Essays in Public Economics

Inaugural-Dissertation
zur Erlangung des Grades eines Doktors
der Wirtschafts- und Gesellschaftswissenschaften
durch die
Rechts- und Staatswissenschaftliche Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von
Rafael Christian Aigner
aus Regensburg

Bonn 2014

Dekan:	Prof. Dr. Klaus Sandmann
Erstreferent:	Prof. Martin Hellwig, PhD
Zweitreferent:	Prof. Dr. Felix Bierbrauer

Tag der mündlichen Prüfung: 2. April 2014

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn
(http://hss.ulb.uni-bonn.de/diss_online) elektronisch publiziert.

Acknowledgments

WRITING A THESIS is a great endeavor. I had the privilege not only to undertake this project, but to undertake it in good company; many great individuals have supported and inspired me along the way. I owe them an enormous amount of gratitude.

Martin Hellwig, my supervisor, gave me the freedom to pursue my own research agenda. His guidance was vital during all stages of my dissertation. I am indebted to his knowledge, support, and advice.

I owe a lot to Felix Bierbrauer who has influenced me significantly in writing this thesis. His excellent course in Public Economic Theory fostered my interest in Public Economics and led me to write Chapter 2. Numerous discussions during my PhD years were of great support and encouragement. Chapter 1 is based on joint work with Felix.

Matthias Lang is the coauthor of the paper presented in Chapter 3. I thank him for investing so much time in this joint project. I also thank Gilat Levy who advised me on this paper during my time at LSE.

All chapters of this thesis have benefited from various discussions and comments. In this respect, I like to thank Sophie Bade, Felix Bierbrauer, Pierre Boyer, Estelle Cantillon, Micael Castanheira, Brian Cooper, Christoph Engel, Emanuel Hansen, Stephen Hansen, Martin Hellwig, Christian Kellner, Marco Kleine, Daniel Krämer, Simon Lange, Mark Le Quement, Gilat Levy, Aniol Llorente-Saguer, Benny Moldovanu, Torsten Persson, Johannes Spinnewijn, Dezsö Szalay, Christian Traxler, Alfons Weichenrieder, and two anonymous referees.

I thank my fellow graduate students, in particular Benjamin Born, Deniz Dizdar, Tilman Drerup, Sebastian Ebert, Andreas Esser, Patrick Hürtgen, Matthias Lang, Stephan Luck, Alexandra Peters, Johannes Pfeifer, Paul Schempp, Gregor Schwerhoff, Christian Seel, Mirko Seithe, Inga van den Bongard, Christoph Wagner, and Florian Zimmermann, as well as my colleagues at the Max Planck Institute, in particular Alia Gizatulina, Olga Gorelkina, Dominik Grafenhofer, Ioanna Grypari, and Berenike Waubert de Puiseau, for giving me such a nice time in Bonn – on and off campus.

I spent my time as a graduate student at the Bonn Graduate School of Economics (BGSE), the London School of Economics and Political Sciences (LSE), and the Max Planck Institute for Research on Collective Goods. I thank these institutions for offering such a great learning and research environment. Special thanks go to Urs Schweizer for making the BGSE what it is, and to Silke Kinzig and Pamela Mertens for their work in the background. The same holds for the whole administration of the MPI. Monika Stimpson deserves special thanks for swiftly handling countless requests.

I am grateful for financial support by the Max Planck Institute for Research on Collective Goods and by the Deutsche Forschungsgemeinschaft (DFG) through the Bonn Graduate School of Economics.

Finally, and most importantly, I thank Lea and my whole family for their love and unconditional support during my PhD years.

Contents

List of Figures

Introduction

This thesis consists of three chapters. Although they are fairly independent, they share a common research approach. All chapters apply microeconomic theory to analyze economic or political questions. Chapters 1 and 2 are contributions to tax theory. Chapter 1 mixes normative and positive analysis. It centers around the question how taxes in the financial sector affect economic outcomes. Chapter 2 takes a mostly normative perspective. It addresses the question how taxes should be set in order to deal with pollution and inequality. Chapter 3 is different; it is a positive analysis of voting behavior in a political economy model with small parties. More precisely, it ask why a rational voter might want to vote for a party that has no chance to enter parliament.

Chapter 1 The first chapter of this thesis is based on joint work with Felix Bierbrauer (Aigner and Bierbrauer, 2013). The chapter addresses the question how taxes in the financial sector affect the real economy. It presents a general equilibrium model with financial intermediation. Banks hire workers and use real inputs to transform risky loans into savings deposits with fixed returns. Households offer labor, save their earnings, and consume. Firms take out loans to run their business and produce consumption goods. The government collects taxes to cover a public budget. The insights of the model are relevant to both the theory of taxing financial intermediation and the respective policy discussion.

The first set of results concerns the tax burden. When the government needs to collect a fixed amount of tax revenues, and equilibrium profits of banks are positive, then one can shift the tax burden from households to banks (and vice versa) by appropriately tweaking tax rates. Tax systems which favor the banks induce a *small* financial sector and distort labor supply downwards. Tax systems which favor the households induce a *large* financial sector and heavily tax profits.

The second set of results concerns the value added tax (VAT). Traditionally, financial services are exempt from the VAT. Various authors consider this an imperfection of the tax system and have proposed reforms to include the financial sector into the VAT scheme. The model presented in Chapter 1 is suited to analyze the general equilibrium effects of these reforms. In part it confirms the claimed reform effects. Yet, it also questions the necessity of these reforms and shows that simple intuitions based on partial equilibrium analyses can be misleading from time to time.

Chapter 2 The second chapter of the thesis is a slightly extended version of a paper which is currently in press (Aigner, forthcoming). The chapter asks how equity concerns affect the optimal level of environmental taxation. The short answer is that higher levels of desired redistribution call for lower levels of green taxes. That is not the whole story, though. I show that the relation of optimal environmental taxes and redistribution depends on the available labor tax instruments. When first-best instruments – which allow for distortion-free redistribution – are available, more redistribution actually requires higher green taxes. The discussion clarifies a crucial difference between optimality rules and optimal levels in the theory of Pigouvian taxation.

The link between redistribution and environmental taxation originates from the marginal cost of public funds. These cost measure how individuals' well-being is affected when they have to pay taxes. When redistribution is distortive, marginal cost of public funds are high and grow with increasing redistribution. High marginal cost of public funds come with a low optimal environmental tax level: The tax designer trades off the social harm from pollution against additional revenues from taxing pollution. The higher the marginal cost of public fund are, the more valuable are revenues from environmental taxation. Thus a smaller amount suffices to cover damages from pollution.

Chapter 3 The third chapter of the thesis is a slightly modified version of a joint paper with Matthias Lang (Aigner and Lang, 2013). It is motivated by an empirical puzzle. In quite a few past elections, small parties gained a significant fraction of votes but failed the election threshold and so were not able to enter parliament. Polls before the election often suggested such an outcome. Thus voters knew to a certain degree that their vote would be wasted – in the sense that it would not count toward representation in parliament.

We offer an explanation that rationalizes such a voting behavior. The voters in our model are interested only in the policy outcome, which is decided in parliament. Nevertheless, they have a strategic incentive to vote for a party that does not enter parliament. The reason is that voters are forward-looking. Voting for the small party at the current election serves as a signal and fosters the party's chances to enter parliament in the next election. Voter trade off influence today against a future benefit. Effectively, they *invest* their vote.

Thirty-plus years ago, when I was a graduate student in economics, only the least ambitious of my classmates sought careers in the financial world. Even then, investment banks paid more than teaching or public service – but not that much more, and anyway, everyone knew that banking was, well, boring.

Paul Krugman, April 2009

1

Taxing Wall Street: The Case of Boring Banking

1.1 Introduction

Financial intermediation accounts for a significant fraction of GDP. Its tax treatment has considerable effects on fiscal revenues and economic efficiency. The revenue aspect in particular has become a salient issue on the global policy agenda since the 2008 financial crisis has burdened public budgets and dramatically increased public debt levels. Various EU countries, like the UK or Germany, have implemented or announced new kinds of 'bank taxes'. The introduction of a financial transactions tax is currently being debated in the EU.[1]

The financial crisis has also reinforced academic interest in excessive behavior, moral hazard, and other market failures associated with the modern financial system. Recent tax theory explores the scope for corrective taxes (for an overview, see Keen, 2011). Yet, the literature struggles with a more basic question. Most of the more than 120 countries in the world that have a value added tax fully or partially exempt the financial sector (Zee, 2005). But why should financial companies

[1]For data on financial sector size, see, e.g., Huizinga (2002) or Lockwood (2014). They report values in the range of 3% to 8% of GDP for a range of developed countries. For a descriptive list of new taxes on financial institutions, see Keen (2011, Box 1). For figures on fiscal exposure and dept level development, see International Montary Fund (2010, p. 2). For a discussion of the financial transaction tax, see Bierbrauer (2013).

be treated differently than regular businesses? Traditionally, the answer would be that the nature of financial intermediation and the prevalence of margin-based pricing makes taxing those services technically difficult. By now, however, the literature provides various ways to overcome these obstacles. In the European Union, the VAT treatment of financial services has been on the policy agenda since the mid-1990. (e.g., de la Feria and Lockwood, 2012; Huizinga, 2002; Poddar and English, 1997)

In this chapter, we propose a model in which the financial sector provides a simple form of credit intermediation. Market failures like excessive risk-taking are absent. There are individual risks, but aggregate returns are deterministic. Banks face neither liquidity nor solvency risks. In other words, *banking is boring*. But because our model is basic, it is able to address a basic question: What are the effects of taxation in the financial sector? Given the state of the political debate and the academic literature, on would expect this question to be answered. But, citing Keen (2011, p. 2), "one striking aspect of these policy developments and of the wider debate ... is that they have been almost entirely unguided by the public finance literature on the topic – because there is hardly any."

The contribution of this chapter is twofold. First, we perform a classical incidence analysis and check to what extent it is possible to let the financial sector bear the burden of funding the public budget. Second, we reconsider the proposals made to reform the VAT treatment of financial transactions and determine their effects on real economic outcomes.

For our analysis we develop a framework with the following features. 1. We have a complete general equilibrium model, based on primitives such as endowments and technologies. 2. Banks have a technology to transform risk. They act as intermediaries between depositors seeking safe returns and businesses issuing risky bonds. Intermediation is costly and requires the input of real resources. 3. Banks may or may not realize profits in equilibrium, depending on assumption on the intermediation technology. 4. Fiscal authorities use linear taxes to create revenue towards a public budget. Due to the general equilibrium nature of the model, the framework allows us to determine how taxes change the equilibrium allocation and how they influence welfare.

A first set of results addresses classical Ramsey-style questions: Suppose the state faces an exogenous revenue requirement to fund public expenditures. Which tax systems yield the desired revenues? Who carries the tax burden? The answers to these questions depend on the intermediation technology.

When financial intermediation is a constant return to scale technology, equilibrium profits are zero. Then banks cannot carry any tax burden, but pass on all taxes to households. So any increase in public expenditures comes at the expense of households. Under some mild elasticity assumptions, it does not matter which taxes are used; all tax systems that raise the same revenues induce the same burden on households. In a sense, the 'overall' tax level is relevant but the tax base choice is not. (Proposition 1.3)

When marginal returns to financial intermediation diminish, banks make positive profits in equilibrium. Different tax systems now have different distributional consequences. Holding revenues constant, the tax designer faces a trade off between burdening the households and burdening the banks. In this situation, making banks better off implies greater use of distortive tax instruments and induces lower financial activity. Making households better off, on the other hand, increases economic output and financial activity. (Proposition 1.4)

We contribute to an existing theoretical literature on the optimal VAT treatment of financial transactions. Auerbach and Gordon (2002) decompose the price of a consumption good into the cost of producing the good and the transactions cost of acquiring it. They find that a VAT including these transaction costs is equivalent to a tax on labor income and conclude that financial services should be taxed under to VAT. Grubert and Mackie (2000) obtain a different conclusion. Based on a view that financial services "provide the funds used to purchase fully taxable consumption goods" they find that "[as] non-consumption goods, such financial services should not be in the base of a consumption tax" (ibid., pp. 23 f.). While Chia and Whalley (1999) support the argument, Boadway and Keen (2003) reject the underlying intuition. They argue that it is a fallacy to believe that goods yielding no direct utility should automatically go untaxed (ibid., pp. 60–61). They also make a case, though, that financial services payed for by means of a spread should go untaxed, while fixed fees should be taxed. Jack (2000) obtains a similar result.

The aforementioned results are of a partial nature. Basically, they check how various tax schemes affect the household budget or rather the after-tax price ratios faced by households. They assume that before-tax fees charged by financial service providers are, more or less, exogenously fixed, i.e., independent of the tax schedule. When analyzing effects of taxation they essentially leave out both the banking and the production sector.[2] Yet, financial services are resource-intensive intermediation services – matching lenders with borrowers. To understand the effects of taxation, a notion of financial sector technology is needed. Some taxation papers do provide micro-foundations for financial sector technology. Lockwood (2014) studies payment services in a dynamic Ramsey model. Caminal (1997) considers a model with endogenous tax collection cost.

[2] Jack (2000) does analyze tax treatment of business user, but uses a separate model, leaving out consumers. Boadway and Keen (2003) verbally argue that businesses should not be taxed referring to Diamond and Mirrlees (1971).

