differences within classes into account, yielding twelve different tax amounts. The lowest amount to be paid per year was $\frac{1}{2}$ Thaler, while the highest was 144 Thaler (Schremmer and Stern, 1989).

The tax function can therefore be summarized as

$$\mathcal{T}^{1821} : T(i,j) = T_{ij}, \quad 1 \leq i \leq 4, 1 \leq j \leq 3$$

where i denotes classes and j subclasses. Here it is $144 = T_{11} > T_{12} > \ldots > T_{43} = \frac{1}{2}$.

Note that the distribution over classes was heavily skewed. The data in Engel (1868) show that in 1851 the second class was twice, the third one ten times and the fourth one even 100 times the size of the first class. Hill (1892) argues that the class tax was perceived as a good compromise between a uniform per capita tax and an income tax. The former was considered as unjust and inappropriate as it could not collect sufficient revenue – the state's financial situation was quite bad and an increase of tax revenue was needed (and achieved). The latter required too much information about tax-payers' financial situations. In fact, it was forbidden by law to use further information about income. As the social class memberships were known by the government, identifying the correct tax payment for somebody was easy and did not need additional information (Hoffmann et al., 1840).

Class taxation is connected to the concept of tagging. Tagging means taxing people depending on observable characteristics that cannot be changed by behavioral responses. If these characteristics are correlated with ability, redistribution of income without distorting work incentives is possible (Weinzierl, 2012). According to the second welfare theorem, differentiated lump sum taxes can be used to reach distributive objectives without having to sacrifice first best efficiency. Assuming social classes to be correlated with abilities to generate income makes the class tax an example of tagging.

The Combination of Class and Income Taxes in 1851

In 1845 there were plans to reform the class tax and to abolish the grist and slaughter tax as its burden for the poor was perceived as too high. However, municipalities did not want to waive the revenue out of this tax and did not believe that the class tax could compensate for it. Furthermore, the members of the chambers – newly installed after the revolution in 1848 – did not want to put a higher tax burden on themselves.

As a compromise the grist and slaughter tax was not abolished and the class tax was kept constant for most of the people, but replaced by a classified income tax for people with incomes above 1000 Thaler (Spoerer, 2015).

In more detail, the first class was abolished, while the number of subclasses for the other three classes was extended to thirteen with the minimum tax per year of $\frac{1}{2}$ and the maximum of 24 Thaler. Everybody with a yearly income above 1000 Thaler – especially the former members of the first class – was assigned to one of 30 income intervals. Each of these intervals was associated to a certain tax payment. These payments should not exceed 3% of income and therefore started with 30 Thaler per year for the lowest interval and ended with 7200 for the highest interval. The tax function can be formalized as

$$\mathcal{T}^{1851}: \begin{cases} T(y) = T_{y_n} \quad \forall y \in [y_n, y_{n+1}), n \in \mathbb{N}^{\leq 30}, & \text{if } i = 1 \\ T(i,j) = T_{ij} \quad \forall j \in \{1,2,3,4,5\}, & \text{if } 2 \leq i \leq 4 \end{cases}$$

where i denotes classes, while j and n denote subclasses. For $i = 1$ tax payments are determined by income intervals y_n to y_{n+1} with $n \in \mathbb{N}^{\leq 30}$. It is $1000 = y_1 < y_2 < \ldots < y_{30} = 240,000$ with increasing interval sizes from 600 up to $60,000$. The tax payment is $T_{y_n} \approx 3\% \cdot y_n$ for all n, yielding decreasing average and zero marginal tax rates within the intervals. For the higher classes it is $24 = T_{21} > T_{22} > \ldots > T_{44} = \frac{1}{2}$.

This reform led to a substantial increase in tax revenue, showing that the class tax had been very low for at least some individuals (Engel, 1868).

Some Adjustments in 1873

In 1873, the Prussian state's fiscal situation was good due to French reparation payments after the war and an economic boom. Thus, it was decided to relieve the poor by abolishing the grist and slaughter tax. At the same time a poverty line was introduced leaving incomes below 420 Mark (140 Thaler[1]) untaxed. In addition, the number of income intervals in the classified part of the income tax system was increased and the upper limit on tax payments, previously 7200 Thaler, was abolished. All together this led to a higher tax burden for the very rich and more redistribution.

There was also a change concerning the identification of class membership for people in the class tax part of the system (with incomes less than 3000 Mark (1000 Thaler)). With the reform in 1873 their estimated incomes were used to decide which class they belonged to. Thus, the former class tax became something very similar to an income

[1]In 1871, a currency reform took place in which Thaler were exchanged for marks at a ratio of one to three.

tax. This procedure did have a starting point earlier as already in 1867 instructions existed which linked certain incomes to tax rates. At this point, however, these were only unofficial reference points. In 1873 the connection between incomes and tax rates became official (Hill, 1892). Nonetheless, the class tax did not only depend on estimated incomes, but also took other factors into account, including a high number of children, illnesses or indebtedness (Engel, 1875). Thus, some historians call it an income tax, while others say it is only close to one. In any case, this reform abolished the sole taxation of people according to their social class. The tax function can be summarized as

$$\mathcal{T}^{1873}: \begin{cases} T(y) = T_{y_n} \quad \forall y \in [y_n, y_{n+1}), n \in \mathbb{N}, & \text{if } i = 1 \\ T(y,z) = T_{y_m}(z) \quad \forall y \in [y_m, y_{m+1}), m \in \mathbb{N}^{\leq 12}, & \text{if } 2 \leq i \leq 4 \end{cases}$$

where i denotes classes, while n and m denote subclasses. For $i = 1$ tax payments are determined by income intervals y_n to y_{n+1} with $n \in \mathbb{N}$. It is $1000 = y_1 < y_2 < \ldots$[2] and $T_{y_n} \approx 3\% \cdot y_n$ for all n. For the higher classes, tax payments are determined similarly, but other factors z are taken into account.

The Progressive Income Tax in 1891

Except for an increase of the poverty line to 900 Mark in 1883, there was no real reform up to the year 1891. Then, the so called Miquelian[3] tax reform took place. It essentially abolished the class tax and the classified income tax and replaced them by a general income tax. The exact tax payment still depended on income intervals and not on the exact income. However, these intervals were much finer than before: The incomes between 3000 and $125,000$ Mark were divided into 75 intervals. Above, each interval consisted of 5000 additional Mark. The income tax was progressive starting with 6 Mark per year for people with a yearly income between 900 and 1050 Mark. An income of 3000 was already taxed by 60 Mark and one of 5000 by 132. These are tax rates of below 1%, of 2% and of 2.64%, respectively (Hill, 1892). The top tax rate of 4% was used for incomes higher than $100,000$ Mark (Königliches Statistisches Bureau, 1898). Less than 3% of tax payers were now responsible for more than 70% of total income tax amount (Buggeln, 2022). The tax function, can be summarized as follows

$$\mathcal{T}^{1891}: T(y) = T_{y_n} \quad \forall y \in [y_n, y_{n+1}), n \in \mathbb{N}$$

[2]Due to Königliches Statistisches Bureau (1898) the highest documented income was 1.6 million Mark, which corresponds to $n = 112$.

[3]Named after Johannes Franz Miquel, Prussian minister of finance.

where tax payments are determined by income intervals y_n to y_{n+1} with $n \in \mathbb{N}$. It is $y_1 = 0$ and $T_{y_1} = 0$ and for example $y_{12} = 3000$ with $T_{y_{12}} = 60$.

People with incomes above 3000 Mark had to hand in a tax return. False statements in these returns were punished by heavy fines. This reform also changed other taxes and even implemented a wealth tax. In total, tax revenue increased.

Why these Reforms?

Some countries, including other German states, implemented income taxes much earlier. The distinction of class tax and classified income tax did not seem appropriate anymore, which could explain the last reform. Concerning the timing, Bismarck, the Prussian prime minister, seems to be essential. He strongly preferred indirect over direct taxation, and therefore blocked reforms earlier. After his retirement in 1890 these reforms became possible.

However, it is still not obvious, why Prussian politicians supported this reform that led to higher taxes for themselves. In general, the literature discusses three motives: state capacity, demand for redistribution and a competition for political power.

In this paper state capacity is used as a synonym for potential tax revenue. The aim to increase it can explain all reforms. Especially in 1820 and 1851 the government needed money and therefore realized the reforms. Later on, the financial situation was much better and not necessarily the reason for the reforms in 1873 and 1891. However, increasing revenue potential can still be reasonable taking a possible future increase of expenditures into account (Besley and Persson, 2011). In section 2.3 we show in detail that the aim to increase state capacity can indeed be an explanation for all reforms.

The demand for more redistribution in times of growing inequality can also be a reason for reforms (Acemoglu and Robinson, 2000). The revolution in 1848 led to a change of the political system including the introduction of a voting system. A similar pressure from left movements existed later on. For example it led to the introduction of social security systems in the years after 1883. Turning to a progressive income tax and using the poverty line can potentially be the reaction to pressure of the people. Seligman (1894) argues that during the first half of the century achieving higher tax revenue was the primary motivation for tax reforms, while equity concerns became the driving motivation later on.

The third possible explanation is a competition for political power (Mares and Queralt, 2020). For this explanation it is crucial to understand the Prussian voting system. After the revolution in 1848, Prussia introduced a three-class franchise. In each district, each voter was aligned to one of three classes depending on his personal tax payment. The first class consisted of the minimum number of people responsible for one third of tax revenue. The minimum number of people other than those of the first class, who were responsible for another third of tax revenue, constituted the second class, while all others belonged to the third class. Therefore, the size of classes differed a lot. However, each class decided about the same number of electoral delegates, who elected the members of one chamber. Each class therefore had the same voting power. Thus, for example in 1849, 5% of the population had the same political power as 80% of the population (Becker and Hornung, 2020).

Each tax reform could change the composition of a voting class and therefore redistribute political power. Especially, a socially more just tax policy increased political inequality. Attempts to keep classes constant by fictitious additional tax payments failed. Thus, in 1888 the first class for example only consisted of 3.6% of voters (Buggeln, 2022). Due to Mares and Queralt (2020) this gave the opportunity for strategic actions between land owners and the rising industrial sector. From the perspective of the landed aristocracy a progressive tax system could push the tax burden to the industrial sector without weakening their own political power.

In the following sections we concentrate on the argument of state capacity. We show that all reforms increase the potential tax revenue even though various appealing features of the prevailing tax system were given up. The reforms can therefore be considered reasonable. Furthermore, we check if a majority benefits due to the reforms. We show that the reforms are politically feasible even thought the political economy of Prussia, with its three-class franchise, had its median voter in the top ten percent.

2.3 The Model

In this section, we use the workhorse model in the theory of optimal income taxation to better understand the evolution of taxes described in the previous section. It is based on Diamond (1998). We link our model to the different Prussian tax systems and show under which conditions a system change leads to a higher tax revenue and is supported by a majority of population and therefore said to be politically feasible.

Consider an economy with a continuum of individuals of mass one who differ in their abilities n. Individuals' utility u depends on consumption c and pretax income y: $u(c, y, n) = c - v(\frac{y}{n})$, with $v(\cdot)$ being an increasing and convex function measuring distaste of labor and $\frac{y}{n}$ describing the hours worked. Consumption depends on working time, individuals' abilities n and taxes T and is therefore given by $c = y - T(y)$. This results in

$$u(y, n, T) = y - T(y) - v\left(\frac{y}{n}\right). \tag{2.1}$$

We assume that preferences satisfy the Spence-Mirrlees single crossing property. Thus, individuals with higher abilities choose higher incomes than lower types. This ordering does not depend on the tax system. Abilities are distributed over an interval $[\underline{n}, \bar{n}]$ with the density function f and the corresponding cdf F.

We assume that the government wants to set taxes to maximize revenue $R = \int_{\underline{n}}^{\bar{n}} T(y)f(n)dn$. In this context, governments must not overtax individuals. Thus, $T(y^*) \leq y^* - v(\frac{y^*}{n}) - e$ needs to hold. Here, y^* denotes the income resulting from the worker's optimal labor choice and e a level of utility that the government needs to guarantee. We can interpret e as well as the utility level of an outlaw who escapes the state. Thus, taxes should not be so high that a worker's after-tax utility falls below this level, making escape from the state tempting. For simplicity, we assume $e = 0$ in the following. This leads to the participation constraints $T(y^*) \leq y^* - v(\frac{y^*}{n})$.[4]

The government knows the share of each ability type, but ability itself is private information of each individual. However, the government can observe individuals' incomes.[5] Each class i is represented by a distribution of abilities F_i over an interval $[\underline{n}_i, \bar{n}_i]$. The highest class is denoted by m.[6] It is assumed to be observable whether an individual belongs to class i or to another class.

With regard to the political feasibility of reforms and the question of whether individuals benefit from a tax change, the use of tax revenue is essential. If it were not taken into account, people could only benefit by reducing tax payments. However, if people benefit in other ways from increased tax revenue, they might also be in favor of increasing tax payments. We assume that all tax revenue is returned to the tax payers with a uniform lump sum transfer. This mechanism can be understood as a basic consumption level, offered by the government, or as the provision of a public good. Denoting old tax regimes by *old* and new ones by *new*, yields the condition for support

[4]One could also link an overtaxation to consumption instead of utility. In this case, the government would need to make sure that consumption is non-negative, yielding the participation constraint $T(y^*) \leq y^*$. This approach allows a different interpretation of individuals' opportunities and the role of a government. Here, avoiding taxes by living as an outlaw is not possible. The government will find the individual, even if she is living as an outlaw, and will collect taxes. One can argue for both approaches and both lead to similar results. In the following, we stick to the non-negative utility.

[5]It is reasonable to assume that the government did not observe incomes while taxing classes in years before 1851. This allowed to save money for administration purposes, what was an additional advantage of the class tax. To keep the analysis comparable for the different settings, we concentrate on a scenario with observable incomes during all tax regimes.

[6]Note that this is not in line with the ordering of social classes in Prussia, but more intuitive taking the increasing scale of abilities into account.

by an individual of ability n

$$R^{new} + y^{*,new} - T^{new}(y^{*,new}) - v\left(\frac{y^{*,new}}{n}\right) > R^{old} + y^{*,old} - T^{old}(y^{*,old}) - v\left(\frac{y^{*,old}}{n}\right). \tag{2.2}$$

In the following, we present scenarios with different tax systems along the historical development in Prussia. We start with a poll tax and continue with a lump sum tax for social classes. We then examine a mixed system of a class tax and an income tax, before evaluating a comprehensive income tax. Thus, we look more closely at system changes at two different levels: First, from lump sum taxes to an income tax and second, from no tagging to tagging and back to no tagging. Sometimes these two developments occur simultaneously. For each scenario, we describe the revenue maximizing tax payments and examine the conditions under which the reforms increase tax revenue and can therefore be understood as a measure to increase state capacity. We also examine if the reforms are supported by a majority of population when tax revenue is rebated lump sum.

2.3.1 A Poll Tax Before 1820

Before 1820 a poll tax was levied in Prussia. In the presence of this poll tax individuals solve $\max_{\{c,y\}} c - v\left(\frac{y}{n}\right)$ subject to $c = y - \tilde{T}$. The solution to this problem does not depend on $\tilde{T}$ and is denoted by $y^*(n)$, where $y^*(n)$ solves $1 = \frac{1}{n}v'\left(\frac{y}{n}\right)$. It is the first-best level of earnings for an individual of type n. The government chooses $\tilde{T}$ to maximize revenue but must ensure that no participation constraint is violated. No worker should be forced to pay a tax amount she cannot actually afford.

Depending on the ability distribution, it could be income-maximizing to choose a high $\tilde{T}$, so that some participation constraints are violated, but the higher incomes of individuals with higher abilities is better exploited. The implied case distinctions about the distribution complicate the analysis without really providing new insights. We therefore assume for the whole following analysis that the government does not want to violate any participation constraint.

Thus, the participation constraint of agents with the lowest ability $\underline{n}$ determines the tax $\tilde{T} = y^*(\underline{n}) - v\left(\frac{y^*(\underline{n})}{\underline{n}}\right)$ and therefore revenue

$$R^P = \int_{\underline{n}}^{\bar{n}} \tilde{T} f(n)dn = y^*(\underline{n}) - v\left(\frac{y^*(\underline{n})}{\underline{n}}\right).$$

2.3.2 The Introduction of a Class Tax in 1820

Between 1820 and 1851, the Prussian system was based on class taxation. This means that the government could set a lump sum tax for each class. The general setting can be described by the government's maximization problem

$$\max_{\{T_i\}_{i=1}^m} \sum_{i=1}^{m} \int_{\underline{n}_i}^{\bar{n}} T_i f_i(n) dn \tag{2.3}$$

$$s.t. \quad \{u = y - T_i - v\left(\frac{y}{n}\right) > 0, \quad \text{for} \quad \underline{n}_i \leq n \leq \bar{n}_i\}_{i=1}^m. \tag{2.4}$$

Again, tax payments are determined by participation constraints of the respective classes. It is $T_i = y^*(\underline{n}_i) - v\left(\frac{y^*(\underline{n}_i)}{\underline{n}_i}\right)$ and thus revenue is given by

$$R^C = \sum_{i=1}^{m} F_i \left[y^*(\underline{n}_i) - v\left(\frac{y^*(\underline{n}_i)}{\underline{n}_i}\right)\right].$$

Evaluating the shift from a poll tax towards a class tax yields the following Proposition:

Proposition 7.

(a) A system change from a poll tax to a class tax increases tax revenue.

(b) For disjunct classes, a system change from a poll tax to a class tax is supported by a majority iff the individual of median ability supports the reform.