Overall, the literature on VAT treatment is rather small. One reason might be the famous productive efficiency result. After all, financial services are, by and large, intermediate goods. Those should go untaxed according to Diamond and Mirrlees (1971). Typically, the value added tax is collected using an invoice-credit method. A universal VAT is then equivalent to a sales tax that applies to final consumers only. If the VAT exempts some businesses, though, the chain of invoice-credits is broken and not all taxes are passed on to final consumers. (We detail on this reasoning in Section 1.5.) Therefore, when a general VAT is in place, productive efficiency requires the inclusion of financial transactions rather than their exemption.

Productive efficiency means that different inputs are efficiently allocated among all production units. Diamond and Mirrlees (ibid.) model production in a rather abstract way. Their result affects financial intermediation only insofar as financial services can be interpreted as a 'standard' input good. In our setting, we explicitly model financial intermediation as a special technology to transform risky claims. One the other hand, the optimal allocation of different primitive inputs is less prominent in our model, because ultimately labor is the only input available in our economy. Another difference to Diamond and Mirrlees (ibid.) is the treatment of profits. The productive efficiency result requires constant returns to scale technologies or a 100% profit tax; positive after-tax profits can imply productive inefficiency (Mirrlees, 1972). We do allow for positive after-tax profits but abstract from incentive considerations; i.e., firms do provide output even if their financial compensation is exactly zero.

Policy proposals The exemption of financial services from the value added tax has led to a discussion on reforming the VAT system. Authors like Poddar and English (1997), Huizinga (2002), and Keen, Krelove, and Norregaard (2010) have put forward proposals to include the financial sector into the VAT scheme. Lost revenues and efficiency concerns are the major motivations for these proposals. Most of the arguments supporting them are based on partial equilibrium insights, basic tax principles or general economic intuition. Yet, few theoretical insights exist on the general equilibrium effects of these proposals, notably the induced tax burden.

We use our model to examine three potential tax reforms. We analyze their effects on economic outcomes and clarify who gains and who looses. Our results vary with the assumptions made upon intermediation technology. In many (but not all) instances we can confirm the expected reform effects (Propositions 1.6, 1.7, and 1.9). But for most cases we cannot confirm the rationales underlying the reform proposals. In particular, when financial intermediation is a constant-return-to-scale technology, none of the reforms is necessary; adjusting existing tax rates could induce the very same effects on economic outcomes (Proposition 1.2).

The chapter is organized as follows. In Section 1.2 we present the model. In Section 1.3, we characterize the competitive equilibrium. In Section 1.4, we detail on taxation. In Section 1.5, we evaluate the effects of reforming the VAT system. Section 1.6 concludes. Omitted proofs are relegated to Appendix 1.A.

1.2 Model

The economy consists of households, food producers, input producers, and banks. There are two periods: the first one is dedicated to production; in the second one, consumption takes place. Food producers employ labor h to produce food x. Households provide labor and consume food. Banks act as intermediaries and provide financial services. They employ labor as well as an input good z, provided by input producers. Fiscal authorities collect taxes used for public consumption.

The need for financial intermediation is two-sided. Households demand banking services because there is no use for their wages in period 1. By putting their earnings into a bank deposit account, they effectively transfer their funds into the second period.[3] Food producers demand banking services because they have no funds to start with and earn no revenues until the second period. To pay their workers in the first period, they must take out loans. In principle, food producers could promise their workers some share of future revenues instead of paying wages. Yet, food production is uncertain (as detailed below). Workers demand a non-random immediate paycheck rather than a risky claim on revenues.[4] The major merit of financial intermediation is thus a transformation of risk. By contracting with many different producers, banks can diversify the production risk and offer depositors a safe return.

We address two sets of questions: First, what are the Pareto-efficient allocations given an exogenous revenue requirement? Second, how do changes in the tax system affect equilibrium outcomes such as fiscal revenue and economic surplus?

[3] We could also allow households to put their earning in some storage technology. We would then focus on equilibria in which households chose not to use it but rather opt for deposits. By assuming right away that the deposit account is the best available use for households' earnings, we remove this layer of complexity.

[4] The underlying idea is that workers are risk-averse and would like to diversify the employer default risk. Doing the diversification on their own (e.g., by working for many different employers) comes with prohibitively high transaction cost, though. To save on notation and reduce complexity, we do not model this decision explicitly. We merely assume that workers have to be paid in period 1. This creates firms' loan demand.

Taxation We consider a set of linear ad valorem taxes. The set contains the 'traditional' taxes on (nonfinancial) commodities, labor, and profits. We also allow for less-traditional instruments, namely a *cash-flow tax* t_C and a *financial activities tax* t_F. Under a cash-flow tax, cash inflows from financial transactions, like receiving a loan or an interest payment, are taxable. Cash outflows, like granting a loan or paying interest, on the other hand, allow for a tax credit. Under a financial activities tax, both wages and profits generated by a financial institution are subject to a tax.

We also address the issue of (non-)deductibility of VAT paid on inputs by means of a variable t_{zC}, which we consider part of the tax system. Banks can deduct t_{zC} from the price of inputs goods. Typically, t_{zC} will either be zero (no deductibility) or will equal the tax on the input good (full deductibility). In practice, exemption of financial products from VAT typically features no deductibility. The cash-flow approach proposed in the literature would include cash flows in the VAT scheme and allow for full deductibility.

Tax revenues are immediately spent on labor or food, depending on the period. Put differently, the government does not trade on financial markets. For given tax rates, prices for real goods and financial services emerge at competitive markets on which the households and all firms as well as the government act a price takers.

Financial services There are two types of financial services, *loans* l and *deposits* d. A loan is a contract between a bank and a producer. The bank hands over $p_l l$ units of cash in period 1. The producer promises to pay back $p_x l$ in period 2. This is the value of l units of food. So the food price serves as the numeraire for financial contracts. A deposit is a contract between a bank and a household. In period one, the household hands over $p_d d$ units of cash to the bank. In period two, it receives $p_x d$ units of cash from the bank. Again this is equivalent to the delivery of d units of the numeraire good food. The implied interest rates are $(p_x/p_d - 1)$ for deposits and $(p_x/p_l - 1)$ for loans.

If used, a cash flow tax t_C applies to all cash flows associated with financial products. Analogous to a value added scheme with tax credits, the each inflow of cash is subject to taxation but each outflow of cash allows for a tax credit. Households, however, are not part of the scheme. They never pay cash flow taxes nor can they claim a tax credit.

The household problem There is a continuum of measure one of homogeneous households who enjoy consumption x_H and dislike working h_H. They choose $A_H = (x_H, h_H, d_H)$ in order to maximize

$$u(x_H, h_H)$$

subject to

$$h_H p_h(1 - t_h) \geq p_d d_H \tag{1.1}$$

$$p_x d_H \geq p_x x_H \tag{1.2}$$

There is one constraint for each period. In the first period, households can deposit no more than their net earnings $h_H p_h(1 - t_h)$. In the second period, they receives p_x for each unit deposited and purchases x_H at the very same price per unit. Utility u is strictly increasing and concave in x_H, and is strictly decreasing and convex in h_H. Furthermore, u is such that the household problem has a unique solution which is strictly greater than zero and bounded above.

The food producer problem There is a continuum of measure one of homogeneous food producers. Food producers have an intertemporal technology to transform labor h into food x. In the first period, food producers employ workers h_F at wage rate p_h. They generate no revenues in period 1, though. Also, they have no initial endowment. To pay their workforce, they have to take out a loan l_F. They receive p_l units of cash for each loan unit. This constitutes a cash inflow from a financial product and they pay cash-flow taxes amounting to $t_C l_F$. In exchange for loans l_F, food producers promise to repay $p_x l_F$ in period 2. The net cash flow of food producers in period 1 must be nonnegative; the respective constraint reads

$$(1 - t_C) p_l l_F \geq p_h h_F\,. \tag{1.3}$$

Labor employed in period 1 transforms into food output, available in period 2. The output level is random. With probability $\alpha \leq 1$ production is successful. In that case, each unit of labor input h_F from period 1 yields one unit of food output which is sold at price p_x in period 2. From the revenues, food producers pay VAT $t_x p_x h_F$, and refund $p_x l_F$ to their creditors, while receiving a tax credit $t_C p_x l_F$ for the cash outflow. With probability $(1 - \alpha)$, production is unsuccessful, inducing zero revenues and a full default on all debt obligations. Food producers maximize the expected net cash flow of period 2, which is

$$\alpha p_x \left[(1 - t_x) h_F - (1 - t_C) l_F\right]\,. \tag{1.4}$$

To sum up, the *food producer problem* is to choose $A_H = (h_F, l_F)$ in order to maximize (1.4) subject to (1.3).

The success of individual producers is independent from other producers. By a law of large numbers, total production is thus nonrandom and equal to αh_F. Notice that the optimal food producer choice does not depend on the price of food, p_x. While period-two revenues do depend on p_x, so do period-two costs because loans are an obligation to deliver the cash necessary to buy food.

The input producer problem There is a continuum of measure one of homogeneous input producers. They choose $A_I = (h_I, z_I)$ in order to maximize

$$(1 - t_z)p_z z_I - p_h h_I$$

subject to

$$h_I \geq z_I$$

Input producers employ labor h_I to produce input good z_I. The rate of transformation equals one. Input producers sell all their production in period one so they do not face any financing issue. Each z-unit sold is subject to a tax t_z.

For many parts of our analysis, the input producers are not necessary. But they provide a taxable good used by banks. As we detail upon in Section 1.5, in the context of VAT exemption for banks, this has been associated with problems of tax cascading as banks pay non-deductible taxes on inputs. The introduction of a physical input good allows us to analyze tax-cascading.

The banking problem There is a continuum of measure one of homogeneous banks. They choose $A_B = (d_B, l_B, h_B, z_B, x_B)$ in order to maximize

$$[(1 - t_C)r(l_B)l_B - (1 - t_C)d_B]\, p_x(1 - t_F - t_\pi)$$

subject to

$$(1 - t_C)p_d d_B \geq (1 - t_C)p_l l_B + (1 + t_F)p_h h_B + (1 - t_{zC})p_z z_B \qquad (1.5)$$

$$h_B \geq \kappa_d d_B + \kappa_l l_B \qquad (1.6)$$

$$z_B \geq \gamma_d d_B + \gamma_l l_B \qquad (1.7)$$

$$x_B = [r(l_B)l_B - d_B]\,(1 - t_C)(1 - t_F - t_\pi)$$

for some parameters $\kappa_d, \kappa_l, \gamma_d, \gamma_l \geq 0$.

In period one, banks collect deposits d_B from households, yielding an after-tax cash flow of $(1 - t_C)p_d d_B$. Banks use the cash to grant loans l_B and pay for further inputs; in order to run their business, they need workers h_B and the input good z_B

in a fixed proportion to the size of loans and deposits. Furthermore, banks have to pay a FAT tax t_F on their wage bill, but can claim tax credits for the other two inputs. Notice though, that the level of tax credits t_{zC} does not need to equal the VAT level t_z nor the cash-flow tax level t_C.

In period two, banks receive a repayment from those food producers that succeeded. Recall that food producers success probability is exogenous from the producers' perspective. From the banks' perspective, it is a potentially endogenous object equal to $r(l_B)$, a function of the total number of loans granted by a bank. We will further discuss this issue in a moment.

Banks grant loans to a continuum of food producers, so they are perfectly insured against individual defaults and receive the expected value of their loan portfolio. This ensures that banks are certainly able to pay off depositors in the second period. The cash outflow to depositors allows for a tax credit $t_C p_x d_B$. The cash inflow from debtor is subject to a cash flow tax $t_C p_x r(l_B) l_B$. Finally, banks pay a financial activities tax, t_F, as well as an ordinary profit tax, t_π, on their before-tax profit (but after cash flow taxation). Whatever is left for the bank is used to consume food. We denote their consumption by x_B, and refer to it as *banking surplus.*

(Very) Boring Banking We analyze two versions of the model: *Boring Banking* and *Very Boring Banking*. They differ in the assumptions made upon the return function $R(l_B) := r(l_B) l_B$. Under Very Boring Banking, financial intermediation exhibits constant returns to scale and the financial sector equilibrium is characterized by a plain zero-profit condition. Banks cannot bear any tax burden but pass on all taxes to consumers. In this sense, this version of the model is *very* boring. Under Boring Banking, we still have a financial sector without aggregate risk, and without excessive behavior of any sort. Banks are traditional intermediaries between depositors and borrowers. In this sense, we still talk about *boring* banking. Yet, the intermediation technology yields decreasing marginal return. This induces positive equilibrium profits. Now, banks may bear a tax burden, depending on supply and demand conditions. The tax incidence is thus more 'exiting' than under Very Boring Banking.

Assumption 1.1 (Very Boring Banking). *The return function $R(l_B) = r(l_B) l_B$ exhibits constant returns to scale with $r(l_B) l_B \equiv \alpha l_B$ for some exogenous $\alpha \in (0, 1]$.*

Under Assumption 1.1, α is an exogenous parameter equal to the success probability of food production. As such, it is the success probability as perceived by both food producers and banks. Assumption 1.1 has the implication that banks make zero profits in equilibrium. This is different under Assumption 1.2, stated below, which yields positive profits in equilibrium and allows for banks to bear some of the tax burden. Under Assumption 1.2, the success probability of food production is an endogenous object, determined by the level of granted loans, l_B, (which is both the number of loans granted by an individual bank and the aggregate level of loans in the economy.) From the perspective of the food producer, however, success of production remains exogenous.