As higher classes and therefore higher skilled workers can now be taxed separately, the system change obviously increases tax revenue.

Part (b) of the Proposition shows that the reform can, moreover, be politically feasible. Here we use inequality (2.2) to check if an individual benefits from the reform. For political feasibility, it is sufficient to show that an agent of median ability benefits. This is due to the monotonicity of the reform whenever class intervals are disjunct[7]. With a poll tax everybody pays the same taxes. Introducing a class tax allows to levy higher taxes on higher classes. Thus, the change of taxes is monotonic in ability. If the agent with median ability benefited from the reform, which means that the increase in tax revenue more than offsets her higher tax payments, this would also be true for lower abilities. Hence, more than half of the population would support the reform.

[7]If classes were not disjunct and the median voter's ability were in an intersection of classes, she would still be the decisive voter for political feasibility, if she were member of the highest of the intersecting classes.

The observation that a median voter theorem applies to monotonic tax reforms is due to Bierbrauer et al. (2021). This paper's contribution is to show that the transition of uniform lump sum taxation to class lump sum taxation actually is monotonic. Thus, the theorem applies. For more details, all proofs can be found in the appendix.

2.3.3 The Combination of Class and Income Taxes in 1851

Following the development of the Prussian tax system, we now move to a situation in which the state combines a class tax and an income tax. After the Prussian reform of 1851, the lower classes faced lump sum taxes as before, while the highest class was taxed depending on income. Thus, we replace the last summand of the maximization problem (2.3) by an income-dependent tax payment $T(y)$ and keep the lump sum taxation in all other classes. Since the government cannot regard workers' abilities, the choice of tax payments must ensure that workers do not want to lie about their abilities. Therefore, incentive constraints for the highest class m must be added to the maximization problem:

$$y^*(n) - T(y^*(n)) - v\left(\frac{y^*(n)}{n}\right) \geq y^*(n') - T(y^*(n')) - v\left(\frac{y^*(n')}{n}\right) \quad \forall\, n, n' \geq \underline{n}_m$$

Note that the participation for the lowest ability type already implies the participation of all higher types if the incentive conditions are met. Thus, the incentive constraints replace the participation constraints except for the one of the lowest ability type:

$$y^*(\underline{n}_m) - T(y^*(\underline{n}_m)) - v\left(\frac{y^*(\underline{n}_m)}{\underline{n}_m}\right) > 0. \tag{2.5}$$

Solving this part of the government's maximization problem leads to the following equation

$$\frac{T'(y(n))}{1 - T'(y(n))} = \frac{1 - F(n)}{f(n)n}\left(1 + \frac{1}{\epsilon}\right), \tag{2.6}$$

where ϵ is the elasticity of labor supply with respect to the net wage rate. It is a variation of the famous ABC formula of Diamond (1998). The marginal tax rate is higher, the lower the hazard rate $\frac{f(n)}{1-F(n)}$ is. Here, $f(n)$ captures the part of the population with negative behavioral responses when $T'(y)$ is increased. For the part $1-F(n)$ this change in marginal tax rate does not lead to behavioral responses, but increases tax revenue. Thus, the lower $f(n)$ and the higher $1 - F(n)$ are, the higher is the optimal marginal tax. However, this effect is scaled by the elasticity of labor supply. The higher ϵ and therefore the behavioral responses, the lower is the optimal marginal tax rate.

Total tax revenue in this tax system can be described by

$$
\begin{aligned}
R^{CI} &= \sum_{i=1}^{m-1} F_i T_i + \int_{\underline{n}_m}^{\bar{n}_m} \left(T(y^*(\underline{n}_m)) + \int_{\underline{n}_m}^{x} T'(y^*(n)) dn \right) f(x) dx \\
&= \sum_{i=1}^{m-1} F_i \left(y^*(\underline{n}_i) - v\left(\frac{y^*(\underline{n}_i)}{\underline{n}_i} \right) \right) \\
&\quad + \int_{\underline{n}_m}^{\bar{n}_m} \left(T(y^*(\underline{n}_m)) + \int_{\underline{n}_m}^{x} \frac{1}{\frac{f(n)n}{1-F(n)} \frac{\epsilon}{1+\epsilon} + 1} dn \right) f(x) dx,
\end{aligned}
$$

where $T(y^*(\underline{n}_m)) = y^*(\underline{n}_m) - v\left(\frac{y^*(\underline{n}_m)}{\underline{n}_m}\right)$ is determined by the participation constraint (2.5).

Evaluating the shift from a class tax towards class and income taxes yields the following Proposition:

Proposition 8.

(a) *A system change from a class tax to class and income taxes increases tax revenue.*

(b) *For disjunct classes, a system change from a class tax to class and income taxes is supported by a majority iff the individuals of median ability benefit from the reform.*

The first part of the Proposition summarizes the reform's revenue increasing effects. For the lower $m-1$ classes, taxes do not change. Thus, they have no effect on tax revenue. For the highest class, this is different. Here, the government can exploit higher incomes due to higher ability. Workers with higher incomes now have to pay higher taxes. This more than offsets the implied distortion of individuals' labor decisions. Technically, the pre-reform lump sum tax is still in the planner's choice set for the new maximization problem. However, she chooses a different tax schedule due to equation (2.6). Therefore, tax revenue increases compared to the previous class tax.

Part (b) of the Proposition shows that similar to Proposition 7, the worker with median ability is crucial for political feasibility. If she profits, the increase in revenue more than outweighs any potential increase in tax payments. Since the change of tax payments in the highest class is monotonic (see equation (2.6)), all agents in this class with a lower ability also benefit from the reform. The same is true for all lower classes, since taxes do not change there. Thus, the median ability is decisive. Again, the assumption of disjunct classes can be replaced by ordering the people by their social class.

2.3.4 Turning to an Income Tax in 1873 and 1891

After further reforms, Prussia completely stopped taxing people based on their social classes. Therefore, this subsection formalizes a tax system based only on incomes. The developments of the last subsection are now extended from the highest class to the entire population. Incentive constraints for all agents need to be taken into account and replace participation constraints except for the one of the lowest ability. The marginal tax rate defined by equation (2.6) now applies to the entire population. This yields the following tax revenue:

$$R^I = \int_{\underline{n}}^{\bar{n}} \left(T(y^*(\underline{n})) + \int_{\underline{n}}^{x} T'(y^*(n))dn \right) f(x)dx$$

The properties of this system change are summarized in the following Proposition:

Proposition 9.

(a) A system change from class and income taxes to an income tax increases tax revenue iff

$$F_m < \frac{\sum_{i=1}^{m-1} \left[(T(y^*(\underline{n})) - T(y^*(\underline{n}_i)))F_i + \int_{\underline{n}_i}^{\bar{n}_i} \int_{\underline{n}}^{x} T'(y^*(n))dnf(x)dx \right]}{T(y^*(\underline{n})) - T(y^*(\underline{n}_m))}, \qquad (2.7)$$

where $T(y^*(n)) = y^*(n) - v\left(\frac{y^*(n)}{n}\right)$ *and* $T'(y) = \frac{1}{1+\frac{f(n)n}{1-F(n)}\frac{\epsilon}{1+\epsilon}}$.

(b) A system change from class and income taxes to an income tax, that increases tax revenue, is supported by a majority if $F_m > \frac{1}{2}$.

In the new system, discrimination between classes is no longer possible, but the government can now discriminate within classes. The first part of the Proposition shows that this can increase tax revenue. However, it depends on the distribution of the population and its exact behavior as there are several counteracting effects. In each former class, tax payments can now vary with incomes. This makes higher tax payments possible.

At the same time, taxes now distort labor decisions. As seen in the former subsection and Proposition 8, this alone would still increase revenue. However, the elimination of all classes limits the possibility to distinguish between different abilities. Formally, the participation constraints in (2.4) are replaced by stricter incentive compatibility constraints. Only the participation constraint for the lowest ability is still part of the maximization problem. Depending on the shape of the marginal tax function and the structure of former classes, revenue could decrease or increase.

For the highest class, however, the effects of the system change are clear. The marginal tax rate has not changed. For individuals of ability $\underline{n}_m$ the participation constraint is replaced by a stricter incentive constraint. Thus, the tax payments of these individuals were at least as high in the old system as in the new system. The entire class faces a reduction or at least no rise in taxes. If the size of this class is small enough and tax revenue increases for the other classes, total tax revenue increases. This is captured by inequality (2.7).

Part (b) of Proposition 9 gives a sufficient condition for political feasibility. If revenue increases, members of the highest class benefit from the reform because they do not have to pay higher taxes. If their size is greater than $\frac{1}{2}$, a majority supports the reform.

2.3.5 Convergence of Classes

Comparing the previous results with the evolution of the Prussian tax system, we see that all reforms might have been useful under the perspective of state capacity. The replacement of the poll tax by a class tax in 1820 definitely led to higher tax revenue. The replacement of the lump sum taxes by the income tax for the top class in 1851 also increased tax revenue. Things were less clear for the reforms of 1873 and 1891, when the remaining lump sum taxes were replaced by an income tax without any tagging. This last reform was the first one that stopped using tagging after its first use in 1820 and it is not obvious why the government implemented this reform.

We have seen in Proposition 9 that these reforms might have increased revenue. However, revenue could have been further increased by additionally using tagging. In the sense of the reform in 1851, one could have distinguished social classes and could have used income taxes inside these classes (see subsection 2.3.3). Tagging is an additional possibility to distinguish individuals' abilities and using this additional information cannot make things worse. Why, then, did the government abolish tagging?

In the previous considerations, we modeled classes as intervals of abilities, each with its own distribution. We also assumed classes to be correlated with abilities, meaning that higher classes contained higher abilities on average. Concerning revenue gains, we did not assume classes to be disjunct[8]. Thus, classes could potentially overlap. There are some signs that this overlapping increased over time or at least the correlation of classes and abilities became weaker. This was, for example, the view of Dr. Ernst Engel, director of *Königlich Preußisches Statistisches Bureau*. In 1878 he wrote:

[8]We only used disjunct classes for political economy statements in Propositions 7 and 8.

> Although it may have been necessary, as J.G. Hoffmann states in his classical theory of taxes, to link the latter to large general characteristics in the transition from indirect to partially direct taxation, this procedure could no longer be maintained after the characteristics themselves had become highly deceptive, and a tax assessment merely according to these characteristics resulted in the greatest inequalities.[9]

In the beginning, tagging was useful, but classes became more and more arbitrary over time. With this development the class tax led to different tax payments for people with similar earnings or - in terms of our model - for same-skilled workers. Horizontal equity in the sense that tax payments should be the same for individuals of the same ability, is therefore violated. This is also the case for similar concepts as, for example, the principle of equal sacrifice (Weinzierl, 2012). This violation could be an explanation for the abolition of class taxation. Moreover, the development can as well weaken the advantages of tagging concerning tax revenue. To show this formally, we use a perturbation approach. In the following, we assume only two classes to exist, each with a distribution over abilities $f_i(n)$, for $i \in \{1,2\}$ and $n \in [\underline{n}, \bar{n}]$. The perturbation of class' i distribution is denoted by $f_i(n) + \delta g(n)$. Here, $g(n)$ is the derivative of $G(n)$ with $G(\underline{n}) = G(\bar{n}) = 0$. This ensures that the properties of a distribution still apply. To compare the case of tagging to the one of no tagging, tax revenue is compared and denoted by R^T and R^{NT}, respectively.

Using this notation, we can state the following Proposition.

Proposition 10.

(a) Assume that the ability distributions over the two existing classes are given by $f_1(n)$ and $f_2(n)$, with $f_2(n) = f_1(n) + \delta g(n)$. Then, it is $\lim_{\delta \to 0} R^T - R^{NT} = 0$.

(b) Assume that the ability distributions over the two existing classes are given by $f_1(n)$ and $f_2(n)$, with $f_2(n) = f_1(n - \delta)$. Then, it is $\lim_{\delta \to 0} R^T - R^{NT} = 0$.

Part (a) of this Proposition is about a perturbation of an existing distribution, while part (b) is about a shift along abilities. Reducing both, and therefore making the classes more and more similar reduces the benefits of tagging, although they are always positive. This result is intuitive. The contribution of this paper is, however, to show it formally. The proofs can be found in the appendix.

A government could for example set a minimum level of additional revenue for tagging to be considered beneficial taking the disadvantages of horizontal equity and feelings

[9] See (Engel, 1868), page 115.

of injustice or higher administrative costs due to tagging into account. Proposition 10 states that converging distributions at some point fall below this level. This may explain why the government no longer used tagging. At some point gains of tagging are small enough to be outweighed by other considerations.

2.3.6 Pareto-Improving Reforms

In this subsection we return to political feasibility. We have seen that a majority of population can benefit from the reforms. This is also true taking the special Prussian voting system into account (see section 2.4). In all these cases, people with low abilities are in favor of the reforms, but at least parts of the high-ability elites are not.

One could argue that these elites are essential for successful reforms although formally there is a political majority without them. In this case it would be a puzzle why they decided to change the tax system and impose higher taxes on themselves. As discussed earlier, they could support the reforms for other reasons, such as strategic considerations towards certain subgroups of elites (Mares and Queralt, 2020). However, one can explain this puzzle without these considerations. One possibility is to state that higher tax revenue increases production and therefore consumption (Barro, 1990). An even more intuitive way is to think of high marginal returns concerning public goods which are financed by taxes. This scenario is especially likely in a situation with a low state activity as it was the case in Prussia at some point.

We formalize this by changing the utility function to the following form:

$$u(y, n, T, R) = \varphi\left(ny - T(y) - v\left(\frac{y}{n}\right)\right) + w(R).$$

Here, $\varphi(\cdot)$ and $w(\cdot)$ are increasing and concave functions measuring utility out of consumption and the public good, respectively. Due to the concavity of the functions, wealthy people can benefit from giving up consumption and getting more of the public good instead: As they have a high consumption in the status-quo and a much lower amount of the public good, a reduction of consumption would possibly not decrease utility as much as more of the public good would increase it. In this case a reform can be Pareto-improving. This of course depends on the curvature of the functions $\varphi(\cdot)$ and $w(\cdot)$. The following Proposition summarizes the requirements for Pareto-improvement on the example of the system change from a class tax to class and income taxes.

Proposition 11. *A system change from a class tax to class and income taxes is Pareto-improving iff*

$$\int_{R^{old}}^{R^{new}} w'(R)dR > \int_{c^{new}(n)}^{c^{old}(n)} \varphi'(c)dc. \;\; \forall n \geq \underline{n}_m, \tag{2.8}$$

with

$$\begin{aligned}
R^{old} &= \sum_{i=1}^{m} F_i \left(y^*(\underline{n}_i) - v\left(\frac{y^*(\underline{n}_i)}{\underline{n}_i}\right)\right),\\
R^{new} &= \sum_{i=1}^{m-1} F_i \left(y^*(\underline{n}_i) - v\left(\frac{y^*(\underline{n}_i)}{\underline{n}_i}\right)\right) + \int_{\underline{n}_m}^{\bar{n}_m} \left(T(y^*(\underline{n}_m)) + \int_{\underline{n}_m}^{x} T'(y^*(n))dn \right) f(x)dx,\\
c^{old}(n) &= y^*(n) - v\left(\frac{y^*(n)}{n}\right) - \left(y^*(\underline{n}_m) - v\left(\frac{y^*(\underline{n}_m)}{\underline{n}_m}\right)\right),\\
c^{new}(n) &= y^*(n) - v\left(\frac{y^*(n)}{n}\right) - \left(y^*(\underline{n}_m) - v\left(\frac{y^*(\underline{n}_m)}{\underline{n}_m}\right) + \int_{\underline{n}_m}^{n} T'(y(n))dn \right),\\
T'(y) &= \frac{1}{\frac{f(n)n}{1-F(n)}\frac{\epsilon}{\epsilon+1} + 1}.
\end{aligned}$$

The Proposition shows, that the reform can be Pareto-improving if the curvatures of the functions are appropriate for the corresponding changes in consumption and tax revenue. Note that inequality (2.8) must hold for all $n \geq \underline{n}_m$, since it is not clear who in this group suffers most from the reduction in consumption. This reduction is highest for ability type $\bar{n}_m$. However, due to the curvature of φ, the utility loss could be higher for lower ability types. At the same time, inequality (2.8) does not need to hold for $n < \underline{n}_m$ since their tax payment and their consumption do not change. They benefit from the reform.

2.4 Empirics

In this section, we aim to connect our theoretical results to historical data and to better understand the various tax reforms. Remember, that we have formulated three Propositions concerning the reform from a poll tax to a class tax, from a class tax to a combination of class and income taxes and from the combination of class and income taxes to a pure income tax. Each Proposition formulates conditions for an increase in revenue as well as for a majority supporting the reform. We now use historical data to check if these conditions hold in reality.

In the Propositions' proofs the monotonicity of reforms plays an essential role. In reality, however, reforms are not always monotonic since reforms are not only about the system change but at the same time also for example about reducing the tax burden for poor or rich people. Thus, we do additionally check the reforms' monotonicity and if they are indeed revenue increasing and politically feasible.

In the last sections, we called a reform politically feasible whenever a majority of people supported it. In Prussia, however, it also makes sense to take the special three-class franchise (see Section 2.1) into account. This makes things different as a voting power majority is now in general not the same as a population majority.