Why should the success probability depend on the choices of banks rather than the food producers? In this very distinctness, this is certainly at odds with reality. But within our model, it does have economic support. On the one hand, food producers run fully leveraged production; they do not lose any equity in case of failure. Furthermore, under equilibrium prices, their profit is zero, independent of production success. Thus producers face weak incentives to foster success, if at all. Banks, on the other hand, fully bear the default risk. Their invested capital is lost completely when their loans go into default, and they still owe funds to depositors. Thus banks have high incentives to make sure loans are paid back.

Assumption 1.2 (Boring Banking). *The return function $R(l_B) = r(l_B)l_B$ is twice continuously differentiable, exhibits decreasing returns to scale, and $r(l_B) \in [0, 1]$. For all l_B, it satisfies*

$$r'(l_B) \leq 0, \qquad R'(l_B) \geq 0, \qquad R''(l_B) < 0, \qquad \left| \frac{R'(l_B)}{l_B R''(l_B)} \right| > 1 \,.$$

Assumption 1.2 has three parts. First, $r(l_B)$ is limited to be between 0 and 1, making sure that it can be interpreted as a probability. Second, the higher the loan volume, the lower the average rate of repayment. Third, the last three inequalities concern the shape of $R(l_B)$: higher loan levels always yield higher return but marginal returns diminish. The last inequality is an assumption upon the elasticity of loan supply.[5] An example of a return function satisfying Assumption 1.2 would be $R(l_B) = [\beta_1/(\beta_1 + l_B)]^{\beta_2} l_B$, with $\beta_1 \in (0, 1)$ and $\beta_2 \in (0, 1/2]$.

The reduced-form model of endogenous loan repayment with diminishing returns is quite useful for our analysis; it allows for positive equilibrium profits while remaining in the domain of competitive markets with price taking behavior. The underlying economic reason for diminishing returns could be a moral hazard problem in the spirit of Holmstrom and Tirole (1997): Banks monitor their debtors in order to induce effort. The more loans banks grant, the less effective becomes the monitoring. This could be due a hidden factor, fundamental capacity

[5]We argue in Section 1.3 that the optimal loan supply maximizes $R(l_B) - Ql_B$, with Q being the 'overall cost of loans', as defined in equation (1.8). Then, by standard arguments, $|R'(l_B)/(l_B R''(l_B))|$ is the elasticity of loan supply with respect to 'cost' Q.

constraints or increasing organizational complexity. Another motivation would be an underlying homogeneity in food producers. When granting the first loan, banks pick the best producers with the highest success probability. With increasing loan volume, banks have to compromise on quality.

In what follows, we always presume that either Assumption 1.1 or Assumption 1.2 is true. A large part of the equilibrium analysis in Section 1.3 is true independently of which of the two it is. When a result needs a particular one, we explicitly state the assumption as a premise. On the other hand, when we state a result without referring to one of the two Assumptions, the result is true if Assumption 1.1 or Assumption 1.2 is true.

Public consumption Fix tax rates. Then the tax revenues amount to

$$T_1 := t_h p_h h_H + t_C \left(p_l(l_F - l_B) + p_d d_B\right) + p_z(t_z z_I - t_{zC} z_B) + t_F p_h h_B$$

in period one, and

$$T_2 := t_x p_x r(l_B) h_F + t_C p_x \left((l_B - l_F) r(l_B) - d_B\right) + (t_F + t_\pi)(r(l_B) l_B - d_B) p_x (1 - t_C)$$

in period two. All tax revenues are ultimately used for food consumption and the respective amount is called *public consumption*. In the second period, food markets are open, and the government uses all tax revenues from this very period to buy food, leading to a public food demand of $x_G = T_2/q_x$.

In the first period, no food has been produced yet. So there is no food market. There is a labor market, though. Just as food producers, the government can employ labor (in period 1) to produce food (in period 2). The technology is akin to food producers technology; each unit of labor yields α units of food (non-randomly). We assume that the government does not engage in financial market, so it spends all first-period tax revenues on labor, leading to a public labor demand of $h_G = T_1/p_h$.

Overall, public consumption c_G is thus the sum

$$c_G := x_G + \alpha h_G \,.$$

Public consumption is a measure of public resource use over the two periods. In some sense, it corresponds to the net present value of inter-temporal tax revenues. The analogy is not complete, though, as the government does not engage in credit markets. Notice that public consumption is measured in units of food, rather than tax revenues which are measured in units of cash and hence vary with the overall price level. In the positive part of our analysis, we ask how different tax schemes affect the level of public consumption. In Section 1.4.2, we exogenously fix some public consumption requirement $\bar{c}_G$ and ask which tax schedules (and implied welfare levels) are compatible with $c_G \geq \bar{c}_G$.

In the equilibrium analysis of Section 1.3, we take taxes and government behavior as given and summarize public demand levels by $A_G := (h_G, x_G)$.

Tax systems Throughout the analysis, we consider nonnegative taxes and do not allow any subsidies. Also, we allow a 100% tax rate only for profit tax t_π. Formally, a vector t is a *tax system* if it holds values $(t_h, t_x, t_z, t_C, t_{zC}, t_F, t_\pi)$ for all tax levels and

$$t_i \in [0, 1) \text{ for } i \in \{h, x, z, C, zC, F\}, \qquad t_{zC} \leq t_z,$$
$$t_\pi \in [0, 1], \qquad t_\pi + t_F \leq 1.$$

The inequalities ensure that the total tax on profits is at most 100% and that the tax credit allowed for inputs does not exceed the VAT imposed on the input, which would be a subsidy.

Equilibrium For any given tax system t, we seek to characterize the competitive equilibrium. An *allocation* is a vector $A = (A_H, A_F, A_I, A_B, A_G)$ holding individual choice variables as well government demand. A *price system* $p = (p_h, p_z, p_d, p_l, p_x)$ holds prices for all markets. A tuple (A, p) is an *equilibrium for t* if

- A_H, A_F, A_I, and A_B are individually optimal given p and t,
- $h_G = T_1/p_h$ and $x_G = T_2/p_x$,
- all markets clear:

$$\alpha h_F = x_H + x_G + x_B, \qquad d_H = d_B$$
$$h_H = h_F + h_B + h_I + h_G, \qquad l_F = l_B$$
$$z_I = z_B$$

- and $\alpha = r(l_B)$.

We also say that A is an *equilibrium allocation for t* if a p exist such that (A, p) is an equilibrium for t. Also, p is an *equilibrium price system for t* if A exist such that (A, p) is an equilibrium for t.

1.3 The Competitive Equilibrium

The goal of this section is to characterize the equilibrium allocation for some given tax system. We center our analysis on one particular market: the market for deposits. The reason is that on this market we have the household deposits supply on the one side, and the banking deposit demand on the other side. These two sectors are the main driving forces of the equilibrium allocation, because all other sectors (food and input good production) are mostly determined by zero-profit conditions.

In the results, we focus on the equilibrium allocation rather than equilibrium prices, because the former is (uniquely) determined while the general price level is undetermined. The crucial step in our characterization is the determination of the loan volume l_B. The loan volume is central for an equilibrium because the shape of banking returns R is a direct function of the loan volume and determines the value of banking surplus. We will also show that all economic activity follows the size of the banking sector as measured by l_B.

We begin with the prices for food and input good production. As both sectors exhibit constant returns to scale, the respective prices must satisfy zero profit conditions.

Lemma 1.1. *If p is an equilibrium price system for t, then*

$$p_z(1-t_z) = p_h, \qquad\qquad p_l(1-t_x) = p_h\,.$$

Now, we derive banks' loan supply and deposit demand. Consider the banking problem. It has five choice variables and four constraints. Yet, we can actually simplify the problem significantly. First, the objective function is equivalent to $R(l_B) - l_B$. Second, the constraints, (1.5), (1.6), and (1.7), must be binding. Third, using the findings from Lemma 1.1, we can collapse the three constraint into one single constraint, namely $d_B = Ql_B$, with

$$Q := \frac{\dfrac{(1-t_C)}{(1-t_x)} + (1+t_F)\kappa_l + \gamma_l \dfrac{(1-t_{zC})}{(1-t_z)}}{\dfrac{(1-t_C)p_d}{p_h} - (1+t_F)\kappa_d - \gamma_d \dfrac{(1-t_{zC})}{(1-t_z)}}\,. \tag{1.8}$$

The constraint tells that for each loan a bank wants to lend, it has to collect Q units of deposits. Thus, from the banks' perspective, Q is the overall cost parameter for granting credits; a loan volume l_B returns $R(l_B)$ and, effectively, costs Ql_B. Q accounts for all the inputs and taxes banks have to pay in order to grant loans. Given Q, the banking problem reduces to an unconstrained maximization problem over l_B. The solution has a standard property: marginal returns must equal marginal cost.

Lemma 1.2. *If (A, p) is an equilibrium for t, then*

$$R'(l_B) = Q, \qquad\qquad d_B = l_B Q.$$

To sum up, once we know the equilibrium loan volume, we know that the equilibrium deposit demand is $d_B = l_B R'(l_B)$.

Next, we derive the households' deposit supply. We can reduce the household problem in a fashion similar to the banking problem. Again, the respective constraints, (1.1) and (1.2), must bind. We combine them and find that $h_B = Px_H$, with

$$P := \frac{p_d}{p_h(1 - t_h)} .$$

P is a tax-including relative price. It tells households how much they have to work in order to afford one unit of food. Knowing this relative price is sufficient to determine household choice. We can thus define the *unconstrained household problem* and its solution $x_H^*(P)$ by

$$x_H^*(P) := \underset{x_H}{\text{argmax}} \quad u(x_H, Px_H). \tag{1.9}$$

By construction, the solution of the household problem is given by $x_H = x_H^*(P)$, $h_H = Px_H^*(P)$ and $d_H = x_H^*(P)$. Price ratio P is thus a sufficient statistic for utility. The lower P, the better off the households are because, by an envelope argument,

$$\frac{\partial u(x_H^*(P), Px_H^*(P))}{\partial P} = x_H^*(P)\frac{\partial u}{\partial h_H} < 0 .$$

Households and banks interact on the deposit market. For an equilibrium, deposit supply d_H needs to equal deposit demand d_B. Individually, d_H is a function of price ratio P and d_B is a function of price Q. But we can express both sides of the market as functions of the equilibrium loan volume l_B. Lemma 1.3 below exploits this fact and characterizes the equilibrium loan volume. The lemma uses function $\tilde{P}$. The function gives the equilibrium value of P, given the equilibrium loan supply. It proves crucial for all the remaining analysis, so we present it as a displayed definition.

Definition 1.1. $\tilde{P}$ is a function, mapping from tax systems t and loans volumes l_B

into $\mathbb{R}$, and

$$\tilde{P}(l_B, t) := \frac{1}{(1-t_h)}\left(\frac{1/R'(l_B)}{1-t_x} + \frac{(1+t_F)\kappa(l_B)}{1-t_C} + \frac{\gamma(l_B)}{1-t_C}\frac{1-t_{zC}}{1-t_z}\right)$$

with

$$\kappa(l_B) := \kappa_d + \frac{\kappa_l}{R'(l_B)} \qquad \gamma(l_B) := \gamma_d + \frac{\gamma_l}{R'(l_B)}.$$

Under Assumption 1.1, $R'(l_B) \equiv \alpha$ is an exogenous constant and we will sometimes simplify notation using $\kappa := \kappa_d + \kappa_l/\alpha$, $\gamma := \gamma_d + \gamma_l/\alpha$ and $\tilde{P} = \tilde{P}(t)$.

Lemma 1.3. *If (A, p) is an equilibrium for t then p is such that price ratio P satisfies*

$$P = \tilde{P}(l_B, t),$$

and loan volume l_B satisfies

$$x_H^*(\tilde{P}(l_B, t)) = l_B R'(l_B). \tag{1.10}$$

Equation (1.10) implicitly defines the equilibrium level of loans as a function of taxes t. The equation resembles the market clearing conditions for deposits. The left-hand side is households' deposit supply d_H, equaling their food demand $x_H = x_H^*(P)$. The right-hand side is banks' deposit demand Ql_B: as marginal cost, Q, equal marginal revenues, $R'(l_B)$, we have $l_B R'(l_B)$ as deposit demand. The missing piece in this reasoning, $P = \tilde{P}(l_B, t)$ is also a consequence of previous considerations. Notice first that P is a sufficient statistic for the household choice, and Q is a sufficient condition for bank choice. By substituting p_d/p_h in (1.8) by $(1-t_h)P$, we find that there is a one-to-one relational between Q and P. We already argued that Q must equal marginal returns $R'(l_B)$ and thus there is a one-to-one relation between $R'(l_B)$ and P. When we rearrange this relation, we end up with $P = \tilde{P}(l_B, t)$.

Lemma 1.3 is helpful because it provides the equilibrium value for l_B and, as it turns out, we can formulate the whole allocation as a function of l_B.