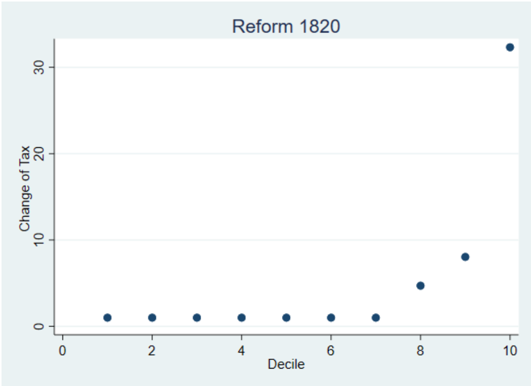
Obviously, this has consequences for the definition of a median voter as well. In the median voter theorem of Bierbrauer et al. (2021) the median voter is of interest as she defines a threshold for population's voting power. Having all people with higher incomes than the median voter and the median voter himself being in favor of a reform results in political feasibility. In the Prussian voting system after 1849, someone with the median income does not define the crucial point for population's voting power anymore as people with higher incomes have higher voting power. Here the threshold is defined by the one with median income in voting class two.

This new definition transfers to our Propositions (7), (8) and (9) as well. In these Propositions the decisive voter for support of a majority was the one with median ability. Taking the three-class franchise into account, the decisive voter now is the one with median ability inside voting class two.

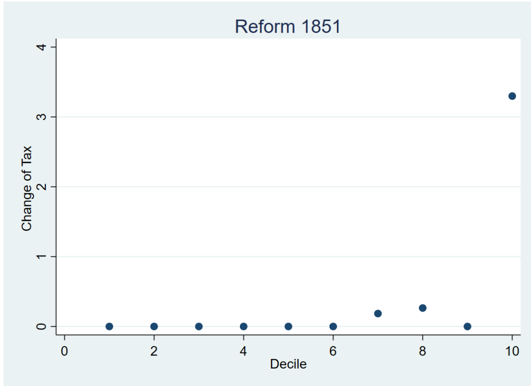
We keep the assumption from section 2.3 that tax revenue is rebated lump sum. Thus inequality (2.2) determines if somebody supports a reform.

For the checks in this section, knowing the distribution of incomes in reform years is necessary. In Engel (1868) and Engel (1875) one can find income distributions that have been used to estimate future tax revenue for 1851 and 1873. Furthermore, Königliches Statistisches Bureau (1898) offers a way more detailed distribution of incomes for 1891. Thus, it is only the reform in 1820 for which we do not have a distribution. Here we use the one of 1851 again.

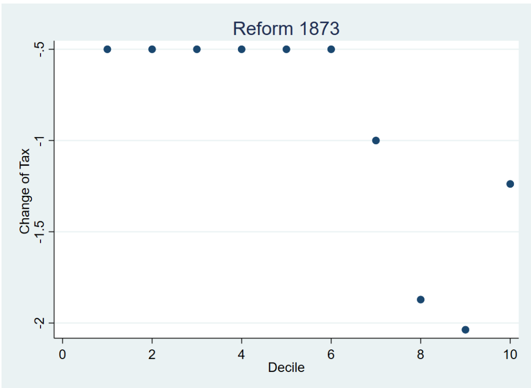
Note, that data is limited in several ways. For example, we do not know how income distributions change after a reform. We therefore take the distributions constant and ignore potential behavioral responses to tax reforms, but only look at mechanical effects. Or to put it in another way, we assume the elasticity of taxable income to be zero. Furthermore, we do not know the residence of people. Thus, we cannot correct our data analysis for the people who live in larger cities and face a different taxation up to 1873 with the grist and slaughter tax. Hence, the following analysis can only be a sketch of reality, but it can still visualize and better explain reforms.



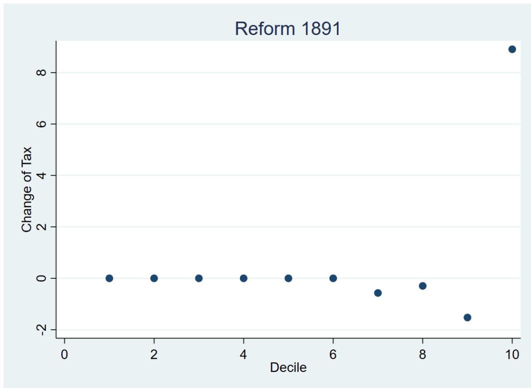
(a) Reform in 1820. Own Calculation Based on Engel (1868) and Schremmer and Stern (1989).



(b) Reform in 1851. Own Calculation Based on Engel (1868) and Schremmer and Stern (1989).



(c) Reform in 1873. Own Calculation Based on Schwarz and Strutz (1902), Engel (1875) and Königliches Statistisches Bureau (1898).



(d) Reform in 1891. Own Calculations Based on Schwarz and Strutz (1902), Engel (1875) and Königliches Statistisches Bureau (1898).

Figure 2.1: Change in Tax Burden Depending on Deciles for Different Reforms

The Introduction of a Class Tax in 1820

The reform in 1820 was from a poll tax to a class tax. In figure 2.1a the average change in tax burden for income deciles is shown based on the distribution in Engel (1868). First of all, we can notice, that there is no tax cut for anybody and that the reform is monotonic. In Proposition 7 we showed that revenue increases and that the reform is politically feasible if the median voter benefits.

Looking at the data, we see that revenue indeed increased. The median voter here is described by the one with median income as the three-class franchise did not exist at this time. She had to pay 0.5 Thaler before the reform and 1.5 Thaler afterwards. Revenue per capita increased from 0.5 to 5.7. Thus, the median voter benefited. At the same time, more than 80% of the population benefited. Therefore, the reform indeed is political feasible and Proposition 7 is in line with reality.

The Combination of Class and Income Taxes in 1851

As mentioned in section 2.2, the financial situation of Prussia before 1851 was bad. Hence, it needed to raise more taxes. In 1851 it changed the system by transforming the class tax for the highest social class into an income tax. Proposition 8 shows that this leads to an increase in tax revenue and is political feasible if and only if the median voter benefits.

Figure 2.1b shows that there was no tax cut for anybody: The lowest five deciles did not face a change in taxes, while taxes increased for higher deciles. Obviously, this indeed led to higher tax revenue.

Taking the three-class franchise into account, the median voter belonged to the ninth decile. She had to pay 4 Thaler per year in taxes before and after the reform. As tax revenue increased, the reform was clearly beneficiary for her.

Since the reform was not completely monotonic, it is not obvious if it was indeed politically feasible. However, Bierbrauer et al. (2021) show that the median voter is often a good estimator for political feasibility if the reform is monotonic at large. Here this holds true, as individuals with a majority of voting power benefited.

Turning to an Income Tax in 1873 and 1891

The reforms in 1873 and 1891 both are parts of a transition to a pure income tax. The situation before these reforms was different to the one in 1851. As explained in section 2.2, the state's financial situation was good. Thus, the reforms were not only about raising taxes, but also about relieving the burden on individuals with lower incomes. Our data show, that tax revenue decreased during the reform in 1873 and increased in 1891. This is in line with Engel (1868) and Spoerer (2015), respectively.

This development is visualized by the graphs in figure 2.1c and 2.1d. For 1873 there is a decrease in tax burden for each decile. Only some parts of the tenth decile face a higher tax burden. At the same time, this leads to a non-monotonic part of the reform. Things are a bit different for 1891. In the first six deciles, there is no change in tax payments. After that, tax burden decreases, but is finally much higher in the tenth decile. Here, the eighth and tenth deciles are exceptions to monotonicity.

Concerning political feasibility, Proposition 9 formulates a sufficient condition ($F_m > \frac{1}{2}$). This condition, however, is not satisfied. As explained before, the median voter can be a good estimator for a reform that is monotonic at large. But this is only the case for the reform in 1891 and not for the one in 1873. For the latter, the median voter was

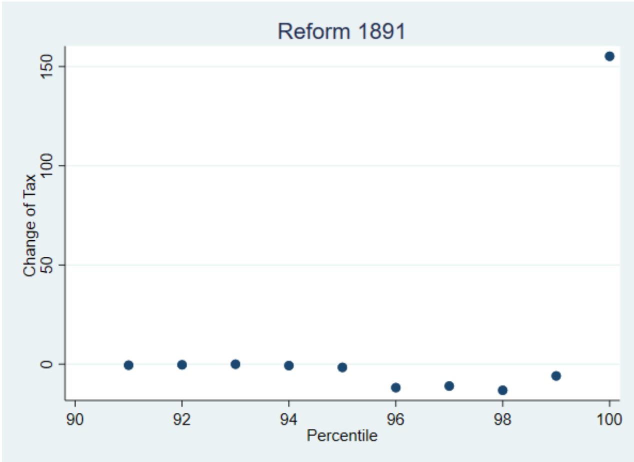


Figure 2.2: Change in Tax Burden Inside the Tenth Decile for the Reform in 1891. Own Calculation Based on Schwarz and Strutz (1902), Engel (1875) and Königliches Statistisches Bureau (1898).

part of the ninth income decile and faced a decrease of tax revenue per capita of 2.7 while her tax payment decreased from 24 to 15 Mark. Thus, she benefited due to the reform. However, individuals with a majority of voting power did not benefit meaning that this reform is not politically feasible.

This looks completely different in 1891. The median voter was part of the tenth decile and the per capita increase of tax revenue was 0.6 Mark while her tax payments per year decreased from 126 to 118. This decrease cannot be seen in figure 2.1d. Zooming in shows how heterogeneous this decile is. Figure 2.2 presents the ten percentiles inside the tenth decile. Only the people with the highest incomes face a dramatic increase of tax burden yielding the increase in total tax revenue. Thus, most of the people support the reform and it is politically feasible. Here, the support of median voter and political feasibility are in line again.

Note, that the shift of median voter's decile from the ninth to the tenth is the result of a higher inequality in tax payments. It can be explained by higher taxes for the rich after the 1873 reform, by zero taxes for the poor after the abolition of the lowest classes in 1883 and by a general change of income distribution between 1873 and 1891.

2.5 Conclusion

This article takes a look at certain Prussian tax reforms. During these reforms, lump sum taxes and tagging were used and abolished after some time, although they are useful instruments from a modern tax theory perspective. Augmenting the model of

Diamond (1998), we explore reasons for this kind of reforms. We find that all reforms can lead to a higher tax revenue and can be politically feasible. Furthermore, we discuss limits of tagging and conditions for Pareto-improvements.

In a final step, we use historical data to check our theoretical results. Most reforms increase tax revenue and are politically feasible taking the Prussian three-class franchise into account.

Future research could try to improve historical data analysis. Further assumptions on income distributions could possibly solve the problems caused by data limitation. Moreover, one could also take the grist and slaughter tax into account as it is connected to income and class taxes.

2.A Appendix

Proof of Proposition 7

(a) Due to the participation constraints, a poll tax is defined by the lowest ability type. It is $\tilde{T} = \underline{n}v'^{-1}(\underline{n}) - v(v'^{-1}(\underline{n}))$. Thus, total revenue is $R^P = \underline{n}v'^{-1}(\underline{n}) - v(v'^{-1}(\underline{n}))$.

Using classes allows for different taxes. Each is defined by the lowest ability in a class. Total revenue therefore is given by $R^C = \sum_{i=1}^{m} \underline{n}_i v'^{-1}(\underline{n}_i) - v(v'^{-1}(\underline{n}_i))$. Utility is an increasing function. Thus, for higher abilities the tax T_i can be set higher. This yields a higher total tax amount.

(b) This is a monotonic tax reform since the additional amount of paid taxes increases monotonically by ability. Thus, the median voter theorem of Bierbrauer et al. (2021) applies.

□

Proof of Proposition 8

(a) After this system change, all individuals face the same tax payments as before, except for the one in the highest class. Here, it is

$$\frac{T'(y(n))}{1 - T'(y(n))} = \frac{1 - F(n)}{f(n)n}\left(1 + \frac{1}{\epsilon}\right).$$

Thus, tax payments for this class differ compared to the ones before the system change. Since the old lump sum taxes are in the possible choice set of the maximization problem, but they are not chosen by the planner, tax revenue increases in the new system.

(b) As tax revenue increases, all lower classes benefit due to the reform. In the highest class, the reform is monotonic. Thus, there exists a threshold in the highest class. Below the threshold everybody benefits, above nobody does. Thus, the reform is supported by a majority if and only if the median ability individual is below this threshold. Otherwise more than half of the population would not benefit and therefore not support the reform.

□

Proof of Proposition 9

(a) In a system of class and income taxation, tax revenue consists of class tax and Income tax. It can be summarized by

$$\begin{aligned} R_{CI} = & \sum_{i=1}^{m-1} [n y_C^*(\underline{n}_i) - v(y_C^*(\underline{n}_i))] F_i \\ & + \int_{\underline{n}_m}^{\bar{n}} \left[n y_I^*(\underline{n}_m) - v(y_I^*(\underline{n}_m)) + \int_{\underline{n}_m}^{x} T'(y^*(n)) dn \right] f(x) dx \end{aligned}$$

with $y_C^*(n) = nv'^{-1}(n) - v(v'^{-1}(n))$ and $y_I^*(n) = nv'^{-1}((1-T'(y))n) - v(v'^{-1}((1-T')n))$. After the reform it is

$$R_I = \int_{\underline{n}}^{\bar{n}} \left[T(y^*(\underline{n})) + \int_{\underline{n}}^{x} T'(y^*(n)) dn \right] f(x) dx.$$

Thus, the change in tax revenue is given by

$$\begin{aligned} \sum_{i=1}^{m-1} & \left[\left(y_I^*(\underline{n}) - v\left(\frac{y_I^*(\underline{n})}{\underline{n}} \right) - \left[y_C^*(\underline{n}_i) - v\left(\frac{y_C^*(\underline{n}_i)}{\underline{n}_i} \right) \right] \right) F_i + \int_{\underline{n}_i}^{\bar{n}} \int_{\underline{n}}^{x} T'(y) dn f(x) dx \right] \\ & + \left(y_I^*(\underline{n}) - v\left(\frac{y_I^*(\underline{n})}{\underline{n}} \right) - \left[y_I^*(\underline{n}_m) - v\left(\frac{y_I^*(\underline{n}_m)}{\underline{n}_m} \right) \right] \right) F_m. \end{aligned}$$

This is positive whenever

$$F_m < \frac{\sum_{i=1}^{m-1} \left[(T(y^*(\underline{n})) - T(y^*(\underline{n}_i))) F_i + \int_{\underline{n}_i}^{\bar{n}_i} \int_{\underline{n}}^{x} T'(y^*(n)) dn f(x) dx \right]}{T(y^*(\underline{n})) - T(y^*(\underline{n}_m))}.$$

(b) Due to the additional incentive constraints, tax payment $T(\underline{n}_m)$ for individuals of type $\underline{n}_m$ cannot increase after the reform, but decreases. As the marginal tax rate does not change for all individuals with $n > \underline{n}_m$, all individuals of group m face lower tax payments. As the total revenue increases, they definitely profit due to the reform. If this group's mass is more than half of the population, a majority supports the reform.

□

Proof of Proposition 10

(a) There are two classes $i = 1, 2$ containing the interval $[\underline{n}, \bar{n}]$. In the first one, individuals are distributed due to the density function $f_1(n)$. For the second one, it is $f_2(n) = f_1(n) + \delta g(n)$, where $g(n)$ is the derivative of $G(n)$. The latter is a continuously differentiable function with $G(\underline{n}) = G(\bar{n}) = 0$ and $\delta \geq 0$ is a scalar. The distribution over the whole population is given by $f(n) = f_1(n) + f_2(n)$ with the corresponding cumulative density function $F(n) = F_1(n) + F_1(n) + \delta G(n)$.