Proposition 1.1 (Equilibrium allocation). *Allocation A is an equilibrium allocation for t if and only if*

$$x_H^*(\tilde{P}(l_B, t)) = l_B R'(l_B),$$

and

$$x_H = d_H = d_B = R'(l_B) l_B, \qquad h_H = \tilde{P}(l_B, t) R'(l_B) l_B,$$

$$h_I = z_I = z_B = \gamma(l_B)R'(l_B)l_B\,, \qquad l_F = l_B\,,$$
$$h_B = \kappa(l_B)R'(l_B)l_B\,, \qquad (1-t_x)h_F = (1-t_C)l_B\,,$$

and

$$x_B = \left[R(l_B) - R'(l_B)l_B\right](1-t_C)(1-t_F-t_\pi)\,,$$
$$h_G = \left(\tilde{P}(l_B,t) - \frac{1-t_C}{1-t_x}\frac{1}{R'(l_B)} - \kappa(l_B) - \gamma(l_B)\right)R'(l_B)l_B\,,$$
$$x_G = \left[R(l_B) - R'(l_B)l_B\right](1-t_C)(t_F+t_\pi) + \frac{t_x(1-t_C)}{1-t_x}R(l_B) - t_C R'(l_B)l_B\,.$$

Proposition 1.1 is a consequence of Lemma 1.3. Equation (1.10) gives the equilibrium value of loan supply l_B, and we can express the whole allocation as a function of l_B, using binding optimization constraints and market clearing conditions. Proposition 1.1 holds for both versions of the model (Assumptions 1.2 or 1.1). Under Very Boring Banking it simplifies. In that case, (1.10) reads

$$x_H^*(\tilde{P}(t)) = l_B\alpha.$$

The left-hand side of the equation is independent of l_B but fully determined by the tax system which induces $\tilde{P}(t)$. Also, α is exogenous under Assumption 1.1, so the equilibrium level of loans simply is $l_B = x_H^*(\tilde{P}(t))/\alpha$. Before-tax banking surplus then amounts to $[R(l_B) - R'(l_B)l_B] = [\alpha l_B - \alpha l_B] = 0$, because constant returns scale technologies drive profits down to zero. Under Boring Banking, on the other hand, profits are positive and banking surplus is

$$\left[R(l_B) - R'(l_B)l_B\right](1-t_C)(1-t_F-t_\pi) = [-2l_B r'(l_B)](1-t_C)(1-t_F-t_\pi) \geq 0\,.$$

Banking surplus is positive because, by Assumption, $r' < 0$.

Under Assumption 1.1, there is exactly one l_B solving equation (1.10). So there exist a unique equilibrium allocation for Very Boring Banking. We conclude this section by providing conditions for existence and uniqueness which also cover the case of Boring Banking.

Lemma 1.4 (Uniqueness). *For any given t there is at most one equilibrium allocation A if one of the following conditions is met:*

- *Assumption 1.1 holds, or*
- *Assumption 1.2 holds and $x_H^*(P)$ is monotonically decreasing in P.*

Lemma 1.5 (Existence). *Fix some tax system t and suppose that the following conditions are met:*

i) x_H^* *is a continuous.*

ii) $R'(0)$ *is bounded.*

iii) There exists $l_B > 0$ *such that* $x_H^*(\tilde{P}(l_B, t)) \leq R'(l_B) l_B$.

Then, there is at least one equilibrium allocation for t.

1.4 Taxation

In this section, we discuss two questions. 1. Which tax systems are equivalent? Textbook-style general equilibrium models with households and CRS producers typically find that a labor tax is equivalent to a uniform commodity tax. Our model features two crucial differences: in addition to producers and households, we have financial intermediaries, and we have diminishing returns to scale. 2. How to set tax rates in order to raise some given amount of tax revenues? How can one distribute the required tax burden?

1.4.1 Equivalent Tax Systems

When are two tax systems equivalent? A restrictive answer would require tax systems to be equivalent when they induce the very same allocation. From a normative perspective, however, the relevant object is not the allocation as such, but the surplus induced by the allocation. Given that input good producers and food producers receive zero surplus in any equilibrium, we can describe payoffs by the following three values. 1. household utility u, 2. banking surplus x_B, and 3. public consumption c_G.

Consequently, we consider two tax systems *equivalent* if they induce the same household utility u, the same banking surplus x_B, and the same public consumption c_G.

Proposition 1.2. *1. Under Very Boring Banking (Assumption 1.1), two tax systems, t and t′, are equivalent if*

$$\tilde{P}(t) = \tilde{P}(t') .$$

2. Under Boring Banking (Assumption 1.2), two tax systems, t and t′, are equivalent if

$$(1 - t_C)(1 - t_F - t_\pi) = (1 - t'_C)(1 - t'_F - t'_\pi)$$

$$\text{and, for all } l_B,\ \tilde{P}(l_B, t) = \tilde{P}(l_B, t') .$$

Proposition 1.2 yields some noteworthy insights. Under Very Boring Banking, the result is particularly strong; two tax systems are equivalent whenever they induce the same price ratio P. The crucial step in the proof is to realize that once the equilibrium price ratio P is determined, the equilibrium values of household utility, banking surplus, and public consumption follow from P, irrespectively of individual tax rates underlying P. This finding is rather obvious for household utility given that the solution to the unconstrained household problem is a function of P, as defined in (1.9). Banking surplus is straightforward as well, because it is zero in any case. For public consumption, the finding requires some more algebraic effort. For details, see Lemma 1.10 in the appendix.

The induced price ratio under Very Boring Banking is

$$\tilde{P}(t) = \frac{1}{(1-t_h)}\left(\frac{1/\alpha}{1-t_x} + \frac{(1+t_F)\kappa}{1-t_C} + \frac{\gamma}{1-t_C}\frac{1-t_{zC}}{1-t_z}\right) \tag{1.11}$$

with $\kappa = \kappa_d + \kappa_l/\alpha$, $\gamma = \gamma_d + \gamma_l/\alpha$, for some exogenous parameter α. Hence, for any given P, there are many tax systems inducing it and thus there are many equivalent tax systems. For instance, a pure labor tax system is equivalent to a broad (universal and uniform) value-added-tax system.[6]

Proposition 1.2 is reminiscent to the findings of Auerbach and Gordon (2002), who also show equivalence of a broad VAT and a labor-income tax. They conclude that therefore the VAT should include all (financial) intermediation. This conclusion, however, is not confirmed in out model. To see why, suppose there is a VAT system which includes sales of 'real' goods ($t_x = t_z > 0$), but exempts financial transactions ($t_C = 0$). Let all other tax rates be zero as well. Then the induced price ratio is $P = 1/((1-t_x)\alpha) + \kappa + \gamma/(1-t_z)$, while a pure labor tax would induce $P = (1/\alpha + \kappa + \gamma)/(1-t_h)$. Yet, by fine-tuning tax rates t_x and t_z, it is possible to induce a price ratio that equals the price ratio under the pure labor tax, making the two tax system equivalent. It is thus not necessary to include financial transactions into the VAT scheme, when the goal is replication of a labor-income tax.

The case of Boring Banking is more intricate. In contrast to Very Boring Banking, banks generate a positive surplus in equilibrium due to diminishing returns to scale. Taxes affect both the pre-tax and the after-tax profit. The last condition, $\tilde{P}(l_B, t) = \tilde{P}(l_B, t')$, guarantees that, similar to Very Boring Banking, the equilibrium price ratio P is identical, yielding identical household utility. It also guarantees identical loan volumes l_B, implying identical before-tax banking profits. The second-to-last condition, $(1-t_C)(1-t_F-t_\pi) = (1-t'_C)(1-t'_F-t'_\pi)$, ensures

[6] Such a "universal" and "uniform" value added tax system would comprise $t_x = t_z = t_C = t_{zC} > 0$ and $t_F = t_h = t_\pi = 0$.

that after-tax profits are equal as well. The equivalent tax system must yield both the same level of economic activity and the same distribution of financial sector surplus.

Therefore, the equivalence results from Auerbach and Gordon (ibid.) and from the Very Boring Banking case change when returns to scale are diminishing; now, a pure labor-income tax is *not* equivalent to a universal and uniform VAT. Such two tax system do induce the same level of economic activity, i.e., the same l_B. But the broad VAT includes a cash-flow tax that cuts into banks' profits, while the labor tax does not.

To retain equivalence, the income tax must be accompanied by a profit tax. This is the case in the following two systems:

1. A uniform tax on labor income and profits: $t_h = t_\pi > 0$, and all other taxes zero.
2. A uniform tax on value added, imposed on all transactions except labor income: $t_x = t_z = t_C = t_{zC} > 0$, and all other taxes zero.

These two tax system are equivalent when the respective uniform tax rates are the same.

1.4.2 Taxation with a Fixed Public Budget Constraint

Suppose tax authorities have to raise some exogenously fixed amount of tax revenues and seek to find suitable instruments. This classical question of tax theory is the subject of this section. As mentioned in the model description, the public budget constraint takes the form

$$c_G \geq \bar{c}_G \,,$$

where $\bar{c}_G \geq 0$ is an exogenous public consumption requirement. Recall that public consumption is the sum $c_G = x_G + \alpha h_G$. Tax revenues must suffice to cover the required expenditures. But again we abstract from revenues and focus directly on public consumption, which is independent of the general price level.

Typically there are many feasible tax systems respecting the public consumption requirement. A feasible tax systems is *efficient* if any other tax system which increases either household utility or banking surplus induces a reduction in the other one or violates the public consumption requirement.

In the case of Very Boring Banking, banking surplus is zero for any tax system, so taxes are efficient if they maximize household utility subject to the public budget constraint. If we restrict parameters such that household demand is relatively inelastic, then there is a straightforward result.

Assumption 1.3. *The (absolute) elasticity ϵ of x_H^*, with respect to relative price P, is at most* 1*, that is, for all P,*

$$\epsilon = \left| \frac{\partial x_H^*}{\partial P} \frac{P}{x_H^*} \right| \leq 1 \,.$$

Assumption 1.3 ensures that higher tax rates always yield higher revenues as the households' behavioral response to tax changes is not too large. Under this assumption, we obtain the following characterization.

Proposition 1.3. *Let Assumptions 1.1 (Very Boring Banking) and 1.3 be true. Then t is an efficient tax system given public consumption requirement $\bar{c}_G$ if and only if*

$$\alpha \left(\tilde{P}(t) - (1/\alpha + \gamma + \kappa) \right) x_H^*(\tilde{P}(t)) = \bar{c}_G \,.$$

The crucial step in the proof is to realize that once we know the equilibrium price ratio P^*, we know that public consumption is

$$c_G = \alpha \left(P^* - (1/\alpha + \gamma + \kappa) \right) x_H^*(P^*) \,, \tag{1.12}$$

irrespectively of the individual tax rates (see Lemma 1.10 in the appendix). So in order to fund public consumption $\bar{c}_G$, the tax designer first needs to find a respective price ratio using (1.12). Under Assumption 1.3, this problem has a unique solution.[7] Once the desired price ratio is found, the tax designer needs to set tax rates in order to induce that price ratio. This second task is characterized by equation (1.11) and has multiple solutions; *all* tax systems with $\tilde{P}(t) = P^*$ yield the desired outcome. There is a high degree of freedom in the choice of the tax base. More to the point, in this version of the model it virtually makes no difference whether the financial sector is included or excluded from taxation.

Proposition 1.3 is only valid for Very Boring Banking. We now turn to the more intricate exercise when banking is just boring.

Boring Banking and the Pareto Frontier

In the case of diminishing marginal returns (Boring Banking), positive banking profits enter the welfare considerations. The question who should fund the state now becomes a substantial one. Which tax instruments are efficient? Does the tax designer have a choice between burdening the banks and burdening the households? Is there an equity-efficiency trade off?

We approach this exercise by posing the following technical question. Which combinations of banking surplus x_B and household utility u are compatible with a

[7] This is true because the RHS of (1.12) is strictly increasing. For details, see the appendix.

public consumption requirement of $\bar{c}_G$ and do not involve any waste of resources? Put differently, which tuples (u, x_B) are Pareto-efficient, given some public budget constraint? When determining these tuples, we account for the fact that there must exist a tax system able to implement it. The answer is a particular type of a Pareto frontier, as defined later in the text.

We can simplify the problem by reducing the considered tax systems. An *earnings tax system* is a tax system where all tax rates are zero except the labor tax t_h and the profit tax t_π. With some abuse of notation, we denote an earnings tax system by (t_h, t_π). Earnings taxes are useful because its two instruments, t_h and t_π, are capable of implementing any possible payoff distribution.

Lemma 1.6. *If tax system t induces public consumption c_G, household utility u, and banking surplus x_B, then there exist an earnings tax system which induces the very same values of c_G, u, and x_B.*

Lemma 1.6 allows us to restrict the following analysis to earnings tax system without loss of generality, because considering more tax instruments would not improve payoffs.

When we assume household demand x_H^* to be strictly decreasing, we can further simplify the problem: For then there is a one-to-one relation between household utility u and loan volume l_B; the higher l_B, the higher is household utility.

Lemma 1.7. *Suppose x_H^* is strictly increasing. Let A and A′ be distinct equilibrium allocations. Let (x_H, h_H, l_B) be part of A, and let (x'_H, h'_H, l'_B) be part of A′. Then*

$$u(x_H, h_H) > u(x'_H, h'_H) \;\Leftrightarrow\; l_B > l'_B \,.$$

Different tax systems give rise to different market prices. Households are well-off when prices are in their favor. Under such prices, households typically consume more and work more.[8] High economic activity by households is only possible with high economy-wide activity. In particular, high loan volumes are needed to fund high production levels. Analogously, a high loan volume can only be an equilibrium value if households are willing to supply many deposits. Typically, households do so only when prices are in their favor, implying high equilibrium household utility.

Recall that a tax system is efficient if no other feasible tax system dominates it in terms of banking surplus and household utility. To illustrate the set of efficient outcomes, we could draw a picture in the u–x_B space depicting all tuples (u, x_B) which correspond to an efficient tax system. But instead we consider the l_B–x_B

[8] With large income effects, this might not be true, which is why we make the assumption of increasing food demand x_H^*.

space and depict all tuples (l_B, x_B) which correspond to an efficient tax system. By Lemma 1.7, the two representations are equivalent. Defining the set of efficient outcomes in the l_B–x_B space rather than the u–x_B is both convenient and appealing. It is convenient because Proposition 1.1 already provides values for both l_B and x_B, as well as their relation. It is appealing because in practice the loan volume l_B is an observable, measurable item, whereas u is not.

Definition 1.2. (l_B, x_B) is an element of the *Pareto frontier for* $\bar{c}_G$, if

- there exist earnings taxes $t_h \in [0, 1)$ and $t_\pi \in [0, 1]$ and public consumption c_G such that l_B, x_B and c_G are part of the equilibrium allocation for (t_h, t_π) and $c_G \geq \bar{c}_G$, and
- there exist no triple (l'_B, x'_B, c'_G) such that
  - $l'_B \geq l_B$ and $x'_B \geq x_B$, with at least one strict inequality,
  - $c'_G \geq \bar{c}_G$, and
  - l'_B, x'_B, and c'_G are part of an equilibrium allocation for some $t_h \in [0, 1)$ and $t_\pi \in [0, 1]$.

Definition 1.2 describes the *Pareto frontier.* It contains all efficient combinations of loan volume and banking surplus for a fixed public budget constraint $\bar{c}_G$. Every point in the set is associated with a particular equilibrium allocation and a corresponding tax system implementing it. Each of these allocations gives rise to a particular value of household utility. By Lemma 1.7, the ranking over household utility levels is equivalent to the ranking over loan volumes. Therefore, an earnings tax system is efficient if and only if the induced allocation has a corresponding element in the Pareto frontier.

As argued, households prefer higher loan volumes; banks prefer higher banking surplus. Hence, when it is possible to increase both at the same time, the tuple cannot be efficient. The following proposition formalizes and confirms this intuition.

Proposition 1.4. *Consider Boring Banking (Assumption 1.2 is true) and suppose* x^*_H *is decreasing. If*

- (x_B, l_B) *and* (x'_B, l'_B) *are elements of the Pareto frontier, with*
- $x_B > x'_B$ *, and*
- (t_h, t_π) *and* (t'_h, t'_π) *are the two corresponding earnings tax systems,*

then

$$l'_B > l_B\,, \qquad t'_h < t_h\,, \qquad t'_\pi > t_\pi\,.$$

Given Lemma 1.7 and the definition of the Pareto frontier, Proposition 1.4 is quite an obvious result. But it highlights a less obvious finding: Banks prefer lower loan volumes rather than higher ones. Although loan volumes measure the size of banking activity and the banks' before-tax profits are indeed increasing in the loan volume, banking surplus decreases in the loan volume. The reason is that higher loan volumes require more household deposit supply. To induce such a supply, the labor tax must be lowered. To compensate for the lower labor tax revenues, the profit tax needs to increase. While before-tax profits increase, after-tax profits decrease, reducing banking surplus.

Proposition 1.4 presumes that a Pareto frontier exists and contains at least two distinct elements. It lacks a full characterization of the frontier. To illustrate that the Pareto frontier typically does exist and to picture its shape, we employ a computationally simple example. It allows us to characterize the whole set quite explicitly. Also, the individual and aggregate effects of tax rates become more visible. The example is based on the following simplifying assumption about utility function u and technology parameters κ_d and γ_d.

Assumption 1.4. $\kappa_d = \gamma_d = 0$ *and* $u = \ln x - h$.

Under Assumption 1.4, utility u is a standard example of quasi-linearity. The intermediation technology is simplified but not trivialized; Recall that κ_d and γ_d correspond to the required inputs per deposit unit. While theses are zero, κ_l and γ_l are still positive. So financial intermediation still requires resources proportional to the loan volume.

Assumption 1.4 is useful because it yields a closed-form solution for the equilibrium loan volume. In fact, under Boring Banking (Assumption 1.2), Assumption 1.4, and a restriction to earnings taxes, Proposition 1.1 simplifies substantially, and we find

$$c_G = t_\pi \left[R(l_B) - l_B R'(l_B)\right] + t_h r(l_B)\,, \tag{1.13}$$

$$x_B = \left[R(l_B) - l_B R'(l_B)\right](1 - t_\pi)\,, \tag{1.14}$$

$$l_B = \frac{1 - t_h}{1 + \kappa_l + \gamma_l}\,. \tag{1.15}$$

Now we can determine the effects of taxes on c_G, x_B, l_B and on household utility u. The loan volume is strictly decreasing in t_h and is unaffected from t_π. Due to the one-to-one relation, the very same qualitative comparative static holds for household utility. This is quite expected; income taxes discourage labor supply and economic activity as a whole. Profit taxes do not change individual choices.

Banking surplus is obviously strictly decreasing in t_π. It also decreases in t_h because l_B decreases in t_h and before-tax profits, $[R(l_B) - l_B R'(l_B)]$, increase in l_B.[9] Public consumption, on the other hand, is increasing in t_π. The effect of t_h on public consumption might be ambiguous. There is a positive direct effect from higher labor tax revenues, and there is an indirect negative effect: more labor taxes reduce economic activity and reduce before-tax profits. Thus they reduce the tax base of the profit tax which reduces profit-tax revenues. Some algebraic effort shows that Assumption 1.2 ensures that the direct effect dominates and public consumption is indeed increasing in t_h.[10]

To describe the Pareto frontier, we need to know the effects of a *simultaneous* change in taxes conducted in a way to keep public consumption constant. As a first step, we derive a function $l_B \mapsto x_B$, which maps from loan volumes to the highest possible banking surplus. The public budget requirement $\bar{c}_G$ enters the function as a parameter. In a second step, we find the feasible domain of the function. Figure 1.1 illustrates this function and the Pareto frontier.

To derive the function, suppose l_B is part of the Pareto frontier and fix its value. Then, by (1.15), the labor tax is $t_h = 1 - \varphi l_B$, with $\varphi := (1 + \kappa_l + \gamma_l)$. The public budget constraint, $c_G \geq \bar{c}_G$, need to hold with equality.[11] Using (1.13) and substituting for t_h, we find

$$\bar{c}_G = t_\pi \left[R(l_B) - l_B R'(l_B)\right] + (1 - \varphi l_B) r(l_B)\,.$$

We combine this equation with (1.14) to obtain

$$x_B = \left[R(l_B) - l_B R'(l_B)\right] - \bar{c}_G + (1 - \varphi l_B) r(l_B)\,. \tag{1.16}$$

Suppose that the banking surplus determined in this way is feasible. Then equation (1.16) gives the equilibrium value of banking surplus x_B for any given l_B under the condition $c_G = \bar{c}_G$. If (l_B, x_B) is indeed part of the Pareto frontier then x_B must be decreasing at that l_B. Otherwise a Pareto-improvement is possible by increasing l_B. So, how does x_B change in l_B? Some straightforward algebra, combined

[9] To see why take the first derivatives to obtain $-l_B R''(l_B)$, which is positive under Assumption 1.2.

[10] In (1.13), take the derivative with respect to t_h, accounting for (1.15). This results in $\partial c_G / \partial t_h = l_B R'' t_\pi / \varphi + r - t_h r' / \varphi$. the last term is positive as $r' \leq 0$. The second term, r, is positive and greater than R'. The first term is negative and $l_B R'' t_\pi / \varphi = -|l_B R''| t_\pi / \varphi > -|l_B R''|$. Hence we also know $l_B R'' t_\pi / \varphi + r > -|l_B R''| + R$. Finally, from Assumption 1.2, we know $R' - |l_B R''| > 0$.

[11] To see why, suppose there was slack in the public budget. If $t_\pi > 0$, then a small reduction in t_π could increase x_G without violating the public budget or changing l_B. If $t_\pi = 0$, we cannot reduce it further. Then, however, a small reduction in t_h would increase both l_B and x_B without violating the public budget. (For $t_h = t_\pi = 0$ the public budget cannot be slack.)

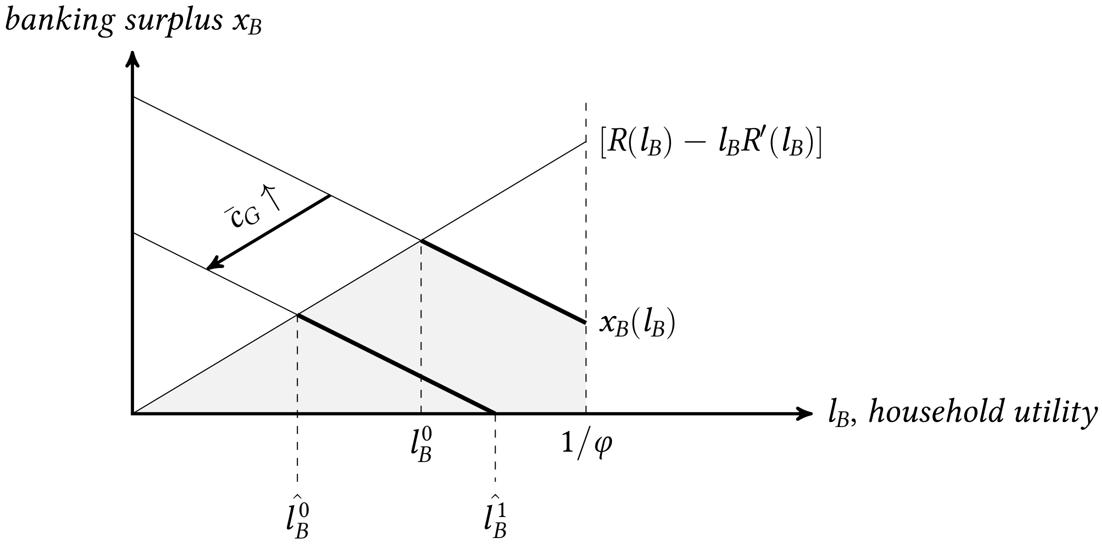

Figure 1.1: The grey area illustrates the feasible combinations of loan volume l_B and banking surplus x_G, for given public budget requirement $\bar{c}_G$. Banking surplus x_B must be between zero and before-tax profits $[R(l_B) - l_B R'(l_B)]$; loan volume l_B must be below $1/\varphi$. Otherwise, positive taxes cannot implement the allocation. The thick lines depict the Pareto frontier for different values of $\bar{c}_G$. If $\bar{c}_G$ is low, the efficient loan levels range from l_B^0 to $1/\varphi$. If $\bar{c}_G$ is high, they range from $\hat{l}_B^0$ to $\hat{l}_B^1$. Notice that the actual functions typically are not linear.

with Assumption 1.2, shows that

$$\frac{\partial x_B}{\partial l_B} = -l_B R'' - \varphi r + r' - \varphi r' l_B = r' - \varphi R' - l_B R'' < -R' - l_B R'' < 0\,.$$

Banking surplus is indeed decreasing in the loan volume. This is good news for our analysis but is also surprising; typically, one would expect that banks are well-off when the financial sector is large. But the opposite is true. A large financial sector comes with low banking surplus.

A higher loan volume is equivalent to a lower labor tax rate. With our assumptions on elasticities, a lower labor tax rate unambiguously leads to lower labor tax revenues. This means that profit taxes have to be increased to balance the public budget. Although higher loan volumes induce higher before-tax profits, the after-tax profits (i.e., the banking surplus) decreases due to the increase in profit taxes. Consequently, there is a trade-off between banking surplus and household utility; one can either have a regime with low labor taxes, a large financial sector, and high profit taxes, or a regime with high labor taxes, a small financial sector, and low profit taxes. The former benefits the households; the latter benefits the banks.

So far, we know that if (l_B, x_B) is on the Pareto frontier, then it satisfies (1.16). To complete the characterization of the Pareto frontier, we need to check which of these tuples (l_B, x_B) are implementable with an earnings tax system, taking account of the fact that we restrict attention to positive tax rates.

First, as $t_h = 1 - \varphi l_B$ and $t_h \in [0, 1)$, we need $l_B \in (0, 1/\varphi]$. Second, banking surplus must be nonnegative. Let l_B^1 be the loan volume such that banking surplus, as determined in (1.16), is zero, i.e.,

$$\left[R(l_B^1) - l_B^1 R'(l_B^1)\right] - \bar{c}_G + (1 - \varphi l_B^1) r(l_B^1) = 0 \,.$$

For $l_B \leq 1/\varphi$, such an l_B^1 exist whenever $\bar{c}_G > [R(1/\varphi) - (1/\varphi)R'(1/\varphi)]$. As banking surplus is decreasing in l_B, l_B^1 is an upper bound for the loan volume.