Without tagging, tax revenue of an income tax is given by

$$R^{NT} = \int_{\underline{n}}^{\bar{n}} \left[T(y(\underline{n})) + \int_{\underline{n}}^{x} \frac{1}{\frac{(2f_1(n)+\delta g(n))n}{1-2F_1(n)-\delta G(n)} \frac{\epsilon}{\epsilon+1} + 1} dn \right] (2f_1(x) + \delta g(x)) dx.$$

With Tagging, each class gets its own marginal tax rate and total revenue is given by

$$\begin{aligned} R^T = & \int_{\underline{n}}^{\bar{n}} \left[T(y(\underline{n})) + \int_{\underline{n}}^{x} \frac{1}{\frac{2f_1(n)n}{1-2F_1(n)} \frac{\epsilon}{\epsilon+1} + 1} dn \right] f_1(x) dx \\ & + \int_{\underline{n}}^{\bar{n}} \left[T(y(\underline{n})) + \int_{\underline{n}}^{x} \frac{1}{\frac{2(f_1(n)+\delta g(n))n}{1-2(F_1(n)-\delta G(n))} \frac{\epsilon}{\epsilon+1} + 1} dn \right] (f_1(x) + \delta g(x)) dx. \end{aligned}$$

Decreasing the differences in distributions between classes is equivalent to $\delta \to 0$. We want to show that $\lim_{\delta \to 0} R^T = \lim_{\delta \to 0} R^{NT}$. To do so, we define a function sequence $h_m(n)$, with

$$h_m(n) = \frac{1}{\frac{2(f_1(n)+\frac{1}{m}g(n))n}{1-2(F_1(n)+\frac{1}{m}G(n))} \frac{\epsilon}{\epsilon+1} + 1}.$$

We replace δ by $\frac{1}{m}$, thus $m \to \infty$ corresponds to $\delta \to 0$. As the fraction in the denominator of h_m is positive, we know that it is $h_m(n) < 1 \equiv H(n)$. Furthermore, h_m converges pointwise to $h(n) = \frac{1}{\frac{2f_1(n)n}{1-2F_1(n)} \frac{\epsilon}{\epsilon+1} + 1}$. The compact interval $[\underline{n}, x]$ and a pointwise converging sequence that is dominated by an integrable function $H(n)$ are conditions of the dominated convergence theorem. It implies that

$$\begin{aligned} \lim_{m \to \infty} \int_{\underline{n}}^{x} \frac{1}{\frac{2(f_1(n)+\frac{1}{m}g(n))n}{1-2(F_1(n)+\frac{1}{m}G(n))} \frac{\epsilon}{\epsilon+1} + 1} dn &= \int_{\underline{n}}^{x} \lim_{m \to \infty} \frac{1}{\frac{2(f_1(n)+\frac{1}{m}g(n))n}{1-2(F_1(n)+\frac{1}{m}G(n))} \frac{\epsilon}{\epsilon+1} + 1} dn \\ &= \int_{\underline{n}}^{x} \frac{1}{\frac{2f_1(n)n}{1-2F_1(n)} \frac{\epsilon}{\epsilon+1} + 1} dn. \end{aligned}$$

A similar argument holds for the outer integral. We define

$$k_m(x) = \left[T(y(\underline{n})) + \int_{\underline{n}}^{x} \frac{1}{\frac{2(f_1(n)+\frac{1}{m}g(n))n}{1-2(F_1(n)+\frac{1}{m}G(n))}\frac{\epsilon}{\epsilon+1}+1}dn\right]\left(f_1(x)+\frac{1}{m}g(x)\right).$$

Again, $k_m(x)$ converges pointwise to $k(x) = \left[T(y(\underline{n})) + \int_{\underline{n}}^{x} \frac{1}{\frac{2f_1(n)n}{1-2F_1(n)}\frac{\epsilon}{\epsilon+1}+1}dn\right] f_1(x)$ and is dominated by the function $K(x) = T(y(\underline{n}))+(x-\underline{n})(f_1(x)+|g(x)|)$. Thus, due to the dominated convergence theorem, we can shift the limit inside the outer integral and it is

$$\begin{aligned}
\lim_{m\to\infty} R^T =& \int_{\underline{n}}^{\bar{n}} \left[T(y(\underline{n})) + \int_{\underline{n}}^{x} \frac{1}{\frac{2f_1(n)n}{1-2F_1(n)}\frac{\epsilon}{\epsilon+1}+1}dn\right] f_1(x)dx \\
&+ \lim_{m\to\infty} \int_{\underline{n}}^{\bar{n}} \left[T(y(\underline{n})) + \int_{\underline{n}}^{x} \frac{1}{\frac{2(f_1(n)+\frac{1}{m}g(n))n}{1-2(F_1(n)+\frac{1}{m}G(n))}\frac{\epsilon}{\epsilon+1}+1}dn\right]\left(f_1(x)+\frac{1}{m}g(x)\right)dx \\
=& \int_{\underline{n}}^{\bar{n}} \left[T(y(\underline{n})) + \int_{\underline{n}}^{x} \frac{1}{\frac{2f_1(n)n}{1-2F_1(n)}\frac{\epsilon}{\epsilon+1}+1}dn\right] f_1(x)dx \\
&+ \int_{\underline{n}}^{\bar{n}} \left[T(y(\underline{n})) + \int_{\underline{n}}^{x} \frac{1}{\frac{2f_1(n)n}{1-2F_1(n)}\frac{\epsilon}{\epsilon+1}+1}dn\right] f_1(x)dx \\
=&2\int_{\underline{n}}^{\bar{n}} \left[T(y(\underline{n})) + \int_{\underline{n}}^{x} \frac{1}{\frac{2f_1(n)n}{1-2F_1(n)}\frac{\epsilon}{\epsilon+1}+1}dn\right] f_1(x)dx.
\end{aligned}$$

For the revenue in the non-tagging case, the argumentation is similar and it yields

$$\begin{aligned}
\lim_{\delta\to 0} R^{NT} =& \lim_{\delta\to 0} \int_{\underline{n}}^{\bar{n}} \left[T(\underline{n}) + \int_{\underline{n}}^{x} \frac{1}{\frac{(2f_1(n)+\delta g(n))n}{1-2F_1(n)-\delta G(n)}\frac{\epsilon}{\epsilon+1}+1}\right](2f_1(x)+\delta g(x))dx \\
=& \int_{\underline{n}}^{\bar{n}} \left[T(\underline{n}) + \int_{\underline{n}}^{x} \frac{1}{\frac{2f_1(n)n}{1-2F_1(n)}\frac{\epsilon}{\epsilon+1}+1}\right] 2f_1(x)dx.
\end{aligned}$$

Thus, it is

$$\lim_{\delta\to 0} R^T - R^{NT} = 0.$$

(b) Let R^T denote the revenue under the usage of tagging and R^{NT} the one without. First of all, notice that using tagging is equivalent to a maximization problem with fewer incentive constraints. Thus, it is always $R^T - R^{NT} \geq 0$.

There exist two classes. The distribution over both classes is exactly the same except for a shift δ in ability. Thus, the distribution over class 2 satisfies $f_2(n) = f_1(n-\delta)$, where $f_1(n)$ is the distribution of class 1. We follow earlier notation and assume the classes to be distributed between abilities $\underline{n}_1$ and $\bar{n}_1$ and $\underline{n}_2$ and $\bar{n}_2$, respectively. Here it is $\underline{n}_2 = \underline{n}_1 + \delta$ and $\bar{n}_2 = \bar{n}_1 + \delta$. It holds $\int_{\underline{n}_1}^{\bar{n}_1} f_1(n)dn = \int_{\underline{n}+\delta}^{\bar{n}+\delta} f_2(n)dn = \frac{1}{2}$. Thus, the distribution over the total population has the usual properties and integrates to one.

Tax revenue with tagging is given by

$$
\begin{aligned}
R^T = & \int_{\underline{n}}^{\bar{n}} \left[T(y(\underline{n})) + \int_{\underline{n}}^{x} \frac{1}{\frac{2f_1(n)n}{1-2F_1(n)}\frac{\epsilon}{\epsilon+1} + 1} \right] f_1(x)dx \\
& + \int_{\underline{n}+\delta}^{\bar{n}+\delta} \left[T(y(\underline{n}+\delta)) + \int_{\underline{n}+\delta}^{x} \frac{1}{\frac{2f_1(n-\delta)n}{1-2F_1(n-\delta)}\frac{\epsilon}{\epsilon+1} + 1} dn \right] f_1(x-\delta)dx \\
= & \int_{\underline{n}}^{\bar{n}} \left[T(y(\underline{n})) + \int_{\underline{n}}^{x} \frac{1}{\frac{2f_1(n)n}{1-2F_1(n)}\frac{\epsilon}{\epsilon+1} + 1} \right] f_1(x)dx \\
& + \int_{\underline{n}+\delta}^{\bar{n}+\delta} \left[T(y(\underline{n}+\delta)) + \int_{\underline{n}}^{x-\delta} \frac{1}{\frac{2f_1(n)(n+\delta)}{1-2F_1(n)}\frac{\epsilon}{\epsilon+1} + 1} dn \right] f_1(x-\delta)dx \\
= & \int_{\underline{n}}^{\bar{n}} \left[T(y(\underline{n})) + \int_{\underline{n}}^{x} \frac{1}{\frac{2f_1(n)n}{1-2F_1(n)}\frac{\epsilon}{\epsilon+1} + 1} \right] f_1(x)dx \\
& + \int_{\underline{n}}^{\bar{n}} \left[T(y(\underline{n}+\delta)) + \int_{\underline{n}}^{x} \frac{1}{\frac{2f_1(n)(n+\delta)}{1-2F_1(n)}\frac{\epsilon}{\epsilon+1} + 1} dn \right] f_1(x)dx.
\end{aligned}
$$

Using similar arguments as in the proof for part (a), we can apply the dominated convergence theorem that yields

$$
\begin{aligned}
\lim_{\delta \to 0} R^T = & \int_{\underline{n}}^{\bar{n}} \left[T(y(\underline{n})) + \int_{\underline{n}}^{x} \frac{1}{\frac{2f_1(n)n}{1-2F_1(n)}\frac{\epsilon}{\epsilon+1} + 1} dn \right] f_1(x)dx \\
& + \int_{\underline{n}}^{\bar{n}} \left[T(y(\underline{n})) + \int_{\underline{n}}^{x} \lim_{\delta \to 0} \frac{1}{\frac{2f_1(n)(n+\delta)}{1-2F_1(n)}\frac{\epsilon}{\epsilon+1} + 1} dn \right] f_1(x)dx \\
= & 2\int_{\underline{n}}^{\bar{n}} \left[T(y(\underline{n})) + \int_{\underline{n}}^{x} \frac{1}{\frac{2f_1(n)n}{1-2F_1(n)}\frac{\epsilon}{\epsilon+1} + 1} dn \right] f_1(x)dx.
\end{aligned}
$$

For the non-tagging case, it is

$$\lim_{\delta\to 0} R^{NT} = \lim_{\delta\to 0} \int_{\underline{n}}^{\bar{n}+\delta} \left[T(y(\underline{n})) + \int_{\underline{n}}^{x} \frac{1}{\frac{(f_1(n)+f_1(n-\delta))n}{1-F_1(n)-F_1(n-\delta)} \frac{\epsilon}{\epsilon+1} + 1} dn \right] (f_1(x) + f_1(x-\delta))dx.$$

Here, things are a more difficult as δ is in the integrand as well in its boundaries. We define

$$h(n,\delta) \equiv \left[T(y(\underline{n})) + \int_{\underline{n}}^{x} \frac{1}{\frac{(f_1(n)+f_1(n-\delta))n}{1-F_1(n)-F_1(n-\delta)} \frac{\epsilon}{\epsilon+1} + 1} dn \right] (f_1(x) + f_1(x-\delta)).$$

As a combination of continuous functions, $h(n,\delta)$ is continuous as well. Due to the Stone-Weierstrass theorem there exists a sequence of polynomials $p_k(n,\delta)$ of the form $p(n,\delta) = \sum_{i=1}^{I} \sum_{j=1}^{J} c_{ij} t^i \delta^j$ with $0 < I, J < \infty$ such that the sequence converges uniformly to $h(n,0)$. For these polynomials, obviously it is

$$\lim_{\delta\to 0} \int_{\underline{n}}^{\bar{n}+\delta} h(n,\delta)dn = \int_{\underline{n}}^{\bar{n}} h(n,0)dn.$$

Thus, it is

$$\begin{aligned}
\lim_{\delta\to 0} \int_{\underline{n}}^{\bar{n}+\delta} h(n,\delta)dn &= \lim_{\delta\to 0} \int_{\underline{n}}^{\bar{n}+\delta} \lim_{k\to\infty} p_k(n,\delta)dn \\
&= \lim_{k\to\infty} \lim_{\delta\to 0} \int_{\underline{n}}^{\bar{n}+\delta} p_k(n,\delta)dn \\
&= \lim_{k\to\infty} \int_{\underline{n}}^{\bar{n}} p_k(n,0)dn \\
&= \int_{\underline{n}}^{\bar{n}} \lim_{k\to\infty} p_k(n,0)dn \\
&= \int_{\underline{n}}^{\bar{n}} h(n,0)dn,
\end{aligned}$$

where we used the uniform convergence to swap the limits as well as limits and the integral. Therefore, we know

$$\lim_{\delta\to 0} R^{NT} = \lim_{\delta\to 0} \int_{\underline{n}}^{\bar{n}+\delta} \left[T(y(\underline{n})) + \int_{\underline{n}}^{x} \frac{1}{\frac{(f_1(n)+f_1(n-\delta))n}{1-F_1(n)-F_1(n-\delta)} \frac{\epsilon}{\epsilon+1} + 1} dn \right] (f_1(x) + f_1(x-\delta)) dx$$
$$= 2 \int_{\underline{n}}^{\bar{n}} \left[T(y(\underline{n})) + \int_{\underline{n}}^{x} \frac{1}{\frac{2f_1(n)n}{1-2F_1(n)} \frac{\epsilon}{\epsilon+1} + 1} dn \right] f_1(x) dx.$$

It follows

$$\lim_{\delta\to 0} R^T - R^{NT} = 0.$$

□

Proof of Proposition 11

Due to the system change from a class tax to the combination of class and income tax, the lower classes face the same tax payments as before. As seen in proposition 8, tax revenue increases and the lower classes therefore benefit from the reform. Thus, one can focus on the highest class in the following. Here, people face higher taxes and therefore a lower consumption, but a higher amount of the public good R. Hence, an individual of ability n profits, whenever the increase in utility out of the public good is bigger than the decrease in utility due to lower consumption. This directly leads to inequality (2.8). The integral boundaries $R^{old}, R^{new}, c^{old}, c^{new}$ result from optimal behavior and tax payments in the old and the new system, respectively. □

Chapter 3

Tax Morale, Identity and the Question of Decentralization

By Anna Kremer and Marco Wysietzki

Abstract

This paper contributes to the literature on fiscal decentralization. We argue that tax morale is important for deciding whether or not to decentralize a system and that place identity is a relevant determinant. In an empirical cross-country analysis, we confirm other literature in the sense that fiscal decentralization and national identity both have a positive impact on tax morale. We further examine the interaction of these effects and investigate the role of other levels of identity, i.e., regional and local identity. We find a significant impact of all levels, with the local identity level being particularly strong. We introduce a theoretical framework to connect our results to previous theoretical considerations. It allows us to assess whether tax morale and place identity make a decentralized or a centralized system preferable. For the example of Germany, we show that a decentralized one is preferred.

3.1 Introduction

The optimal design of tax systems is a fundamental issue in public economics with many economic implications. A key element of design is whether the tax system is centralized or decentralized. In the first case, taxes are collected and allocated by a national government, while in the second case, subnational governments collect and allocate substantial amounts of taxes within their jurisdictions.

A well-known result in the context of the debate on centralized or decentralized tax systems is the theorem of Oates (1972). It states that a decentralized system is better aligned with regional preferences in the provision of public goods, but has higher costs due to scale effects. According to Oates (1972) a decentralized system should be preferred whenever the gain from matching preferences outweighs the additional costs.

The theorem is often debated. For example, Besley and Coate (2003) argue that heterogeneous preferences cannot only be aligned with a decentralized system, but also with a centralized one. In general, it is difficult to measure preference heterogeneity and thus assess whether a centralized or a decentralized system is preferable. In this paper, we therefore refrain from focusing on preference heterogeneity, but instead focus on tax morale.

Various studies find that a decentralized tax system decreases justification of evading taxes (e.g., Güth et al., 2005; Torgler and Werner, 2005; Torgler et al., 2007; Lago-Peñas and Lago-Peñas, 2010; López-Laborda et al., 2021). Tax morale, as this justification is commonly referred to, is measured in certain surveys. It is highly correlated with actual tax compliance (Torgler, 2004) and therefore a variable of interest. One might expect the impact of decentralization on tax morale to depend on individuals' place identity, which is an umbrella term for local, regional and national identity. People with strong identity might benefit more from a decentralized system and thus increase their tax morale. This could especially be the case, if they feel connected to and responsible for their local or regional communities. They might appreciate the better met needs in a decentralized system (Tiebout, 1956) and increase their tax morale. This could be the case independent of their national identity. We therefore include place identity in our analysis. Existing studies show that an individual's strong national identity or being more patriotic is related to higher tax morale (Konrad and Qari, 2012, Andreoni, 1990, MacGregor and Wilkinson, 2012). We add to the literature by providing a closer look at regional and local identity.

To formalize these effects, we introduce a stylized model that builds on feelings of morale concerning tax evasion depending on the fiscal system and the degree of place identity. It allows us to analyze how tax evasion and welfare differ between a decentralized and a centralized system. We connect this analysis to the theorem of Oates (1972) and identify parameter ranges for which the theorem would suggest a centralized system, while welfare would actually be higher in a decentralized system given the effects on tax compliance.

In a next step, we empirically investigate the impact of decentralization and place identity on tax morale. We quantify the effects using survey data from the European Value Study (EVS), the World Value Survey (WVS) and fiscal data (from OECD and IMF). We confirm the results of other studies concerning the positive effect of decentralization on tax morale. Moreover, we find that national, local and regional identity have a positive effect. We are the first to measure the impact of regional and local identity in this context. The effect of local identity seems to be the largest numerically, although we cannot reject the hypothesis that all identity levels have the same influence. The coefficient on the interaction of local identity and decentralization is negative, meaning that decentralization has a higher positive impact on people with lower local identity. We argue that this can be explained, among others, by scale effects.

To better understand size and impact of our results, we finally turn to the example of Germany. We calibrate our model and check if the existing decentralized system in Germany is preferred over a centralized one. Under reasonable assumptions concerning administrative costs, the decentralized system yields higher welfare compared to a centralized system.

The remainder of this paper is organized as follows: In section 3.2, we present a broad overview of the related literature. Section 3.3 introduces a model that links tax morale, identity and the theorem of Oates. In section 3.4, we first describe the used data and our empirical strategy, before we summarize our results and discuss an additional IV approach. We calibrate our model in section 3.5 and we conclude in section 3.6.

3.2 Literature Overview

Our paper contributes to two strands of literature. First, it relates to the literature that seeks to find out about determinants of tax morale. Here, we concentrate on the impact of a tax system being centralized or decentralized and on the individual's place identity. Second, we contribute to the literature on the question of whether a centralized or decentralized tax system is preferable.

Concerning the first strand, Güth et al. (2005) show in an experimental study that a centralized tax system induces a lower tax morale. Likewise, Torgler and Werner (2005) confirm this effect for German municipalities, while controlling for a detection rate and socio-economic factors. In a very recent study, López-Laborda et al. (2021) confirm this connection for the example of Spain. As an additional factor in a decentralized tax system, Lago-Peñas and Lago-Peñas (2010) show that rich European federal regions have a lower tax morale, which they attribute to inter-regional transfers. In the same context, Torgler et al. (2007) study Switzerland in the 1990s and find that cantons with a higher direct democratic participation or local autonomy have a higher tax morale. In this paper, we confirm the positive relation between decentralized tax systems and tax morale. Furthermore, we connect this to the impact of place identity on tax morale and check if there is an interaction.