Next, if

$$\lim_{l_B \to 0} \left(\left[R(l_B) - l_B R'(l_B)\right] - \bar{c}_G + (1 - \varphi l_B) r(l_B)\right) < 0 \,,$$

implying $\lim_{l_B \to 0} r(l_B) < \bar{c}_G$, then there is no loan volume such that the public budget requirement is compatible with positive banking surplus. In this case, $\bar{c}_G$ is simply too high to be feasible.

As we restrict t_π to be nonnegative, there cannot be a subsidy to banks. This restriction becomes binding when the revenues from labor taxation suffice to cover the public consumption requirement. Labor taxes contribute an amount of $(1 - \varphi l_B) r(l_B)$ towards public consumption.[12] It is straightforward to show that this value decreases in l_B. Thus we have a lower bound on loan volumes, l_B^0, defined by[13]

$$(1 - l_B^0 \varphi) r(l_B^0) = \bar{c}_G \,.$$

We can now summarize:

Proposition 1.5. *Consider Boring Banking and let Assumption 1.4 be true. Then (l_B, x_B) is an element of the Pareto frontier for $\bar{c}_G$, if*

$$\begin{aligned} &l_B \in \left(l_B^0, \bar{l}_B\right] \\ &x_B = \left[R(l_B) - l_B R'(l_B)\right] - \bar{c}_G + (1 - \varphi l_B) r(l_B) \,, \end{aligned}$$

[12] Labor tax revenues are $t_h p_h h_h$ and amount to $t_h p_h PR'(l_B) l_B$ in equilibrium. These revenues allow the government to employ $t_h PR'(l_B) l_B$ units of labor in period one, producing $r(l_B) \cdot t_h PR'(l_B) l_B$ units of food toward the public budget in period two. Next, equilibrium condition (1.10) reads $1/P = l_B R'(l_B)$. So the food produced through labor taxation is $t_h r(l_B) = (1 - \varphi l_B) r(l_B)$.

[13] Notice that such an l_B^0 does indeed exist. We have already argued that we require $\bar{c}_G < \lim_{l_B \to 0} r(l_B) = \lim_{l_B \to 0} (1 - l_B \varphi) r(l_B)$. Also, $\lim_{l_B \to 1/\varphi} (1 - l_B \varphi) r(l_B) = 0$.

where the boundaries, l_B^0 and $\bar{l}_B$, are (implicitly) defined by

$$\bar{c}_G = (1 - l_B^0\varphi)r(l_B^0),$$
$$\bar{l}_B = \min\{l_B^1, 1/\varphi\}, \qquad \bar{c}_G = \left[R(l_B^1) - l_B^1 R'(l_B^1)\right] + (1 - \varphi l_B^1)r(l_B^1)\,.$$

Proof. See main text. □

Tax Rates at the Pareto frontier

What does Proposition 1.5 imply for tax rates? If we move along the Pareto frontier, outlined in Figure 1.1, profit and labor taxes move simultaneously: one decreases, the other increases. But what happens in the extreme cases, in which the tax designer maximizes either household utility or banking?

First, let household utility be the objective function. Then we need to pick the largest possible loan volume. If $\bar{c}_G$ is low, the maximal loan volume equals $1/\varphi$. This corresponds to a labor tax rate of zero; all revenues are raised through profit taxation. Typically, profit taxes are less than 100% (and banks enjoy a positive surplus), because the public budget only needs a certain amount of revenues. If the public budget requirement increases, the profit tax rate increases accordingly. This is feasible up to a profit tax of 100%. To further sustain public budget balance, labor taxes have to be raised in addition to a 100% profit tax. In this case, the optimal labor tax rate induces a loan volume of l_B^1 as defined above. The following corollary summarizes these considerations:

Corollary 1.1. *Consider the case of Boring Banking and let Assumption 1.4 be true. Let (t_h^*, t_π^*) be the earnings tax system that maximizes household utility for a given level $\bar{c}_G < r(0)$ of public consumption. Then the following holds:*

- *If* $\bar{c}_G \leq [R(1/\varphi) - R'(1/\varphi)/\varphi]$, *then*

$$t_\pi^* = \bar{c}_G/[R(1/\varphi) - R'(1/\varphi)/\varphi]$$
$$t_h^* = 0\,.$$

- *If* $\bar{c}_G > [R(1/\varphi) - R'(1/\varphi)/\varphi]$, *then*

$$t_\pi^* = 1$$
$$t_h^* = (1 - \varphi l_B^1)\,,$$

with $\left[R(l_B^1) - l_B^1 R'(l_B^1)\right] + (1 - \varphi l_B^1)r(l_B^1) = \bar{c}_G$.

Proof. This is a corollary to Proposition 1.5, as detailed in the main text. □

Second, we consider the other extreme: maximizing banking surplus. Now we need to pick the highest x_B from the Pareto frontier. This is equivalent to picking the lowest possible loan volume. Again, that is counter-intuitive: the best possible outcome for banks is an allocation where the level of banking activity (measured by the loan volume) is lowest. Naive intuition would suggest the contrary, namely that banks prefer tax systems which lead to high trade volumes. Proposition 1.5 states that the lower bound of loan volumes is l_B^0. So we have the following corollary.

Corollary 1.2. *Consider the case of Boring Banking and let Assumption 1.4 be true. Let* (t_h^*, t_π^*) *be the earnings tax system that maximizes banking surplus for a given level* $\bar{c}_G < r(0)$ *of public consumption. Then the following holds:*

$$\begin{aligned} t_h^* &= (1 - \varphi l_B^0), \quad \textit{where } \ \bar{c}_G = (1 - l_B^0 \varphi) r(l_B^0) \\ t_\pi^* &= 0 \end{aligned}$$

Proof. This is a corollary to Proposition 1.5, as detailed in the main text. □

The optimal tax rates are quite intuitive: when we want banks to be well-off, the profit tax is zero and the labor tax is at a rate such that its revenues exactly cover the public budget. When we want households to be well-off, banks are taxed heavily and labor is taxed only if banking profits alone cannot fund public consumption. When we want banks and household to share the tax burden, both profit and labor taxes are in between the extremes.

Profit and labor taxes not only differ in their distributive impact, though. Something else sets them apart. The labor tax is distortive. It changes individual decision making and reduces total economic activity. The profit tax, on the other hand, is not distortive, as it does not change individual decision making. When the tax designer wants to maximize economic output, the profit tax has an edge over the labor tax. But when the banks' well-being enters the consideration, the potential of the non-distortive tax is not fully exploited. Instead, a distortive instrument – the labor tax – is used to satisfy distributive goals. This result is reminiscent to the classical equity-efficiency trade off (Mirrlees, 1971), where distorting income-taxes might dominate non-distorting lump-sum taxes due to equity concerns. As a further remark, our findings are at odds with conventional claims that a 'business-friendly' policy fosters growth. In our model, a 'business-friendly' policy requires lessening output through distortive taxation.

1.4.3 Discussion

In Section 1.4, we have performed a classical tax exercise. For a set of linear tax instruments, we have provided conditions for the equivalence of tax systems and we have outlined the efficient possibilities of funding state expenditures.

With respect to equivalence, our results are twofold. On the one hand, a labor tax is equivalent to a uniform VAT including financial transactions. (When banks make positive profits, the labor tax must be accompanied by a profit tax.) In principle, this confirms the findings of Auerbach and Gordon (2002). On the other hand, we cannot conclude that financial transactions should or should not be included in a VAT scheme. In particular, when banks make zero profits in equilibrium, the optimal tax base is indeterminate; any desired allocation (of payoffs) can be implemented with many different tax bases, as long as rates are fine-tuned appropriately. This is in contrast to some previous contribution with partial equilibrium analyses, which call for or against particular tax bases (e.g., Auerbach and Gordon, 2002; Boadway and Keen, 2003; Grubert and Mackie, 2000; Jack, 2000).

With respect to efficient tax systems, we have shown that it is sufficient (but not necessary) to use direct taxes (labor and profits) and leave all other commodities untaxed. This result is reminiscent of the classical results of direct vs. indirect taxation by Atkinson and Stiglitz (1976). Furthermore, when banks make positive profits, fiscal authorities face a trade off between burdening the banks and burdening the households. Tax systems favoring the banks induce lower loan volumes, i.e., lower banking activity.

Given our findings, there is a two-step procedure of designing a 'good' tax system for some exogenous public consumption requirement. First, pick some efficient combination of loan volume and banking surplus. The choice depends on the welfare function and determines the equilibrium payoffs of banks and households. It also yields the corresponding earnings tax system. Under Boring Banking the choice is degenerated because only the household can carry the tax burden. Second, choose a particular tax system implementing the chosen allocation. The second choice comprises some degree of freedom. Often, one can choose among different tax systems which are equivalent in terms of economic outcomes.

So far, our analysis has been driven by an interest in the fundamental properties of linear taxes in a model of financial intermediation. But our model is also suited to contribute to the more applied policy discussion on how to reform the VAT system. The discussion is based on presumed shortcomings of current real-world systems and considers various reform proposals. In the next section, we use our framework to analyze the general equilibrium effects of these proposals.

1.5 VAT Reforms Reconsidered

The value added tax treatment is a major topic in both the political and academic discussion of financial sector taxation. While the VAT is one of the largest sources of fiscal revenues, financial services are exempt from VAT in more or less all countries that otherwise employ a VAT (Honohan, 2003, p. 9). The exempt treatment typically is such that banks do not have to collect values added taxes on their services, but in turn are not allowed to recover taxes paid on input goods.[14]

The classical argument for exemption is of a technical nature and based on the lack of a proper tax base. By design, the value added tax applies to the difference of input and output prices. But this difference is not easily observable for financial transactions which typically use margin-based-pricing. The 'buyer' of financial products, be it a household opening a checking account or a firm taking out a loan, 'pays' by accepting some interest rate spread rather than paying a distinct fee. The supply side is similar. Deposits are the main input of financial services, and again, they do not come with a simple price tag. This makes it difficult to identify the actual value added embedded in the transaction. It is particularly difficult when returns are uncertain, because is is hard for fiscal authorities to observe and verify the risk of an individual transaction (Huizinga, 2002). Furthermore, margin-based pricing makes it difficult to properly attribute value added between the two sides of a transaction, namely the lender/depositor who provides fund to the intermediary and the borrower who receives funds from the intermediary. An appropriate allocation of the tax burden is necessary because registered business should receive a tax credit while finals consumers should not. (e.g., de la Feria and Lockwood, 2012; Honohan, 2003; Keen, Krelove, and Norregaard, 2010)

But despite the widespread use of exemption and the technical challenges, the exempt treatment has many critics and there is an ongoing discussion on how to overcome the technical issues. Honohan (2003, p. 10), for instance, describes the exemption as a "historical inheritance without much political or economical rational." In the European Union, it has been a returning subject on the policy agenda since the mid-1990 (de la Feria and Lockwood, 2012). At the G-20 level, it is part of a broader discussion on how to make the financial sector contribute to the cost of public interventions during and after the 2008 financial crisis (International Montary Fund, 2010).

Most concerns fall in one of three groups. First, the exemption might reduce tax revenues. Genser and Winker (1997) estimate that the German fiscal revenue

[14] In practice, some services, like safety box rentals do require VAT collection and do allow for VAT credits on input (Huizinga, 2002). The appropriate devision of overhead cost into recoverable and non-recoverable input is an obvious problem for tax collection in practice. We abstract from this issue.

loss of VAT exemption was 10 billion DM (5.11 billion euro) in 1994, using an approach based on aggregate balance sheet and interest rate data. For the UK, the HM Treasury (2009) reports an estimated loss of £2.8 billion in 2009-10. For the European Union as a whole, Huizinga (2002, Table 5) estimates the revenue potential of including financial services into VAT to be in the range of 9.5 to 15 billion euro, depending on demand elasticity assumptions.

A second concern is about a loss in competitiveness due to the unrecoverable VAT on input goods. In a report for the European Commission, Pricewaterhouse-Coopers (2006, p. 43/369) states that "embedded VAT costs continue to have a clear bearing on the cost efficiency of the EU25 financial services sector." Along these lines, Borselli (2009) writes that the Italian banking industry had to pay 2.2 billion euro in non-recoverable VAT in 2006. Mitigating such tax burdens was among the goals of a 2006 EU directive (Borselli, 2009; de la Feria and Lockwood, 2012).[15]

Third, and most importantly from an economist's perspective, many authors doubt the economic efficiency of the exemption.[16] Keen, Krelove, and Norregaard (2010, p. 119), for instance, consider it a "reasonably benchmark" that "'ideal' policy would comprise a uniform rate of tax on final consumption of all commodities, including financial services".

The most articulated reasoning for inefficiency argues that exemption of financial products makes private use less expensive and business use more expensive as compared to a VAT that includes the financial sector. The argument is as follows: A universal VAT with deductibility would effectively tax the value added at each stage of production. Similar to a sales tax, the final consumer should bear a tax on the total value of the product. The production sector, in contrast, should go untaxed, following the logic of productive efficiency (Diamond and Mirrlees, 1971). But when financial products are exempt, the value which is added in the financial sector goes untaxed. This reduces the consumer price of financial products. Compared to a universal VAT, household products are 'under-priced'.

In the case of business use, the invoice-credit mechanism allows businesses to deduct all VAT paid on its inputs. Thus VAT typically does not enter as a cost in the price calculation. A bank exempted from VAT, however, cannot reclaim VAT paid on inputs and considers it a cost which likely increases the banks output

[15] Council Directive 2006/112/EC of 28 November 2006 on the common system of value added tax.

[16] Zee (2005, p. 82), writes: "Needless to say, this paper takes as given the more conventional view that financial intermediation services are no different from other consumable goods and services, and as such should not be excluded from the base of a broad-based consumption tax like the VAT." Similarly, de la Feria and Lockwood (2012, p. 175) state: "But, nevertheless, the lesson that has been drawn from this academic literature is that inclusion of financial services in VAT is desirable." Some papers conclude differently, though, claiming that at least some finanical services should be excluded for efficiency reasons. Among them are Grubert and Mackie (2000), and Jack (2000).

price. As the bank charges no VAT, a business using bank products cannot deduct any VAT from the respective invoice. The effective business price for financial products is thus higher as compared to an all-including VAT. What is more, the higher price for financial inputs might be reflected in the business's own output price which *is* fully taxable. Insofar as it includes the unrecoverable VAT paid by the bank, tax-cascading can occur. So, overall, the exemption effectively leads to an 'over-taxation' of businesses and an 'under-taxation' of households.[17]

There are at least three prominent reform proposals: the financial activities tax, the cash-flow approach, and zero rating. The three proposals roughly follow the three major complaints against exemption. The financial activities tax (FAT) proposal, put forward by Keen, Krelove, and Norregaard (2010), is part of the IMF report to the G-20 on how to raise a 'fair and substantial contribution' from the financial sector (International Montary Fund, 2010). Zero-rating would allow banks to fully deduct VAT paid on inputs.[18] The cash-flow approach is the most comprehensive reform, aiming to make value added in the financial sector fully taxable, while sustaining compatibility with existing VAT schemes and avoiding the technical difficulties of a financial VAT.

Practical and administrative challenges of these proposals have been discussed extensively (e.g., by Borselli, 2009; de la Feria and Lockwood, 2012; Edgar, 2001; Huizinga, 2002; Poddar and English, 1997; Zee, 2005). But little theoretical insights exist on their general equilibrium implications. With our framework, we contribute to fill this gap. Though our model is simple in many respects and the banking sector is 'boring', we do provide a setting that allows to analyze the effects in terms of real economic payoffs. We consider variations of the following exercise: 1. Start with a tax system that exempts financial transaction. 2. Implement a reform proposal. 3. Compare tax revenues, household utility, and banking surplus before and after the tax reform.

First, we need to define the benchmark (or status quo) tax system. Within our model, exempt treatment means a zero cash-flow tax $t_C = 0$ paired with non-deductibility of input-VAT, i.e., $t_{zC} = 0$. In the simplest version, the value added tax is otherwise uniform and all other taxes are zero. Let t_{VATex} be a uniform VAT rate outside the financial system. Then a tax system t is a *VATex tax system* if

$$t_x = t_z = t_{VATex} > 0 \text{ and } t_C = t_{zC} = t_F = t_\pi = 0 \,.$$

In the following, we compare the proposed reforms to this benchmark.

[17] The over-/under-taxation line of though can be found in the contributions by Huizinga (2002, Box 3), Keen, Krelove, and Norregaard (2010, pp. 119f), Borselli (2009), Keen (2011) and others.

[18] See Huizinga (2002) or Edgar (2001).

1.5.1 Cash-flow Taxation

The idea behind the cash-flow approach is to treat financial transactions in a way analogous to non-financial transactions. Cash-inflows from financial transactions, like receiving a loan or an interest payment, are treated like a sales revenue and are subject to a tax. Cash-outflows, like granting a loan or paying interest, are treated like a cost and allow for a tax credit. The cash flow tax is paired with full deductibility of VAT paid on inputs. The cash-flow approach is fully compatible to the invoice credit method of the usual VAT system. The plain cash-flow approach is often quoted as a method which, despite some practical complications, is able fully to include the financial transactions in the VAT while avoiding the technical difficulties associated with the standard VAT.[19]

In terms of our model, the simplest version of the cash-flow approach consists of a uniform tax, say t_{VATcf}, imposed on all transaction, including financial ones (but excluding labor). Tax credits are generally granted. Formally, a tax system t is a *VATcf tax system* if

$$t_x = t_z = t_{zC} = t_C = t_{VATcf} > 0 \text{ and } t_h = t_F = t_\pi = 0.$$

Within our formal framework the VATcf system is closest to the notion of a uniform commodity taxation including financial transactions. Thus, a comparison of a VATcf system with a VATex system is the best way to track down the effects of exempting financial transactions from the VAT system. The following does such a comparison, holding constant the non-zero tax rates across regimes. So on top of a VATex system with, say, a 20% tax rate, we introduce cash flow taxes (and universal deductibility) with the very same rate of 20%.

Proposition 1.6. *Suppose that x_H^* is strictly increasing. Consider a switch from a VATex tax system with tax rate $t_{VATex} > 0$ to a VATcf tax system with the same rate $t_{VATcf} = t_{VATex}$. Then:*

1. *Household utility u decreases if $\kappa_l + \kappa_d > 0$. It remains constant if $\kappa_l + \kappa_d = 0$.*
2. *Banking surplus x_B decreases under Boring Banking (Assumption 1.1) and remains constant (equal to zero) under Very Boring Banking (Assumption 1.2).*
3. *Public consumption c_G increases under Boring Banking if $\kappa_l + \kappa_d = 0$. Under Very Boring Banking, it remains constant if $\kappa_l + \kappa_d = 0$, and it increases if Assumption 1.3 is true and $\kappa_l + \kappa_d > 0$.*

[19] For the plain cash-flow approach, see, e.g., de la Feria and Lockwood (2012), Huizinga (2002), or International Montary Fund (2010). Poddar and English (1997) propose an extension, the "Truncated Cash-Flow Method with Tax Calculation Account", to deal with some practical complications.

4. *The price for deposits, p_d, increases if $\kappa_l + \kappa_d > 0$, (that is, household earn less interest on their savings). If $\kappa_l + \kappa_d = 0$, the price remains constant.*

5. *The price of business loans, p_l, remains constant.*

Proposition 1.6 implies a couple of noteworthy findings. First, the intuition that exemption and the non-deductibility of inputs leads to 'overpricing' of business loans and 'underpricing' of consumer products does not apply in our setting. The price of business loans is not affected at all by the exemption. The price of deposits does change under some conditions. But when banks need no labor ($\kappa_l = \kappa_d = 0$) it does not. Yet, the original intuition does not refer to labor input but to input goods which are subject to VAT. If the intuition was true, it would need to hold in particular for $\kappa_l = \kappa_d = 0$.

Second, our theory confirms the intuition that a cash flow taxation creates revenues or, equivalently, that the exemption reduces tax revenues. Notice that revenues may be created even if banking operates with a CRS technology implying zero profits; in contrast to a profit tax, a cash flow tax generates revenues even if banking revenues equal costs.

Third, the change in payoffs typically does follow intuition. Banks and households suffer from additional taxation while the government (typically) gains. (This statement is, of course, abstracting from any economy-wide gains of spending tax revenues.)

Proposition 1.6 is valuable mostly from a policy perspective. It answers the question 'what happens when we add cash flow taxation to the VAT system?'. From a normative perspective, it is less valuable. For it gives no indication of the systems' relative efficiency (one generates more tax revenues, one generates more utility). A normatively more relevant question would be 'what happens if we add cash flow taxation to the VAT system in a revenue-neutral way?'. For the case of CRS technology we can provide a straightforward answer.

Lemma 1.8. *Let Assumptions 1.1 and 1.3 be true. Consider a VATcf tax system and a VATex tax system. If both yield the same public consumption c_G, then they give rise to the same utility level u and the same banking surplus $x_G = 0$.*

When Assumptions 1.1 and 1.3 are true, it does not really matter whether or not financial transactions are exempt from VAT or included via a cash flow tax. The introduction of cash flow taxes raises revenues under certain conditions. But simply increasing existing tax rates to raise the additional revenues has the very same effect on economic surplus. So if technical or administrative reasons make it difficult to include financial transaction into the VAT scheme, this version of our model suggest that there is no fundamental problem to dispense with VAT in the financial sector.[20]

Under Boring Banking, it is typically not possible to find a VATex system which would be equivalent to a VATcf system. To see why, suppose the two systems induce the same utility level u. This implies that they also induce the same loan volume l_B.[21] When the loan volume is the same, the before-tax banking profit is also identical. But then the after-tax profits differ; the cash-flow tax cuts into banks' profits, whereas a VATex system does not burden banking surplus (c.f. Proposition 1.1).

1.5.2 Zero-rating

The HM Revenue and Customs of the UK names four types of VAT rates: standard, reduced, exempt, and zero. Goods that are zero-rated are taxable, but their rate is zero per cent. In contrast to exempted goods, the producer of a zero-rated good can reclaim VAT on its inputs.[22] The proposal to zero-rate financial transactions is thus a suggestion to make VAT paid on inputs fully deductible. The main motivation is avoidance of tax cascading.[23]

To evaluate the effects of zero-rating, we consider a tax system with a uniform value added tax excluding financial transactions but allowing full deductibility. More precisely, a tax system t is a *VATzero tax system* if

$$t_x = t_z = t_{zC} > 0 \text{ and } t_C = t_F = t_\pi = 0\,.$$

Proposition 1.7. *Suppose $\gamma_l + \gamma_d > 0$. Consider a switch from a VATex tax system with VAT rate $t_{VATex} > 0$ to a VATzero system with the same VAT rate, i.e., with $t_x = t_z = t_{zC} = t_{VATex}$. If Assumptions 1.1 (Very Boring Banking) and 1.3 are true, household utility increases and public consumption decreases. If Assumption 1.2 (Boring Banking) is true and x_H^* is decreasing, household utility increases and banking surplus increases as well.*

Zero-rating only has any effects if banks use real inputs (which are subject to

[20]Notice, though, that the technical/administrative difficulties mentioned are not addressed in our model. Including them could possibly change our result with respect to the effects of cash flow taxation.

[21]See Section 1.4.2 and Lemma 1.7 in the Appendix for details on the one-to-one relation of u and l_B.

[22]See www.hmrc.gov.uk/vat/forms-rates/rates/rates.htm, retrieved on August 13th, 2013.

[23]Huizinga (2002) puts forward a reform that zero-rates business use of financial services but makes household use fully taxable. This would require banks to verify their clients' status. Huizinga (ibid.) argues that this has become feasible. Nevertheless, we analyze 'pure' zero-rating in order to isolate its very effects.

VAT), so $\gamma_l + \gamma_d > 0$ is the relevant case. Zero-rating then effectively reduces the tax burden by broadening deductibility. If banking technology is CRS, so that banks make zero profits in any equilibrium, the effects are unambiguous: households gain, fiscal authorities lose. Under diminishing returns to scale, both households and banks gain from the lower tax burden. Typically, this comes at the expense of tax revenues. Under certain demand condition, however, pre-reform taxes could be 'on the right-hand side of the Laffer curve', such that reducing taxes would increase tax revenues.[24] Under these circumstances, zero-rating could even raise tax revenues. Yet, if this was the case, the pre-reform VAT rate would be inefficient in the set of VATex tax system. Thus, overall, the effects are in line with naive intuition and confirm the expected effects of the reform. Yet again, this tells nothing about whether such a reform is actually desirable from the perspective of economic efficiency. The following proposition does.

Proposition 1.8. *For any VATzero tax system, there exist a VATex system such that the two systems are equivalent.*

Proposition 1.8 shows that zero-rating cannot induce allocations that would be superior to the exempt treatment. This also means that if there are distortions induced by exemption, zero-rating is not able to remove them. We conclude that zero-rating typically affects households, banks, and tax revenues, in a particular and expected way. These effects might be desirable. Efficiency alone, however, does not provide any reason to zero-rate financial transactions.

1.5.3 The Financial Activities Tax

In its report to the G-20 on financial sector taxation, the IMF addresses the issue of value added tax exemption. Among other concerns, the report notes the risk that the financial sector becomes "unduly large because of its favorable treatment under existing VATs" (International Montary Fund, 2010, p. 20). The report suggests an alternative way to treat value added in the financial sector. The idea is that values added is just the sum of profits and wages. Taxing this sum should thus principally be equivalent to a VAT. Such a tax is called *Financial Activities Tax* or just *FAT*.[25]

While the basic principle – taxing the sum of profits and wages – is quite simple, it becomes less straightforward in an intertemporal setting like ours. Both profits and wage payments are well defined in our frameworks but accrue in different

[24] For CRS technology, we exclude this possibility by means of Assumption 1.3.

[25] For an extensive discussion of the FAT, see Keen, Krelove, and Norregaard (2010). They consider various versions of a FAT with different goals. We only discuss the version which they call FAT1. The aim of FAT1 is to replicate a VAT.

periods, so it is not obvious how to take the sum and when to tax it. In our model, we opt for the approach to tax twice, i.e., in both periods. This makes our analysis more clear-cut.