Various studies show that the decision whether or not to pay taxes depends on various personal psychological aspects. These include personal ethics or social norms (e.g., Grasmick and Bursik (1990), Bosco and Mittone, 1997; Wenzel, 2004) or perception of fairness (e.g., Murphy, 2003). A more extensive discussion is found in Wenzel (2007). One of these factors is national identity. Its positive impact on tax morale has been proven in multiple previous studies: Torgler and Schneider (2004) show this relation for different European countries and Torgler (2004) looks at Asia. Torgler (2005) focuses on Latin America, while Martinez-Vazquez and Torgler (2009) provide evidence by the example of Spain. All studies check the impact of multiple variables on tax morale, including national pride. They employ different proxies from survey data, such as the EVS and the WVS, and find a link between higher pride in national origin and higher tax morale.

Konrad and Qari (2012) evaluate the impact of national pride on tax compliance. They refer to patriotism instead of nationalism because the latter implies the rejection of other groups, e.g. migrants. Using ISSP data, they find a robust positive association between patriotism and tax morale for eight mostly Western countries. Furthermore, they try to resolve simultaneity through an IV approach with election participation and membership in a sports club as instruments. They find the effect to be causal. They admit that their instruments are not very strong, however.

MacGregor and Wilkinson (2012) are also specifically interested in patriotism and criticize the literature's focus on people's acceptance of tax evasion. They examine whether taxpayers' views are influenced by patriotism in a broader sense. Using survey data from an Income Tax Assistance program, they find a significant association between patriotism and taxpayer attitudes. Among other aspects, patriotic individuals are more likely to support higher taxes and a progressive tax system.

All previous studies analyzing tax morale and identity have examined patriotism or identity at national level. In this paper, we confirm this relationship. Besides, we add the consideration of local and regional identity. Using an experimental approach, Gangl et al. (2016) show a relation between local identity and tax morale. Their reasoning corresponds with our assumption that the identity of individuals often differs between local, regional and national levels. We aim to contribute to their findings and can, opposed to Gangl et al. (2016), build on a data set that directly encompasses local and regional identity.

As a second strand, we contribute to the literature on the question if centralized or decentralized tax systems should be preferred. It is theoretically investigated in the seminal papers of Tiebout (1956) and Oates (1972). Tiebout (1956) argues that local spending could better reflect local preferences than national spending. Oates (1972) on the other hand places greater emphasis on the more efficient provision of public goods in a centralized system due to scale effects.

We link this topic to our previously outlined considerations, specifically we integrate tax morale into the assessment of a decentralized tax system. We are the first to check whether the level of place identity acts as a mediator for tax morale depending on the decentralization of the tax system. If this is the case, it would provide new arguments for the question of whether a centralized or decentralized system is preferable. We therefore connect our analysis to the theorem of Oates (1972).

3.3 Theoretical Considerations

In this section we formalize our previous considerations and introduce a simple model based on Besley (2020). The strength of our model is the focus on the channels between our variables of interest and the refraining from other standard explanations in the literature. This allows us to calibrate our model in section 3.5.

Our model enables us to relate to the theorem of Oates (1972). It formulates a trade-off regarding decentralization: A decentralized system can better fulfill the needs of a population, but is associated with higher costs due to scale effects. We aim to specify under which conditions identity and tax morale impact the decision in favor of a decentralized or a centralized system. Furthermore, we investigate how this impact changes Oates' theorem.

We consider a country populated by a continuum of individuals of measure 1. The government may choose to tax and provide public goods either centrally or decentrally, denoted by the index C and D, respectively. Following the considerations of Oates (1972), the latter improves the quality of public goods, but entails additional costs.

Each individual gets utility from consumption c and a public good G. To finance both, individuals have an endowment e and are expected to pay taxes at a rate t. However, each individual can decide whether to evade a share n of the intended taxes. Consumption is thus given by $c = e - e(1 - n)t$. One could decide to fully freeride, choose $n = 1$ and consume the full endowment. Similar to standard models for tax evasion (e.g. Allingham and Sandmo (1972)), detected evasion implies a penalty. Furthermore, moral feelings shape the individuals decisions. Both leads to a choice of $n \in [0, 1]$. In a system j, with $j \in \{C, D\}$, utility is therefore given by

$$U^j = G^j + e - e(1-n)t - eH(n) - \mu_1\gamma n - \mu_2\gamma n\mathbb{1}^D - \mu_3 n\mathbb{1}^D - \sigma n. \qquad (3.1)$$

First, note that the public good G^j differs for a centralized and a decentralized system. This takes into account that a decentralized system can offer public goods in a region-specific way, yielding a higher quality and, thus, a higher utility for people in these regions. At the same time, the decentralized system is associated with higher costs δ^D. It is therefore $G^C = T^C - \delta^C$ and $G^D = \alpha(T^D - \delta^D)$, with $\alpha > 1$, $\delta^D > \delta^C$ and $T^j = et(1-n)$ describing tax revenue.

Evading taxes may be costly if evasion is detected. The penalty due is given by the function $H(n)$, and scaled by endowment e. Following Besley (2020), we use $H(n) = \frac{n^2}{2}$ in the rest of the analysis.

The next part of equation (3.1) is about regional identity γ. The stronger the identity, the worse an individual feels if she evades taxes. This is especially true if collected taxes are used in a region-specific way. Thus, we subtract γn and $\gamma n\mathbb{1}^D$, scaled by μ_1 and μ_2, respectively. The function $\mathbb{1}^D$ equals one if the system is decentralized and zero otherwise.

In a centralized system, tax revenue is spent on the whole country and not only on the taxpayer's home region. In a decentralized system, the individual can be sure to get back a good portion of her taxes by the public goods (Torgler, 2004). Thus, she feels comparatively worse for evading taxes in a decentralized system. This is captured by $n\mathbb{1}^D$ and scaled by μ_3. With σn, the very last part of equation (3.1) sums up all the remaining potential impact of tax evasion on utility. This is necessary for connecting our theoretical and empirical results in section 3.5.

Utility maximization yields

$$n^j = \frac{et - \mu_1\gamma - \mu_2\gamma\mathbb{1}^D - \mu_3\mathbb{1}^D - \sigma}{e}. \qquad (3.2)$$

To ensure an inner solution, we assume $e > et - \mu_1\gamma - \mu_2\gamma\mathbb{1}^D - \mu_3\mathbb{1}^D - \sigma > 0$. A higher regional identity reduces evasion as $\frac{\partial n^j}{\partial \gamma} = -\frac{\mu_1 + \mu_2\mathbb{1}^D}{e} < 0$ for a centralized and a decentralized system, respectively. However, the latter further amplifies this effect.

Due to their impact on tax morale, the tax system and regional identity influence tax revenue and therefore the public good. It is

$$G^C = t(e - et + \mu_1\gamma + \sigma) - \delta^C, \tag{3.3}$$
$$G^D = \alpha t(e - et + (\mu_1 + \mu_2)\gamma + \mu_3 + \sigma) - \alpha\delta^D. \tag{3.4}$$

To assess whether a decentralized system is preferred by individuals, we do not compare the public goods on offer, but rather the utility in both situations. This results in the following Proposition:

Proposition 12. *A decentralized tax system is preferred by individuals if the additional costs $\Delta\delta$ satisfy*

$$\Delta\delta < \bar{A} \equiv A_1 + A_2 + A_3, \tag{3.5}$$

with $A_1 = (\alpha - 1)t(e - et + (\mu_1 + \mu_2)\gamma + \mu_3 + \sigma)$, $A_2 = \frac{(\mu_2\gamma+\mu_3)(et-\mu_1\gamma-\sigma)}{e} - \frac{(\mu_2\gamma+\mu_3)^2}{2e}$, $A_3 = \frac{(\mu_1\gamma+\sigma)(\mu_2\gamma+\mu_3)}{e} - \frac{\mu_2\gamma+\mu_3}{e}(et - (\mu_1 + \mu_2)\gamma - \mu_3 - \sigma)$ *and* $\Delta\delta = \alpha\delta^D - \delta^C$.

Proposition 12 summarizes three main effects of a system change. In A_1 the higher quality of the public good in a decentralized system is captured. Term A_2 corresponds to the decrease in the expected penalty for detected tax evasion. It decreases due to the lower tax evasion n. The change in moral feelings is captured by A_3 and consists of two parts. The first part corresponds to the decrease in moral disutility due to lower tax evasion in a decentralized system. While the second part is the additional disutility that occurs only in a decentralized system. Whenever the sum of these effects is greater than the increase in costs, a decentralized system is preferred.

In summary, a decentralized system exhibits several advantages due to the induced decrease in tax evasion. However, it entails additional financial costs and additional moral costs for individuals, as they feel particularly bad about not contributing to a decentralized system. Oates (1972) claims that the decision in favor of a centralized or a decentralized system is (under certain assumptions) only about the better public good versus the higher costs. Here, we argue that it is also about lower tax evasion induced by individuals' moral feelings. Thus, under certain parameters Oates (1972) would have proposed a centralized system, as the additional costs are higher than the public good benefits, while our model proposes a decentralized system. This is summarized in the following Proposition:

Proposition 13. *A decentralized tax system is preferred although the direct improvement of the public good is outweighed by the higher costs, whenever* $\Delta\delta \in [(\alpha - 1)t(e - te + \mu_1\gamma + \sigma), \bar{A}]$*, with* $\bar{A}$ *defined as in Proposition 12.*

In the following section, we turn to our empirical analysis. In particular, we examine whether our hypothesis concerning identity and decentralization and their impact on tax morale can be confirmed. In terms of our model, this means we attempt to quantify μ_1, μ_2, μ_3 and σ.

3.4 Empirical Analysis

3.4.1 Data and Empirical Strategy

For the empirical part, we rely on various publicly available data sets. We use individual level data on tax morale, local, regional and national identity, as well as individual socioeconomic control variables[1] and information on town size from the European Value Study (EVS) and the World Value Survey (WVS). At country level, we merge data from the International Monetary Fund (IMF) and the Organisation for Economic Co-operation and Development (OECD) on the decentralization of the tax system. The latter two data sets contain information on revenues and expenditures at local, state and central level as indicators for decentralization. We therefore focus on fiscal decentralization without further enquiring about the decentralization of the political system. As additional country level controls, we obtain information on total expenditure to measure the size of governments, GDP, and population size from the OECD. We also monitor government effectiveness and corruption control in countries with data from the World Bank's Worldwide Government Indicators.

The surveys contain the question "Please tell me [...] whether you think it can always be justified, never be justified, or something in between: Cheating on tax if you have the chance" on a scale between 1 and 10, where 1 represents never and 10 always.[2] This assessment of tax evasion behavior is referred to in the literature as 'tax morale'. It is often used as a synonym for tax compliance, although it does not measure actual action. Torgler (2004) presents a high correlation between these terms using experiments. We apply this result to establish a link between tax morale and tax compliance[3] in section 3.5.

Local, regional and national identity is reflected in the answers to the question "People have different views about themselves and how they relate to the world. [...] Would

[1]Income level measured in 3 levels, employment status in 8 categories, sector (primary, secondary, tertiary), gender (male, female), age, education level in 3 levels, religion in 10 categories, marital status in 6 categories.

[2]The distribution is heavily right-skewed since most people refrain from (admitting) to justify tax evasion.

[3]For a further discussion of these terms, see for example Luttmer and Singhal (2014).

you tell me how close do you feel to...? your town or city; your [county, region, district]; [country]" on a four-level-scale representing Very Close (1), Close (2), Not Very Close (3) and Not Close at All (4). The data used provide this distinction only for the most recent wave of surveys. Therefore, our analysis below relies on a cross-country comparison rather than on a panel.[4] We distinguish the levels of identity because the attachment of individuals to their town, their region and their country can vary. Accordingly, tax morale can also be influenced differently. A more local level of identity can be particularly important for tax morale in conjunction with decentralization.

To classify decentralization we create a dummy[5] (analogous to Belmonte et al. (2018)) based on the fiscal structure. If two of the following three criteria apply, a country is considered decentralized: Subnational expenditure share is greater than the sample mean, subnational revenue share is greater than the sample mean, and inter-governmental transfer revenue share relative to expenditure of the subnational levels is greater than the sample mean.

For the empirical method we adhere to the majority of the literature and estimate a Weighted Ordered Probit model (compare Torgler et al., 2007, Konrad and Qari, 2012, Lago-Peñas and Lago-Peñas, 2010), summarized by the following equation:

$$Morale_{i,c} = \beta_1 Ident_i + \beta_2 Ident_i \times Decentral_c + \beta_3 Decentral_c + x_i'\beta_4 + y_c'\beta_5 + \epsilon_{i,j} \tag{3.6}$$

This models the tax morale ($Morale_{i,c}$) of individual i in country c as a function of the individual's identity ($Ident_i$), a dummy $Decentral_c$ for a country's degree of fiscal decentralization, and their interaction. All other controls are summarized in the vectors x_i and y_c.

Descriptive statistics of the included variables can be found in Table 3.1. Many of the variables are categorical but their ordinal distribution is still meaningful and they are therefore depicted with a reference category to understand the magnitudes.[6] Our data set contains information on individuals from the following countries: Turkey, Greece, Czech Republic, Hungary, Slovenia, Slovak Republic, Estonia, Mexico, New Zealand, Poland, Chile, Lithuania (all considered as centralized), Spain, Norway, Sweden, Finland, France, Italy, Austria, Iceland, Denmark, Japan, Colombia, Switzerland, Australia, Netherlands, Germany, United States, Canada (all decentralized).

[4] For a further analysis, we extend the data in time by using a proxy for national identity (see subsection 3.4.2)

[5] In the appendix we offer a robustness check concerning alternative decentralization measures.

[6] Means of marital and employment status, the sector of employment and religious denominations are missing because these categorical variables do not have a natural order and thus cannot be summarized by averages.

	count	mean	sd	min	max	reference category of mean (if needed)
Tax morale	27509	9.053328	1.805554	1	10	10: never justifiable
Local identity	27509	3.262387	.7405576	1	4	3: feeling close
Regional identity	27509	3.147697	.760724	1	4	3: feeling close
National identity	27509	3.354611	.7067408	1	4	3: feeling close
Decentralization dummy	27509	.7315061	.443184	0	1	1: decentral
Government size	27509	38.33156	16.95806	0	55.6487	
Employment status	27509					not ordinal
Employment sector	27509					not ordinal
Income level	27509	1.95656	.702441	1	3	2: medium
Gender	27509	1.529354	.4991467	1	2	2: female
Age	27509	49.85746	17.24994	17	98	
Marital status	27509				not ordinal	
Education level	27509	2.251918	.7121837	1	3	2: middle
Religion	27509					not ordinal
Townsize	27509	4.415828	2.407337	1	8	4: 10,000-20,000 (EVS) /10,000-25,000 (WVS)
GDP	27509	46854.94	14007.77	15439.04	69807.23	
Population	27509	5.31e+07	7.66e+07	338349	3.24e+08	
Government effectiveness	27509	1.31663	.6254288	-.2636025	2.050234	-2.5: less effective; 2.5: more effective
Corruption control	27509	1.323498	.8398482	-.9295135	2.232307	-2.5: weak; 2.5: strong

Table 3.1: Descriptive Statistics of Variables, Except for Marital and Employment Status, Sector and Religion

3.4.2 Results

Baseline results

Estimating a Weighted Ordered Probit model we obtain the results shown in Table 3.2. The following should be noted for the interpretation: Tax morale is numerically higher when people have a lower justification for tax evasion. Higher identity numbers refer to a stronger identity. The decentralization dummy equals 1 if a country's fiscal structure is relatively decentralized.

We exhibit different specifications of the regression[7] in the different columns of Table 3.2: In the first, we simultaneously control for all levels of place identity, while in the others, we consider only local, regional or national identity, respectively. When only one level of identity is considered, the impact of this level on tax morale is positive and highly significant. When all levels are simultaneously controlled for, the statistically significant influence of local identity continues to apply, while regional and national identity have only a positive yet insignificant effect. This suggests that any feeling of identity increases tax morale, but that local identity may be the most important and robust one. However, Wald tests show that we cannot reject the hypothesis, that the coefficients are equal, either jointly or pairwise. To illustrate this, the corresponding confidence intervals are shown in Figure 3.1.

Our estimations demonstrate a positive and significant correlation between decentralization and tax morale in each specification, which is in line with the literature. Thus, in decentralized places citizens consider cheating on taxes to be less acceptable.

We find negative effects for the interaction of identity and decentralization. This is true regardless of which level of identity we consider. If we only include local or regional identity, this negative effect is significant even at the 5% level. Thus, moving from a centralized to a decentralized system has a positive effect on tax morale, whereas this effect is smaller for individuals with relatively strong place identity. This result is not intuitive at first glance. One could expect individuals with strong place identity to care even more about their region in a decentralized system, leading to a high tax morale. As a result, the interaction could actually be positive. We further discuss this and provide possible explanations in the following.

[7]The corresponding overall marginal effects are indicated in Table 3.8 in the Appendix.

	(1) Tax morale	(2) Tax morale	(3) Tax morale	(4) Tax morale
Local identity	0.129***	0.184***		
	(0.0463)	(0.0368)		
Regional identity	0.0489		0.145***	
	(0.0446)		(0.0345)	
National identity	0.0494			0.131***
	(0.0423)			(0.0364)
Decentralization dummy	0.802***	0.729***	0.686***	0.614***
	(0.193)	(0.155)	(0.145)	(0.164)
Decentralization*Local identity	-0.0498	-0.109**		
	(0.0575)	(0.0454)		
Decentralization*Regional identity	-0.0765		-0.105**	
	(0.0581)		(0.0440)	
Decentralization*National identity	-0.00884			-0.0752
	(0.0537)			(0.0470)
controls	yes	yes	yes	yes
N	27521	27595	27565	27565

Standard errors in parentheses

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

Table 3.2: Baseline Regression Results

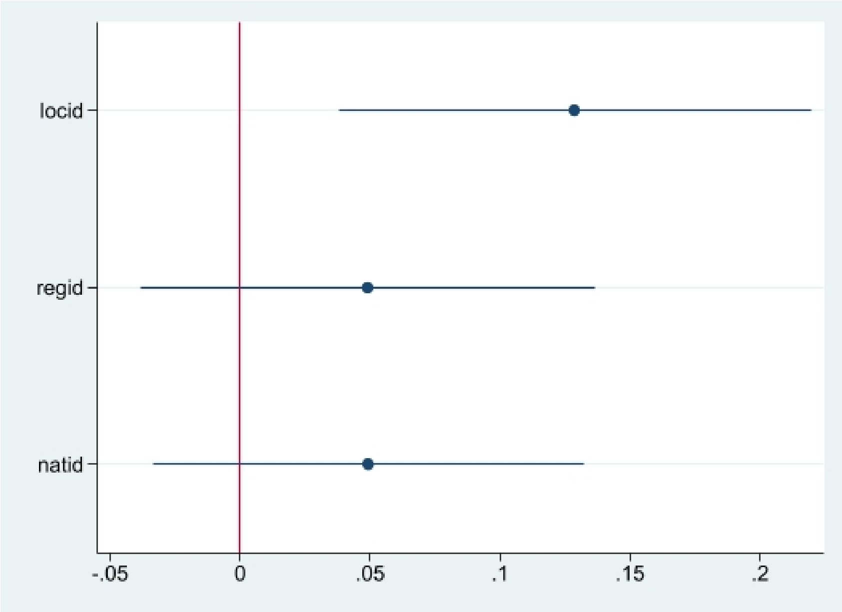


Figure 3.1: Confidence Intervals for Regression Estimates: Local, Regional and National Identity

Discussion of the Negative Interaction Term

In our model in section 3.3, the interaction of decentralization and identity is captured by μ_2. We have determined how μ_2 affects tax evasion and therefore the decision whether or not to decentralize a system in this framework. Moreover, we have revealed that it has consequences for the theorem of Oates (1972). A negative μ_2 mitigates these consequences. Therefore, it is of particular interest to understand why the interaction is negative and whether there are reasons to doubt this result. We discuss three possible explanations below.

A first explanation are scale effects. The general level of tax morale is quite high with an average value of 9.02 out of 10. The values are higher for those with strong identity. Thus, it is evident that a change from a centralized to a decentralized system cannot simply increase tax morale among those with strong identity as much as those with weak identity. Similarly, in a decentralized system, tax morale could be so high that identity has no additional increasing effect. To show more formally that this scale effect for identity[8] exists, we extend our data set with several waves over time. This has the disadvantage of losing data on the different levels of identity, since these were only collected in the last survey wave. However, most waves include a question on national pride. We use this question as a proxy for national identity.[9] Since individuals are not

[8]It is not possible to perform a similar strategy for decentralization because there is only minor variation in these countries' fiscal system.

[9]Gangl et al. (2016) do the same. A discussion of the difference between nationalism and patriotism

	(1) Change in Tax morale
Tax morale (initial)	-0.0781
	(0.152)
Change in national pride	-1.141
	(1.672)
Decentralization dummy	11.04
	(7.820)
0.Decentralization*National pride	3.873
	(2.744)
1.Decentralization*National pride	0.721
	(0.731)
N	25
R^2	0.720

Standard errors in parentheses. Standard errors clustered at the country level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.3: Effect of Original Tax Morale Level on its Change

tracked separately in the surveys, we cannot rely on individual panel data. Therefore, we aggregate the data at country level to form a country panel enabling us to look at the change in average tax morale in a country relative to the change in identity. We run a regression similar to our baseline regression, but at country level and controlling for changes in national pride and several other variables.[10] The results are summarized in Table 3.3.[11] In this regression, the initial level of tax morale negatively impacts the change in tax morale, corroborating the hypothesis of scale effects. Higher levels of tax morale do not change as much as lower levels when a tax system is transformed into a decentralized one.

A second explanation is a substitution effect between identity and decentralization as reflected in the negative interaction effect. Individuals who identify less with the different levels may care more about the quality of the public good, which Oates (1972) contends is better in a decentralized system. Therefore, they are more willing to pay

can be found in Konrad and Qari (2012).

[10]In order to control for several non-ordinal variables, we need to update our data set: For marital status at country level we use share of people married or living as if married, for religious denomination we consider the percentage of people being part of a religious organization and for employment status we control for the share of unemployment in the country sample. Owing to the data structure, we do not control for sectors here.

[11]The findings are based on a subsample of only 25 country-year observations, but regression diagnostics show that the analysis is appropriate.

their taxes if they receive better value in return. Whereas individuals who identify more strongly might derive higher utility from legitimately paying their taxes, regardless of the quality of the public good. To prove this, we would need a variable in the individual dataset that captures the perceived utility of public goods to the individual. However, there is no real suitable proxy for this variable in the surveys.

A third explanation for the negative interaction is related to the literature on yardstick competition (following Besley and Case, 1992). Individuals in a decentralized country might compare their tax rates and the quality of public goods with those of their neighbors. If they see that other regions are doing better, their willingness to pay taxes decreases. This problem occurs only in the decentralized case, and could be more pronounced among individuals with stronger identity, since they might be more interested in politics and therefore compare their region more with others. We try to control for this effect by adjusting the regression with a variable capturing political interest. If yardstick competition effect is responsible for the negative interaction, it should be weakened by this introduction.[12] In the results in Table 3.4, estimated with the individual data set, the interaction effect is slightly attenuated compared to our baseline estimate, but to such a small extent that it appears negligible. Since the proxy of political interest for determining jealousy of other regions appears far from perfect, we cannot reject the hypothesis. Yardstick competition could still have an effect, but we do not find clear evidence for it.

3.4.3 Causality of the Correlation

While we have already demonstrated the positive correlation between place identities and tax morale, the direction of causality in our regressions remains unclear. Higher tax morale could affect identity levels, which would differ from the mechanism we propose. In this section, we attempt to confirm causality. To do so, we use a well-established instrument: The ruggedness of a country. Data on ruggedness of the countries stem from Shaver et al. (2019).

This instrument is based on the finding that geographic profile correlates with people's identity.[13] The reasoning is as follows: In a rugged country, transportation and communication is more difficult and therefore more valuable. For that reason, those who live in a rugged place tend to think of themselves more as a community. Meanwhile, ruggedness should be an external variable that does not directly affect tax morale

[12]Political interest is only correlated between 2 to 8% with the identity levels and to 16% with decentralization.

[13]The mean territorial ruggedness of a country correlates between 8 and 15 % with the identity at the different levels.

	(1) Tax morale	(2) Tax morale
Decentralization dummy	0.683***	0.648***
	(0.174)	(0.168)
Political interest	-0.00335	0.00523
	(0.0309)	(0.0310)
0.Decentralization*Political interest	0	0
	(.)	(.)
1.Decentralization*Political interest	0.0132	0.00643
	(0.0379)	(0.0380)
Local identity	0.184***	
	(0.0369)	
0.Decentralization*Local identity	0	
	(.)	
1.Decentralization*Local identity	-0.107**	
	(0.0455)	
Regional identity		0.142***
		(0.0347)
0.Decentralization*Regional identity		0
		(.)
1.Decentralization*Regional identity		-0.101**
		(0.0441)
N	27516	27487
R^2		

Standard errors in parentheses

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

Table 3.4: Controlling for the Effect of Political Interest/Yardstick Competition

because there is no link between these two concepts.

Using ruggedness as an instrument, we estimate a Weighted Ordered Probit IV regression. This estimation also uses Ordered Probit in the first stages and controls for the same control variables as in the baseline model. The second- and first-stage results are reported in Table 3.5. The second-stage regressions confirm our basic results regarding the direction of effects, except for the effect of decentralization in the national specification. They confirm the positive effect of identity and (mostly) decentralization on tax morale. However, in the second-stage IV regressions, the effects are no longer significant.

First-stage regressions show that local identity is positively affected by decentralization and by the mean value of ruggedness separately. However, the interaction of ruggedness and decentralization seems to have a negative impact on local identity. This could be due to a scale effect similar to the one discussed earlier. Looking at the first-stage regression for local identity exclusively in decentralized countries confirms that ruggedness has a negative effect. The first-stage results at the regional level are similar when considering the determinants of regional identity. They differ only with respect to regional identity in a decentralized country. Here, a positive constant is established for decentralized countries, whereas a negative effect of the interaction cancels out some of the positive effect of ruggedness. Looking at the first-stage regression for national level, we determine that all variables are positively related to national identity (and the interaction with decentralization), but all are insignificant.

Taken together, the IV confirms the direction of our baseline results and thus shows that at least part of the causality runs from the notion of identity to tax morale. However, due to insignificance, it is uncertain whether this direction describes the main effect. It seems likely that higher tax morale also affects the feeling of place identity. One explanation could be that individuals justify their tax morale by overreporting the level of identity to rationalize their low levels of tax evasion, as explained by Konrad and Qari (2012).

	(1)	(2)	(3)
2nd stage: Tax morale			
Decentralization dummy	0.0586	0.0103	-0.0591
	(.)	(7.301)	(0.364)
Local identity	0.166		
	(0.263)		
1.Decentralization*Local identity	-0.0988		

	(0.203)		
Regional identity		0.123	
		(0.583)	
1.Decentralization*Regional identity		-0.0674	
		(0.300)	
National identity			0.113
			(0.300)
1.Decentralization*National identity			-0.0567
			(0.0993)
1st stage: Local identity			
Decentralization dummy	1.266***		
	(0.154)		
Ruggedness	0.00410***		
	(0.000610)		
Ruggedness*1.Decentralization	-0.00263***		
	(0.000595)		
1st stage: Local identity*1.Decentralization			
Decentralization dummy	0		
	(.)		
Ruggedness	-0.00291***		
	(0.000218)		
Ruggedness*1.Decentralization	0		
	(.)		
1st stage: Regional identity			
Decentralization dummy		1.394***	
		(0.299)	
Ruggedness		0.00531***	
		(0.00125)	
Ruggedness*1.Decentralization		-0.00313***	
		(0.000771)	

1st stage: Regional identity*1.Decentralization			
Decentralization dummy		8.854***	
		(2.954)	
Ruggedness		0.00394	
		(.)	
Ruggedness*1.Decentralization		-0.00158***	
		(0.000511)	
1st stage: National identity			
Decentralization dummy			0.250
			(0.452)
Ruggedness			0.00119
			(0.00181)
Ruggedness*1.Decentralization			0.000231
			(0.000746)
1st stage: National identity*1.Decentralization			
Decentralization dummy			8.102
			(7.843)
Ruggedness			0.00110
			(.)
Ruggedness*1.Decentralization			0.000392
			(0.00137)
N	27738	27738	27738

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.5: IV Regression

3.5 Decentralization of Germany

In this section, we calibrate our model and use the empirical results to describe in more detail under which conditions a decentralized system yields welfare gains compared to a centralized system. We conduct this analysis for the example of Germany. Since the

German system is already decentralized, the question in this case is whether switching to a centralized system would lead to welfare losses.

Summarizing the previous results, we have seen that decentralization, place identity and the interaction of both, impact on tax morale. This leads to a higher quantity of public goods and has direct utility effects in our model. Decentralization can thus lead to higher welfare. However, it is also associated with costs. As discussed in the previous sections, Oates (1972) explores a trade-off between better aligning individual preferences and scale effects concerning a system change. In a decentralized system, public goods can be provided in a more region-specific way. However, in a centralized system scale effects lead to, for example, lower administrative costs or fewer externalities. In this section we abstain from preference heterogeneity and define a range of scale effects in which a decentralized system yields welfare gains. Thus, we define situations in which a decentralized system could be beneficial even without taking regional preferences into account.

Table 3.6: Calibrated Parameters for Germany

Parameter	Meaning	Value	Source/Target
e	individual income	4401 €	Destatis (2022b)
t	total tax rate	39%	Destatis (2022b)
γ	local identity	3.3	panel data
μ_1	impact of identity	$242,94$	Table 3.7; $\mu_1 = \beta_1 e$
μ_2	impact of interaction	$-192,32$	Table 3.7; $\mu_2 = \beta_2 e$
μ_3	impact of decentralization	$919,81$	Table 3.7; $\mu_3 = \beta_3 e$
σ	additional impact	383,08	calculation to meet n
n^D	evasion in dec. system	0.056	panel data
δ^D	costs in dec. system	114 €	Destatis (2022a)

To calibrate our model for Germany, we use our empirical results as well as some exogenous sources. All values can be found in Table 3.6. The average pre-tax monthly income before taxes for a single German is 4401 €. Net earnings are 61% of this value, which gives $t = 39\%$. From the EVS data set, we know that local identity in Germany has the value $\gamma = 3.3$.

For the estimates of identity and decentralization we use a simple OLS regression. The results can be found in Table 3.7. These estimates can be interpreted quantitatively, opposed to those from the Ordered Probit model, which is our baseline estimation.[14]

[14] We do not use the marginal effects of the Probit model because we have a multinomial outcome variable with 10 levels and interpretation of the estimates depends on holding the other variables constant. It is therefore simpler to use the OLS because it does not measure the effect on the probability on a given level of tax morale, but estimates the effect on tax morale of increasing the independent variables by one unit.

	(1) ln Tax morale	(2) ln Tax morale	(3) ln Tax morale	(4) ln Tax morale
Local identity	0.0411*	0.0552***		
	(0.0239)	(0.0181)		
Regional identity	0.00869		0.0406**	
	(0.0232)		(0.0164)	
National identity	0.0172			0.0398**
	(0.0207)			(0.0174)
1.Decentralization dummy	0.256***	0.209***	0.187***	0.186***
	(0.0816)	(0.0679)	(0.0603)	(0.0683)
0.Decentralization*Local id	0	0		
	(.)	(.)		
1.Decentralization*Local id	-0.0233	-0.0437**		
	(0.0259)	(0.0196)		
0.Decentralization*Regional id	0		0	
	(.)		(.)	
1.Decentralization*Regional id	-0.0203		-0.0398**	
	(0.0256)		(0.0181)	
0.Decentralization*National id	0			0
	(.)			(.)
1.Decentralization*National id	-0.0152			-0.0373*
	(0.0227)			(0.0195)
N	27489	27563	27533	27533
R^2	0.078	0.077	0.076	0.076

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.7: OLS Estimation for Calibration

Although regression diagnostics show that OLS regression is not well placed to reproduce our model, the results are qualitatively comparable to those of Probit regression. Based on the latter, we believe the estimates are reasonably close to the true values.

In this calibration, we focus on column (2) of Table 3.7, which is about local identity. For an appropriate matching of the empirical results and our model, equations (3.6) and (3.2) are crucial. We rearrange the latter to

$$1 - n^j = \frac{e - et + \mu_1\gamma + \mu_2\gamma\mathbb{1}^D + \mu_3\mathbb{1}^D + \sigma}{e}. \tag{3.7}$$

In this format, both equations show the impact of the variables of interest on tax morale and the share of taxes that are not evaded, respectively. To match these equations, we assume that the share $1 - n$ and tax morale are equal. We can then compare the prefactors of the variables of interest. It is $\beta_i = \frac{\mu_i}{e}$ for all $i \in \{1, 2, 3\}$. Multiplying the values of column (2) in Table 3.7 with e yields the values in Table 3.6.

To ensure that the equations match, σ is a variable that captures all other effects on $1 - n$ except for our variables of interest. From the panel data we know, that the average tax morale in Germany has the value 0.944. This yields $n^D = 0.056$. Equation (3.7) allows us to identify $\sigma = 383.08$.

With all the parameters in mind, we can calculate tax evasion in Germany if it was a centralized system. It yields $n^C = 0.1208$ which is higher than in the decentralized case. Substituting all values into the utility function (3.1), we get the change in utility

$$\Delta U = U^D - U^C \tag{3.8}$$
$$= \delta^C - \delta^D + 86.02. \tag{3.9}$$

It depends on the relation of costs in a centralized and in a decentralized system, whether the latter yields higher utility,. A change in the German system from a decentralized to a centralized system would have to be associated with lower costs of at least 86.02 € per person to achieve a benefit. As discussed earlier, we expect costs to be higher in a decentralized system due to scale effects. However, these scale effects are difficult to measure. We use administrative costs to assess the meaning of these 86.02 €. It is difficult to accurately describe administrative costs, too, but at least a lower and an upper bound can be given.

Considering only costs that are definitely incurred for administrative purposes[15] yield 114 € per worker (Destatis (2022c), Destatis (2022a)). Using instead simply all governmental expenditures yields 1433 € (Destatis, 2023). Based on these numbers, the

[15]Specifically, we add up "Politische Führung und zentrale Verwaltung" and "Finanzverwaltung".

introduction of a centralized system would need to be associated with a cost reduction of between 6% and 75% in order to be beneficial.

Other literature attempts to determine the additional costs of decentralization by looking at the size of government (Feld et al., 2003). In this context, the literature is inconclusive. Oates (1972) is unable to find a robust correlation between centralization and size of government, measured by tax revenue (compare Feld et al., 2003). In a more recent paper, Cassette and Paty (2010) examine a panel of European countries and find that greater tax autonomy does in fact increase total public expenditure. With respect to Germany, there are several studies relying on difference-in-difference or synthetic control designs that look at administration costs. Yet, they focus on the local level. Roesel (2017) examines district mergers in Germany. He considers the example of Saxony and shows that there is no significant effect of centralization on expenditure. This is consistent with the finding of Blesse and Baskaran (2016) for Brandenburg. Fritz (2015) studies the state of Baden-Wuerttemberg and concludes that municipal mergers actually increase costs.

In summary, it is not even clear that centralization in Germany would be associated with a cost reduction of 6%, which is the lower bound discussed above. It can be concluded that the current decentralized system is preferable to a centralized system because of its impact on tax morale and its direct impact on utility. This is especially true if the heterogeneity of preferences across regions would be additionally taken into account.

3.6 Conclusion

This paper revisits the debate about centralized or decentralized fiscal systems. The famous theorem of Oates places the focus on preference heterogeneity. We argue that tax morale is important in this context as well. Thus, we examine the issue of decentralization or centralization by linking it to tax morale and its relationship to place identity.

We confirm that national identity and decentralization improve tax morale. We complement the literature by focusing on regional and local identity, and determine that they positively impact on tax morale as well, with local identity seemingly having the largest impact. However, we cannot reject the hypothesis that all effects are of equal magnitude. Using an instrumental variable estimation allows us to examine the causal relationship between identity and tax morale to some extent.

We introduce a model to connect our considerations to the theorem of Oates. We first theoretically demonstrate how tax morale can change the content of the theorem.

We then calibrate the model to Germany using our empirical results. We find that centralization in Germany would have to be associated with a substantial cost reduction to outweigh the negative impact on tax morale.

Future research could increase the precision of this analysis and especially improve the model as it is very stylized. In particular, other forms of the utility function could be taken into account. In addition, it would be very interesting to include heterogeneity of preferences as well as externalities in the evaluation. Finally, we reserve a more detailed look at the reasons for the negative interaction of decentralization and identity for future research.

3.A Appendix

3.A.1 Marginal Effects

	(1)	(2)	(3)	(4)
	all identities	local	regional	national
Local identity				
1._tax morale*0._dec	-0.00822**	-0.0117***	-0.00899***	-0.00813**
	(-2.62)	(-4.24)	(-3.75)	(-3.18)
1._tax morale*1._dec	-0.00235*	-0.00222**	-0.00121	-0.00164
	(-2.25)	(-2.78)	(-1.47)	(-1.88)
2._tax morale*0._dec	-0.00168*	-0.00245***	-0.00191**	-0.00169**
	(-2.38)	(-3.43)	(-3.16)	(-2.79)
2._tax morale*1._dec	-0.000600*	-0.000583*	-0.000317	-0.000421
	(-2.15)	(-2.48)	(-1.38)	(-1.81)
3._tax morale*0._dec	-0.00409**	-0.00584***	-0.00456***	-0.00413***
	(-2.63)	(-4.32)	(-3.78)	(-3.35)
3._tax morale*1._dec	-0.00158*	-0.00150**	-0.000816	-0.00111
	(-2.20)	(-2.60)	(-1.43)	(-1.76)
4._tax morale*0._dec	-0.00398**	-0.00573***	-0.00446***	-0.00407***
	(-2.67)	(-4.47)	(-3.87)	(-3.36)
4._tax morale*1._dec	-0.00170*	-0.00163**	-0.000873	-0.00120
	(-2.24)	(-2.68)	(-1.43)	(-1.83)
5._tax morale*0._dec	-0.00465**	-0.00664***	-0.00523***	-0.00472***
	(-2.68)	(-4.52)	(-3.92)	(-3.41)
5._tax morale*1._dec	-0.00218*	-0.00207**	-0.00112	-0.00152
	(-2.20)	(-2.71)	(-1.47)	(-1.89)
6._tax morale*0._dec	-0.00926**	-0.0132***	-0.0105***	-0.00946***
	(-2.77)	(-4.95)	(-4.17)	(-3.55)
6._tax morale*1._dec	-0.00505*	-0.00479**	-0.00259	-0.00353

	(-2.29)	(-2.76)	(-1.44)	(-1.85)
7._tax morale*0._dec	-0.00406**	-0.00583***	-0.00463***	-0.00419***
	(-2.76)	(-4.88)	(-4.09)	(-3.59)
7._tax morale*1._dec	-0.00259*	-0.00248**	-0.00133	-0.00182
	(-2.26)	(-2.69)	(-1.43)	(-1.83)
8._tax morale*0._dec	-0.00571**	-0.00818***	-0.00659***	-0.00592***
	(-2.80)	(-5.08)	(-4.20)	(-3.68)
8._tax morale*1._dec	-0.00431*	-0.00410**	-0.00222	-0.00303
	(-2.28)	(-2.71)	(-1.42)	(-1.84)
9._tax morale*0._dec	-0.00468**	-0.00676***	-0.00552***	-0.00493***
	(-2.74)	(-4.77)	(-4.00)	(-3.63)
9._tax morale*1._dec	-0.00522*	-0.00500**	-0.00268	-0.00367
	(-2.32)	(-2.78)	(-1.44)	(-1.86)
10._tax morale*0._dec	0.0463**	0.0664***	0.0524***	0.0472***
	(2.79)	(5.09)	(4.25)	(3.62)
10._tax morale*1._dec	0.0256*	0.0244**	0.0132	0.0180
	(2.30)	(2.79)	(1.45)	(1.86)
Regional identity				
1._tax morale*0._dec	-0.00312			
	(-1.10)			
1._tax morale*1._dec	0.000821			
	(0.73)			
2._tax morale*0._dec	-0.000638			
	(-1.08)			
2._tax morale*1._dec	0.000210			
	(0.74)			
3._tax morale*0._dec	-0.00156			
	(-1.09)			

3._tax morale*1._dec	0.000554
	(0.73)
4._tax morale*0._dec	-0.00151
	(-1.09)
4._tax morale*1._dec	0.000594
	(0.73)
5._tax morale*0._dec	-0.00177
	(-1.09)
5._tax morale*1._dec	0.000762
	(0.73)
6._tax morale*0._dec	-0.00352
	(-1.10)
6._tax morale*1._dec	0.00177
	(0.73)
7._tax morale*0._dec	-0.00154
	(-1.09)
7._tax morale*1._dec	0.000907
	(0.73)
8._tax morale*0._dec	-0.00217
	(-1.09)
8._tax morale*1._dec	0.00151
	(0.74)
9._tax morale*0._dec	-0.00178
	(-1.08)
9._tax morale*1._dec	0.00183
	(0.74)
10._tax morale*0._dec	0.0176
	(1.10)

10._tax morale*1._dec	-0.00895
	(-0.73)
National identity	
1._tax morale*0._dec	-0.00315
	(-1.14)
1._tax morale*1._dec	-0.00121
	(-1.25)
2._tax morale*0._dec	-0.000644
	(-1.12)
2._tax morale*1._dec	-0.000309
	(-1.25)
3._tax morale*0._dec	-0.00157
	(-1.16)
3._tax morale*1._dec	-0.000815
	(-1.20)
4._tax morale*0._dec	-0.00153
	(-1.15)
4._tax morale*1._dec	-0.000873
	(-1.24)
5._tax morale*0._dec	-0.00178
	(-1.16)
5._tax morale*1._dec	-0.00112
	(-1.25)
6._tax morale*0._dec	-0.00355
	(-1.16)
6._tax morale*1._dec	-0.00260
	(-1.24)
7._tax morale*0._dec	-0.00156

	(-1.17)			
7._tax morale*1._dec	-0.00133			
	(-1.23)			
8._tax morale*0._dec	-0.00219			
	(-1.18)			
8._tax morale*1._dec	-0.00222			
	(-1.24)			
9._tax morale*0._dec	-0.00180			
	(-1.18)			
9._tax morale*1._dec	-0.00269			
	(-1.24)			
10._tax morale*0._dec	0.0178			
	(1.17)			
10._tax morale*1._dec	0.0132			
	(1.24)			
N	27489	27563	27533	27533

The first column refers to all identity levels, stacked,

the other columns refer to the respective identity level at the top.

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.8: Marginal Effects of Baseline Regression Results

3.A.2 Robustness

In this section, we perform some robustness checks on the decentralization measure to further validate our previous results. It is possible that decentralization in general is not a determinant of tax morale, but rather the share of expenditures or tax revenue on a subnational level. The decentralization dummy we use in our main analysis includes these shares. As a robustness check, we replace the decentralization dummy with a dummy for above-average subnational tax revenue and a dummy for above-average subnational expenditure, respectively.

The results are summarized in Table 3.9. The first four columns show the results for the dummy describing expenditure, while the results in the last four columns are based on the dummy for revenue. The variable decentralization describes either one or the other. Comparing the results with thoses of the baseline regression in Table 3.2, we find that using a dummy for the subnational revenue share does not change much. This is due to a very high correlation between this dummy and our previous decentralization dummy. Looking at the first four columns and thus at subnational expenditure, the situation changes. The effect of identity is weakened a lot and basically insignificant. Moreover, the sign of decentralization reverses, although it is still significant. The interaction is now positive and insignificant. It is not clear why the effect of decentralization is now negative. However, the other effects are only smaller or in a similar range, but basically point in the same direction as the baseline regression.

To summarize, using other measures of decentralization can change the results. In particular, looking at expenditures as a measure of decentralization, changes the interpretation. This is of special interest because our previous hypotheses were essentially based on the supply and quality of public goods, which depend on the expenditures. However, the results of this robustness check do not deviate so much that we would need to question our baseline results.

	(1) Tax morale	(2) Tax morale	(3) Tax morale	(4) Tax morale	(5) Tax morale	(6) Tax morale	(7) Tax morale	(8) Tax morale
Local identity	0.0384	0.0547*			0.129***	0.184***		
	(0.0416)	(0.0304)			(0.0463)	(0.0368)		
Regional identity	-0.0390		0.0434		0.0489		0.145***	
	(0.0498)		(0.0329)		(0.0446)		(0.0345)	
National identity	0.105**			0.104***	0.0494			0.131***
	(0.0449)			(0.0348)	(0.0423)			(0.0364)
decentralization	-0.789***	-0.861***	-0.780***	-0.590***	0.802***	0.729***	0.686***	0.614***
	(0.208)	(0.175)	(0.179)	(0.183)	(0.193)	(0.155)	(0.145)	(0.164)
0.decentralization*Local identity	0	0			0	0		
	(.)	(.)			(.)	(.)		
1.decentralization*Local identity	0.0560	0.0517			-0.0498	-0.109**		
	(0.0540)	(0.0412)			(0.0575)	(0.0454)		
0.decentralization*Regional identity	0		0		0		0	
	(.)		(.)		(.)		(.)	
1.decentralization*Regional identity	0.0341		0.0254		-0.0765		-0.105**	
	(0.0607)		(0.0423)		(0.0581)		(0.0440)	
0.decentralization*National identity	0			0	0			0
	(.)			(.)	(.)			(.)
1.decentralization*National identity	-0.0665			-0.0360	-0.00884			-0.0752
	(0.0542)			(0.0449)	(0.0537)			(0.0470)

Standard errors in parentheses

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

Table 3.9: Regression with Different Dummies for Decentralization. Columns (1) -(4) Use a Decentralization Dummy Concerning Subnational Expenditure; Columns (5)-(8) Use One Concerning Subnational Revenue.

Bibliography

Acemoglu, D. and Robinson, J. A. (2000). Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective. *The Quarterly Journal of Economics*, 115 (4): 1167–1199.

Acemoglu, D., Ticchi, D., and Vindigni, A. (2011). Emergence and Persistence of Inefficient States. *Journal of the European Economic Association*, 9 (2): 177–208.

Acharya, V., Engle, R., and Richardson, M. (2012). Capital Shortfall: A New Approach to Ranking and Regulating Systemic Risks. *American Economic Review*, 102 (3): 59–64.

Acharya, V., Gale, D., and Yorulmazer, T. (2011). Rollover Risk and Market Freezes. *Journal of Finance*, 66 (4): 1177–1209.

Acharya, V., Pedersen, L. H., Philippon, T., and Richardson, M. (2017). Measuring Systemic Risk. *Review of Financial Studies*, 30 (1): 2–47.

Adam, K., Marcet, A., and Beutel, J. (2017). Stock Price Booms and Expected Capital Gains. *American Economic Review*, 107 (8): 2352–2408.

Adam, K. and Nagel, S. (2022). Expectations Data in Asset Pricing. *Working Paper*.

Adrian, T. and Brunnermeier, M. K. (2016). CoVar. *American Economic Review*, 106 (7): 1705–1741.

Akerlof, G. A. (1978). The Economics of "tagging" As Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning. *The American Economic Review*, 68 (1): 8–19.

Aliber, R. and Kindleberger, C. (2015). *Manias, Panics, and Crashes: A History of Financial Crises*. Palgrave Macmillan, 7th edition.

Allingham, M. G. and Sandmo, A. (1972). Income Tax Evasion: A Theoretical Analysis. *Journal of Public Economics*, 1 (3-4): 323–338.

Andreoni, J. (1990). Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving. *The Economic Journal*, 100 (401): 464–477.

Atmaz, A. (2014). A Dynamic Model of Short Interest. *Working Paper, London Business School.*

Atmaz, A. and Basak, S. (2016). Belief Dispersion in the Stock Market. *Working Paper.*

Barro, R. J. (1990). Government Spending in a Simple Model of Endogeneous Growth. *Journal of Political Economy*, 98 (5, Part 2): 103–125.

Becker, S. O. and Hornung, E. (2020). The Political Economy of the Prussian Three-Class Franchise. *The Journal of Economic History*, 80 (4): 1143–1188.

Belmonte, A., Dell'Anno, R., and Teobaldelli, D. (2018). Tax Morale, Aversion to Ethnic Diversity, and Decentralization. *European Journal of Political Economy*, 55: 204–223.

Bernanke, B. and Gertler, M. (1990). Financial Fragility and Economic Performance. *Quarterly Journal of Economics*, 105 (1): 87–114.

Besley, T. (2020). State Capacity, Reciprocity, and the Social Contract. *Econometrica*, 88: 1307–1335.

Besley, T. and Case, A. (1992). Incumbent Behavior: Vote Seeking, Tax Setting and Yardstick Competition. *Nber Working Paper*, 4041.

Besley, T. and Coate, S. (2003). Centralized Versus Decentralized Provision of Local Public Goods: A Political Economy Approach. *Journal of Public Economics*, 87: 2611–2637.

Besley, T. and Persson, T. (2009). The Origins of State Capacity: Property Rights, Taxation, and Politics. *American Economic Review*, 99 (4): 1218–44.

Besley, T. and Persson, T. (2010). State Capacity, Conflict, and Development. *Econometrica*, 78 (1): 1–34.

Besley, T. and Persson, T. (2011). *Pillars of Prosperity.* Princeton University Press, Princeton.

Bianchi, J. (2011). Overborrowing and Systemic Externalities in the Business Cycle. *American Economic Review*, 101 (7): 3400–3426.

Bianchi, J. and Mendoza, E. G. (2018). Optimal Time-Consistent Macroprudential Policy. *Journal of Political Economy*, 126 (2): 588–634.

Bierbrauer, F. J., Boyer, P. C., and Peichl, A. (2021). Politically Feasible Reforms of Nonlinear Tax Systems. *American Economic Review*, 111 (1): 153–91.

Bisias, D., Flood, M., Lo, A. W., and Valavanis, S. (2012). A Survey of Systemic Risk Analytics. *Annual Review of Financial Economics*, 4 (1): 255–296.

Blesse, S. and Baskaran, T. (2016). Do Municipal Mergers Reduce Costs? Evidence from a German federal state. *Regional Science and Urban Economics*, 59: 54–74.

Blume, L. E., Cogley, T., Easley, D. A., Sargent, T. J., and Tsyrennikov, V. (2018). A Case for Incomplete Markets. *Journal of Economic Theory*, 178: 191–221.

Boadway, R. and Pestieau, P. (2006). Tagging and Redistributive Taxation. *Annales D'economie Et De Statistique*, pages 123–147.

Bosco, L. and Mittone, L. (1997). Tax Evasion and Moral Constraints: Some Experimental Evidence. *Kyklos*, 50: 297–324.

Brunnermeier, M., Simsek, A., and Xiong, W. (2014). A Welfare Criterion for Models With Distorted Beliefs. *Quarterly Journal of Economics*, 129 (4): 1753–1797.

Brunnermeier, M. K. and Oehmke, M. (2013). Bubbles, Financial Crises, and Systemic Risk. *Handbook of the Economics of Finance.*

Brunnermeier, M. K. and Pedersen, L. H. (2009). Market Liquidity and Funding Liquidity. *Review of Financial Studies*, 22 (6): 2201–2238.

Buggeln, M. (2022). *Das Versprechen der Gleichheit: Steuern und Soziale Ungleichheit in Deutschland von 1871 bis Heute*. Suhrkamp Verlag.

Caballero, R. J. and Krishnamurthy, A. (2003). Excessive Dollar Debt: Financial Development and Underinsurance. *Journal of Finance*, 58 (2): 867–893.

Caballero, R. J. and Lorenzoni, G. (2014). Persistent Appreciations and Overshooting: A Normative Analysis. *IMF Economic Review*, 62 (1): 1–47.

Cassette, A. and Paty, S. (2010). Fiscal Decentralization and the Size of Government: A European Country Empirical Analysis. *Public Choice*, 143: 173–189.

Cheng, I.-H., Raina, S., and Xiong, W. (2014). Wall Street and the Housing Bubble. *American Economic Review*, 104 (9): 2797–2829.

Cremer, H., Gahvari, F., and Lozachmeur, J.-M. (2010). Tagging and Income Taxation: Theory and an Application. *American Economic Journal: Economic Policy*, 2 (1): 31–50.

Cvitanic, J. and Malamud, S. (2011). Price Impact and Portfolio Impact. *Journal of Financial Economics*, 100 (1): 201–225.

Dávila, E. (2023). Optimal Financial Transaction Taxes. *Journal of Finance*, 78 (1): 5–61.

Dávila, E. and Walther, A. (2023). Prudential Policy with Distorted Beliefs. *American Economic Review*, 113 (7): 1967–2006.

Destatis (2022a). Fachserie 14/Reihe 3.1. Finanzen Und Steuern, Table 3.2.1.

Destatis (2022b). Fachserie 16/Reihe 2.5. Verdienste Und Arbeitskosten.

Destatis (2022c). Fachserie 1/Reihe 4.1. Erwerbsbeteiligung Der Bevölkerung - Ergebnisse Des Mikrozensus Zum Arbeitsmarkt.

Destatis (2023). Statistisches Bundesamt, Rechnungsergebnisse des öffentlichen Gesamthaushalts, Tabelle 71141 und Fortschreibung des Bevölkerungsstandes, Tabelle 12411.

Diamond, P. A. (1998). Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates. *American Economic Review*, pages 83–95.

Dávila, E. and Korinek, A. (2018). Pecuniary Externalities in Economies with Financial Frictions. *Review of Economic Studies*, 85 (1): 352–395.

Engel, E. (1868). Die Ergebnisse der Classensteuer, der classificirten Einkommensteuer und der Mahl- und Schlachtsteuer im Preussischen Staate. *Zeitschrift Des Königlich Preussischen Statistischen Bureaus*, 8.

Engel, E. (1875). Die Klassen- und klassificirte Einkommensteuer und die Einkommensvertheilung im preussischen Staat. *Zeitschrift Des Königlich Preussischen Statistischen Bureaus*, 15.

EVS (2022). EVS Trend File 1981-2017: Integrated Dataset (EVS1981-2017). Gesis Data Archive, Cologne. Za7503 Data File Version 3.0.0, Doi:10.4232/1.14021.

Feld, L. P., Kirchgässner, G., and Schaltegger, C. A. (2003). Decentralized Taxation and the Size of Government: Evidence from Swiss state and local governments. *Cesifo Working Paper*, 1087.

Fisher, I. (1933). The Debt-Deflation Theory of Great Depressions. *Econometrica*, 1 (4): 337–357.

Fritz, B. (2015). *The Political Economy of Amalgamation: Evidence from a German state*. PhD thesis, Albert-Ludwigs-Universität Freiburg.

Gangl, K., Torgler, B., and Kirchler, E. (2016). Patriotism's Impact on Cooperation with the State: An Experimental Study on Tax Compliance. *Political Psychology*, 37: 867–881.

Gayer, G., Gilboa, I., Samuelson, L., and Schmeidler, D. (2014). Pareto Efficiency with Different Beliefs. *Journal of Legal Studies*, 43 (S2): 151–171.

Geanakoplos, J. (1996). Promises Promises. *Cowles Foundation Discussion Papers 1143*.

Geanakoplos, J. (2003). *Liquidity, Default, and Crashes: Endogenous Contracts in General Equilibrium*, volume 2 of *Econometric Society Monographs*, pages 170–205. Cambridge University Press.

Geanakoplos, J. (2010). The Leverage Cycle. *Nber Macroeconomics Annual*, 24 (1): 1–66.

Geanakoplos, J. D. and Polemarchakis, H. M. (1986). Existence, Regularity, and Constrained Suboptimality of Competitive Allocations when the Asset Market is Incomplete. *Cowles Foundation Paper 652*.

Gennaioli, N. and Shleifer, A. (2018). *A Crisis of Beliefs*. Princeton University Press.

Gilboa, I., Samuelson, L., and Schmeidler, D. (2014). No-Betting-Pareto Dominance. *Econometrica*, 82 (4): 1405–1442.

Grasmick, H. G. and Bursik, R. J. (1990). Conscience, Significant Others, and Rational Choice: Extending the Deterrence Model. *Law & Society Review*, 24: 837–862.

Gromb, D. and Vayanos, D. (2002). Equilibrium and Welfare in Markets with Financially Constrained Arbitrageurs. *Journal of Financial Economics*, 66 (2-3): 361–407.

Güth, W., Levati, V., and Sausgruber, R. (2005). Tax Morale and (De-)centralization: An Experimental Study. *Public Choice*, 125: 171–188.

Harrison, J. M. and Kreps, D. M. (1978). Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations. *Quarterly Journal of Economics*, 92 (2): 323–336.

Hart, O. D. (1975). On the Optimality of Equilibrium when the Market Structure is Incomplete. *Journal of Economic Theory*, 11 (3): 418–443.

Hill, J. A. (1892). The Prussian Income Tax. *The Quarterly Journal of Economics*, 6 (2): 207–226.

Hoffmann, J. G. et al. (1840). *Die Lehre Von Den Steuern als Anleitung zu Gründlichen Urtheilen über das Steuerwesen: Mit besonderer Beziehung auf den Preussischen Staat*. Nicolaische Buchhandlung.

Jeanne, O. and Korinek, A. (2010). Excessive Volatility in Capital Flows: A Pigouvian Taxation Approach. *American Economic Review: Papers and Proceedings*, 100 (2): 403–407.

Jeanne, O. and Korinek, A. (2019). Managing Credit Booms and Busts: A Pigouvian Taxation Approach. *Journal of Monetary Economics*, 107: 2–17.

Jeanne, O. and Korinek, A. (2020). Macroprudential Regulation Versus Mopping Up After the Crash. *Review of Economic Studies*, 87 (3): 1470–1497.

Kaplan, G., Mitman, K., and Violante, G. L. (2020). The Housing Boom and Bust: Model Meets Evidence. *Journal of Political Economy*, 128 (9): 3285–3345.

Keynes, J. M. (1936). *The General Theory of Employment, Interest and Money*. Macmillan. 14th edition, 1973.

Kim, J. H. J. and Kim, B.-C. (2021). A Welfare Criterion with Endogenous Welfare Weights for Belief Disagreement Models. *Journal of Economic Behavior & Organization*, 191: 312–333.

Kiyotaki, N. and Moore, J. (1997). Credit Cycles. *Journal of Political Economy*, 105 (2): 211–248.

Königliches Statistisches Bureau (1898). Statistisches Handbuch für den Preußischen Staat. 2.

Konrad, K. A. and Qari, S. (2012). The Last Refuge of a Scoundrel? Patriotism and Tax Compliance. *Economica*, 79 (315): 516–533.

Lago-Peñas, I. and Lago-Peñas, S. (2010). The Determinants of Tax Morale in Comparative Perspective: Evidence from European Countries. *European Journal of Political Economy*, 26: 441–453.

Lorenzoni, G. (2008). Inefficient Credit Booms. *Review of Economic Studies*, 75 (3): 809–833.

Luttmer, E. F. P. and Singhal, M. (2014). Tax Morale. *Journal of Economic Perspectives*, 28 (4): 149–68.

López-Laborda, J., Vallés-Giménez, J., and Zárate-Marco, A. (2021). Personal Income Tax Compliance at the Regional Level: The Role of Persistence, Neighborhood, and Decentralization. *International Regional Science Review*, 44 (2): 289–317.

MacGregor, J. and Wilkinson, B. (2012). The Effect of Economic Patriotism on Tax Morale and Attitudes toward Tax Compliance. *Advances in Taxation*, 20: 159–180.

Mankiw, N. G. and Weinzierl, M. (2010). The Optimal Taxation of Height: A Case Study of Utilitarian Income Redistribution. *American Economic Journal: Economic Policy*, 2 (1): 155–76.

Mares, I. and Queralt, D. (2020). Fiscal Innovation in Nondemocratic Regimes: Elites and the Adoption of the Prussian Income Taxes of the 1890s. *Explorations in Economic History*, 77.

Martinez-Vazquez, J. and Torgler, B. (2009). The Evolution of Tax Morale in Modern Spain. *Journal of Economic Issues*, 43: 1–28.

Mendoza, E. G. (2002). Credit, Prices, and Crashes: Business Cycles with a Sudden Stop. In Edwards, S. and Frankel, J. A., editors, *Preventing Currency Crises in Emerging Markets*, NBER Conference Report Series, pages 335–383. University of Chicago Press.

Mian, A. and Sufi, A. (2022). Credit Supply and Housing Speculation. *Review of Financial Studies*, 35 (2): 680–719.

Minsky, H. P. (1977). The Financial Instability Hypothesis: An Interpretation of Keynes and an Alternative to "Standard" Theory. *Challenge*, 20 (1): 20–27.

Minsky, H. P. (1986). Stabilizing an Unstable Economy. *Hyman P. Minsky Archive. Paper 144.*

Murphy, K. (2003). Procedural Justice and Tax Compliance. *Australian Journal of Social Issues*, 38: 379–407.

Oates, W. E. (1972). Fiscal Federalism. *Books.*

Reinhart, C. M. and Rogoff, K. S. (2008). Is the 2007 U.S. Subprime Crisis so Different? An International Historical Comparison. *American Economic Review*, 98 (2): 339–344.

Roesel, F. (2017). Do Mergers of Large Local Governments Reduce Expenditures? Evidence from Germany Using the Synthetic Control Method. *European Journal of Political Economy*, 50: 22–36.

Scheinkman, J. A. and Xiong, W. (2003). Overconfidence and Speculative Bubbles. *Journal of Political Economy*, 111 (6): 1183–1220.

Schremmer, D. E. and Stern, W. (1989). *Taxation and Public Finance: Britain, France, and Germany*, volume 8 of *The Cambridge Economic History of Europe*, chapter 5, pages 315–494. Cambridge University Press.

Schremmer, E. (2013). *Steuern und Staatsfinanzen während der Industrialisierung Europas: England, Frankreich, Preußen und das Deutsche Reich 1800 bis 1914*. Springer-Verlag.

Schwarz, O. and Strutz, G. (1902). *Der Staatshaushalt und die Finanzen Preussens unter Benutzung Amtlicher Quellen*, volume 1. J. Guttentag.

Seligman, E. R. (1894). The Income Tax. *Political Science Quarterly*, pages 610–648.

Shaver, A., Carter, D. B., and Shawa, T. W. (2019). Terrain Ruggedness and Land Cover: Improved Data for Most Research Designs. *Conflict Management and Peace Science*, 36 (2): 191–218.

Shleifer, A. and Vishny, R. (1992). Liquidation Values and Debt Capacity: A Market Equilibrium Approach. *Journal of Finance*, 47 (4): 1343–1366.

Shleifer, A. and Vishny, R. (2011). Fire Sales in Finance and Macroeconomics. *Journal of Economic Perspectives*, 25 (1): 29–48.

Simsek, A. (2013). Belief Disagreements and Collateral Constraints. *Econometrica*, 81 (1): 1–53.

Simsek, A. (2021). The Macroeconomics of Financial Speculation. *Annual Review of Economics*, 13: 335–369.

Spoerer, M. (2015). Steuerlast, Steuerinzidenz und Steuerwettbewerb. In *Steuerlast, Steuerinzidenz und Steuerwettbewerb*. Akademie Verlag.

Stiglitz, J. E. (1982). The Inefficiency of the Stock Market Equilibrium. *Review of Economic Studies*, 49 (2): 241–261.

Stiglitz, J. E. (1989). Using Tax Policy to Curb Speculative Short-Term Trading. *Regulatory Reform of Stock and Futures Markets: A Special Issue of the Journal of Financial Services Research*, pages 3–17.

Summers, L. H. and Summers, V. P. (1989). When Financial Markets Work Too Well: A Cautious Case For a Securities Transactions Tax. *Journal of Financial Services Research*, 3: 261–286.

Söderlind, P. (2009). Why Disagreement May Not Matter (Much) for Asset Prices. *Finance Research Letters*, 6 (2): 73–82.

Teschemacher, H. (1912). *Die Einkommensteuer und die Revolution in Preußen: Eine Finanzwissenschaftliche und Allgemeingeschichtliche Studie über das Preußische Einkommensteuerprojekt von 1847*. Laupp.

Tiebout, C. M. (1956). A Pure Theory of Local Expenditures. *Journal of Political Economy*, 64: 416–424.

Tilly, R. (2010). The Distribution of Personal Income in Prussia, 1852 to 1875: An Exploratory Study. *Jahrbuch Für Wirtschaftsgeschichte/economic History Yearbook*, 51 (1): 175–194.

Tobin, J. (1978). A Proposal for International Monetary Reform. *Eastern Economic Journal*, 4 (3): 153–159.

Torgler, B. (2004). Tax Morale in Asian Countries. *Journal of Asian Economics*, 15 : 237–266.

Torgler, B. (2005). Tax Morale in Latin America. *Public Choice*, 122:133–157.

Torgler, B. and Schneider, F. (2004). Does Culture Influence Tax Morale? Evidence from Different European Countries. *Science*.

Torgler, B., Schneider, F., and Schaltegger, C. A. (2007). With or Against the People? The Impact of a Bottom-Up Approach on Tax Morale and the Shadow Economy. *Crema Working Paper*.

Torgler, B. and Werner, J. (2005). Tax Morale and Fiscal Autonomy: Evidence from Germany. *Public Finance Management*, 5.

Weinzierl, M. C. (2012). Why Do We Redistribute so Much but Tag so Little? The Principle of Equal Sacrifice and Optimal Taxation. Technical report, National Bureau of Economic Research.

Wenzel, M. (2004). An Analysis of Norm Processes in Tax Compliance. *Journal of Economic Psychology*, 25: 213–228.

Wenzel, M. (2007). The Multiplicity of Taxpayer Identities and Their Implications for Tax Ethics. *Law and Policy*, 29: 31–50.

WVS (2022). Haerpfer, C., Inglehart, R., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano, J., Lagos, M., Norris, P., Ponarin, E., Puranen, B., et al. (eds.). World Values Survey Trend File (1981-2022) Cross-National Data-Set. Data File Version 3.0.0. *Jd Systems Institute Wvsa Secretariat.*, doi:10.14281/18241.

Xiong, W. (2013). Bubbles, Crises, and Heterogeneous Beliefs. *Nber Working Paper*.

List of Applied Software

1. Matlab R2022A: Applied for the numerical example in chapter 1.
2. Excel 2016: Used for data cleaning and the descriptive analysis of historical Prussian data in chapter 2.
3. StataMP 17: Applied for the empirical analysis in chapter 3 and the graphs in chapter 2.
4. Texmaker 5.1.4: Used to compile the single papers and the thesis.

Marco Wysietzki

Center for Macroeconomic Research
University of Cologne
Albertus-Magnus-Platz
50923 Cologne, Germany

E-mail: wysietzki@wiso.uni-koeln.de
Phone: +49 221 470 1247

Current Positions

Center for Macroeconomic Research, University of Cologne
Teaching Assistant

Education

University of Cologne, PhD candidate, Economics, since Oct 2019

University of Bonn, M.Sc. Economics, Oct 2016 – Sep 2019

Louisiana State University, Semester Abroad, Economics, Aug 2017 – Dec 2017

University of Bonn, B.Sc. Economics, Oct 2014 – Apr 2016

University of Bonn, B.Sc. Mathematics, Oct 2011 – Sep 2014

Gymnasium Lindlar, Abitur (High School Certificate), Jul 2011

Teaching Experience

Center for Macroeconomic Research, University of Cologne,
Exercise: Competition in Economics and Politics
Exercise: Growth and Technical Change

Exercise: Introduction to Public Economics

Institute for Macroeconomics and Econometrics, University of Bonn,
Tutorial: Macroeconomics B
Tutorial: Macroeconomics A

Institute for Microeconomics, University of Bonn,
Tutorial: Microeconomics A

Institute for International Economic Policy, University of Bonn,
Tutorial: Introduction to Economics B

Scholarships

Cologne Graduate School, PhD Fellowship, Oct 2019 – Sep 2020

Cusanuswerk Foundation, Scholarship, Jan 2016 – Apr 2019

Other Qualifications

Languages: German, English, French
Computer skills: Matlab, LaTeX, Stata, R

Cologne, February 18, 2024

Eidesstattliche Erklärung nach § 9 Abs. 5 der Promotionsordnung vom 01.08.2022

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Weitere Personen, neben den ggf. in der Arbeit aufgeführten Koautorinnen und Koautoren, waren an der inhaltlich-materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten fremden Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind als solche kenntlich gemacht.

Ich versichere an Eides statt, dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von den angegebenen Teilpublikationen und eingebundenen Artikeln und Manuskripten - noch nicht veröffentlicht worden ist, sowie, dass ich eine Veröffentlichung der Dissertation vor Abschluss der Promotion nicht ohne Genehmigung des Promotionsausschusses vornehmen werde. Die Bestimmungen dieser Ordnung sind mir bekannt.

Darüber hinaus erkläre ich hiermit, dass ich die Leitlinien der Universität zu Köln zur Sicherung guter wissenschaftlichen Praxis gelesen und sie bei der Durchführung der der Dissertation zugrundeliegenden Arbeiten und der schriftlich verfassten Dissertation beachtet habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen. Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht. Ich versichere, dass ich nach bestem Wissen die reine Wahrheit gesagt und nichts verschwiegen habe.

Köln, 18.02.2024