In order to determine the effects of a FAT, we consider the introduction of a FAT on top of an existing value added tax scheme with exemption. A tax system t is a *FAT tax system* if

$$t_x = t_z > 0 \ , \ \ t_F > 0 \ \text{ and } \ t_C = t_{zC} = t_\pi = 0 \, .$$

Proposition 1.9. *Consider a switch from a VATex tax system with some tax rate $t_{VATex} > 0$ to a FAT tax system with the same value added tax rates $t_x = t_z = t_{VATex}$. Then*

1. *Household utility u decreases if $\kappa_l + \kappa_d > 0$. It remains constant if $\kappa_l + \kappa_d = 0$.*
2. *Banking surplus x_B decreases under Boring Banking and remains constant (equal to zero) under Very Boring Banking.*
3. *Public consumption c_G increases under Boring Banking if $\kappa_l + \kappa_d = 0$. Under Very Boring Banking, it remains constant if $\kappa_l + \kappa_d = 0$, and it increases if Assumption 1.3 is true and $\kappa_l + \kappa_d > 0$.*

When financial intermediation does not require labor input, i.e., $\kappa_l + \kappa_d = 0$, the FAT is effectively the same as a pure profit tax. Then it can only have any effect if equilibrium profits are positive. If so, public revenues increase at the expense of banks.

When financial intermediation does require labor input, i.e., $\kappa_l + \kappa_d > 0$, the FAT induces a wedge on the labor market and creates a distortive burden on households. As a consequence, households reduce labor supply and reach a lower utility level. Banks are only affected if they make positive profits. If they do, the FAT reduces the before-tax profits. Furthermore, banks must now pay a tax on their profits. So banking surplus certainly declines. Public revenues typically increase, but sometimes they may decrease; while the FAT is an additional source of revenues, it also reduces economic activity, thereby reducing the tax base of the value added tax. Under Very Boring Banking, Assumption 1.3 is a sufficient condition for increasing public revenues.

Overall, the comparative static effects are qualitatively identical to the effects following a cash flow tax reform (cf. Proposition 1.6). Hence, our model confirms the intuition that the FAT is an alternative way to target the value added of financial intermediation. FAT and cash flow taxation are not exactly equivalent, though; while they have the same effect on profits, their quantitative impact on equilibrium price ratio P and on individual decisions differ.

1.6 Conclusion

This chapter develops a framework to study the general equilibrium effects of taxation in the financial sectors. It allows for a comprehensive incidence analysis while remaining quite tractable. The banking sector in our setting is a 'boring' one, in the sense that there are no market imperfections or aggregate risks. For instance, neither liquidity nor solvency is ever an issue on the aggregate level. The study of a 'boring' banking sector is a useful first step to enhance our understanding of how financial taxes take effect and who gains and loses from them.

With respect to the tax burden, our model identifies a trade-off. When the state needs to collect a fixed amount of revenues, it can shift the tax burden from households to banks and vice versa. (This result holds for the version of the model in which equilibrium banking profits are positive.) The existence of this trade-off is not very surprising. But it is accompanied with a less intuitive effect: when the tax system makes banks better off, the financial sector shrinks; when it makes households better off, the financial sector expands.

Real-world tax systems typically exempt financial services from value added taxes. Many authors consider this an imperfection and have proposed reforms aimed at including financial services in the VAT scheme. Our framework contributes to this debate by providing general equilibrium insights concerning the reform proposals. On the one hand, we by and large confirm the expected effects. On the other hand, our findings question the necessity of any such reform. In many instanced the reform effects could also be achieved by tweaking the rates of existing tax systems. Our model is too simple to provide for straight policy advice. But it highlights the fact the tax reforms should be based on thorough general equilibrium analysis, because partial insights or plain intuition might be misleading.

The framework proposed in this chapter suggests further research, most notably 'the case of exciting banking'. As banking is 'boring' in the current version of the model, there is no scope for regulation; taxes only serve the goal of collecting revenues. A more 'exciting' version of the financial sector would introduce market imperfections that call for regulatory measures. As our framework allows for a tax incidence, one could analyze the distributional effects of corrective taxes. The financial sector size might be another object of study. It is sometimes argued that the financial sector is too large. If theory supports this claim, there could be an interesting trade-off as the insight from 'boring banking' suggests to enlarge the financial sector in order to make banks contribute more to the public budget.

Appendix 1.A Proofs

1.A.1 Proofs for Section 1.3

Proof of Lemma 1.1. Food producers and input good producers use CRS technologies. By standard arguments, equilibrium prices must imply zero profits. Otherwise supply/demand of these producers would be unbounded or zero. Accounting for the fact that the optimization constraints must bind, we find the stated price relations. □

Proof of Lemma 1.2. Claim 1: $l_B > 0$. $l_B \leq 0$ cannot be an equilibrium because it would imply zero production and zero household consumption. Yet, by assumption on u, some consumption is always optimal for the household, even if prices are high. Claim 2. $l_B < \infty$. $l_B = \infty$ cannot be an equilibrium as it would require unbounded deposit supply. That is never optimal for the household, again by assumption on u. Claims 1 and 2 imply that under equilibrium prices there must be an interior optimum for the banking problem. We know from Lemma 1.1 how p_z and p_l must be relative to p_h. Given these pieces of information and the definition of Q in the main text, $d_B = l_B Q$ follows from the binding constraints of the banking problem. Next, we can reduce the banking problem to $\max_{l_B} [R(l_B) - Ql_B)]$. Any interior solution to this reduced problem satisfies $Q = R'(l_B)$. Consequently, also $d_B = R'(l_B) l_B$. □

Proof of Lemma 1.3. By Lemma 1.2, $Q = R'(l_B)$. In (1.8) we substitute for Q and for $p_d/p_h = (1 - t_h)P$. Rearranging yields the first claim: $P = \tilde{P}(l_B, t)$. Next, if (A, p) is an equilibrium for t, then $d_H = d_B$ (deposit market clears), and $d_H = x_H$ (binding household budget), so $d_B = x_H$. We know from (1.9) that $x_H = x^*_H(P)$; from Lemma 1.2 we know $d_B = l_B R'(l_B)$. Hence, $x^*_H(P) = l_B R'(l_B)$. □

Proof of Proposition 1.1. First, we know from Lemma 1.3 that if A is an equilibrium then $x^*_H(\tilde{P}(l_B, t)) = l_B R'(l_B)$. It is then straightforward to show that market clearing conditions and binding maximization constraints imply the remaining entries of the allocation as stated in the Proposition.

Second, to show that the stated allocation is an equilibrium, we have to find an appropriate price system and check all equilibrium conditions. Let prices p be such that they satisfy the (zero-profit) requirement of Lemma 1.1 and that relative price P equals $\tilde{P}(l_B, t)$. Then:

1. Choices are individually optimal: Household choices are optimal as $x_H = x^*_H(P)$ is optimal by definition, and so are $d_H = x^*_H(P)$ and $h_H = Px^*_H(P)$. Input producer choices are optimal as $h_I = z_I$ is feasible and yields zero profits given prices p. Food producer choices are optimal as $h_F(1 - t_x)/(1 - t_C) = l_F$ is feasible

and yields zero profits given prices p. Bank choices are optimal as $P = \tilde{P}(l_B, t)$ implies $Q = R'(l_B)$ and the latter is the optimality condition for banks' loan supply under the prices of Lemma 1.1. By construction of Q, $d_B = Ql_B$ is optimal and hence, $d_B = R'(l_B)l_B$. It is straightforward to check that z_B, d_B, and x_B then satisfy the respective constraints with equality and thus are optimal as well.

2. To show that government demand (h_G, x_G) corresponds to tax revenues, we use the definitions of T_1 and T_2, plug in the values for prices and the allocation, simplify the expressions, and end up with the stated values.

3. Finally, we need to show that all markets clear. This is trivially true for the deposit, loans and input good markets. Next consider the labor market. Supply is $h_H = \tilde{P}(l_B, t)R'(l_B)l_B$. Demand is

$$\begin{aligned} h_F + h_B + h_I + h_G &= \frac{1 - t_C}{1 - t_x} l_B + \kappa(l_B)R' l_B + \gamma(l_B)R' l_B \\ &\quad + \left(\tilde{P}(l_B, t) - \frac{1 - t_C}{1 - t_x}\frac{1}{R'} - \kappa(l_B) - \gamma(l_B)\right) R' l_B = \tilde{P}(l_B, t)R' l_B = h_H\,. \end{aligned}$$

Now, consider the food market. Food producers employ h_F units of labor, yielding an aggregate food supply of αh_F. Given $\alpha = r(l_B)$ and the stated value of h_F, supply is $r(l_B)l_B(1 - t_C)/(1 - t_x) = R(l_B)(1 - t_C)/(1 - t_x)$. Demand is

$$\begin{aligned} x_H + x_B + x_G &= R' l_B + \left[R - R' l_B\right](1 - t_C)(1 - t_F - t_\pi) \\ &\quad + \left[R - R' l_B\right](1 - t_C)(t_F + t_\pi) + \frac{t_x(1 - t_C)}{1 - t_x} R - t_C R' l_B \\ &= R' l_B \cdot (1 - (1 - t_C) - t_C) + R \cdot \left((1 - t_C) + \frac{t_x(1 - t_C)}{1 - t_x}\right) = R \cdot \frac{1 - t_C}{1 - t_x}\,, \end{aligned}$$

so it does equal supply. □

Proof of Lemma 1.4. Proposition 1.1 implies that any two equilibrium allocations with the same l_B are actually identical (for any given t). Also, if A is an equilibrium allocation for t then it satisfies condition (1.10), $x_H^*(\tilde{P}(l_B, t)) = l_B R'(l_B)$, by Lemma 1.3. Thus it suffices to show that this equation has at least one solution.

First, suppose Assumption 1.1 holds. Then $\tilde{P}(l_B, t) = \tilde{P}(t)$. The left-hand side of (1.10) is $x_H^*(\tilde{P}(t))$ and is constant in l_B. The right-hand side is αl_B, with α exogenous. Hence, there is indeed exactly one l_B such that (1.10) holds.

Second, suppose Assumption 1.2 holds. The Assumption ensures that the RHS of (1.10) is strictly increasing in l_B because

$$\frac{\partial l_B R'(l_B)}{\partial l_B} = R' + l_B R'' = |R'| - |l_B R''| = |l_B R''| \left(\left|\frac{R'}{l_B R''}\right| - 1\right) > 0\,.$$

The LHS of (1.10) decreases in l_B: increasing l_B decreases $R'(l_B)$ which increases $\tilde{P}(l_B, t)$. This, by premise, decreases $x_H^*(\tilde{P}(l_B, t))$. Consequently, LHS and RHS of (1.10) can intersect at most once. □

Proof of Lemma 1.5. The proof consists of two steps.

Claim 1: If (1.10) has a solution $l_B^* > 0$, then there exists an equilibrium allocation. Claim 1 follows from Proposition 1.1, which explicitly states an equilibrium allocation conditional on an l_B satisfying (1.10).

Claim 2: There exist $l_B^* > 0$, such that l_B^* is a solution to (1.10), i.e.,

$$x_H^*(\tilde{P}(l_B^*, t)) = l_B^* R'(l_B^*).$$

Claim 2 is true because (a) for $l_B = 0$ the RHS is $R'(0) \cdot 0 = 0$ if $R'(0)$ is bounded, whereas the LHS is strictly positive due to the assumption that the household problem always has an interior solution. So the RHS is below the LHS for $l_B = 0$. (b) By premise iii), there exist $l_B > 0$ so that $x_H^*(\tilde{P}(l_B, t)) \leq R'(l_B) l_B$, i.e., for some $l_B > 0$ the RHS exceeds the LHS. By continuity, there must exist $l_B^* > 0$ such that LHS and RHS are equal. □

1.A.2 Proofs for Section 1.4

Before proving the results of Section 1.4, we establish two helpful lemmas.

Lemma 1.9. *If (A, p) is an equilibrium for t then the sum of banking surplus and public consumption is*

$$x_B + c_G = R'(l_B)R(l_B)\left[P - \left(\frac{1}{r(l_B)} + \gamma(l_B) + \kappa(l_B)\right)\right]$$

with $P = \tilde{P}(l_B, t)$.

Proof. Lemma 1.9 builds on Proposition 1.1, which states the equilibrium values for x_B, x_G, and h_G. We seek the sum $x_B + c_G = x_B + x_G + r(l_B)h_G$. To simplify notation, we drop some dependencies on l_B. The first part of the sum is

$$\begin{aligned} x_B + x_G &= \left[R - R' l_B\right](1 - t_C) + \frac{t_x(1 - t_C)}{1 - t_x} R - t_C R' l_B \\ &= \frac{1 - t_C}{1 - t_x} R - R' l_B\,. \end{aligned}$$

The second part is

$$r h_G = \left(\tilde{P}(l_B, t) - (1 - t_C)/[(1 - t_x)R'] - \kappa(l_B) - \gamma(l_B)\right) R' R\,,$$

given that $rl_B = R$. We sum up the two parts and rearrange the terms to obtain

$$\begin{aligned} x_B + c_G &= \frac{1-t_C}{1-t_x}R - R'l_B + \tilde{P}(l_B, t)R'R - \frac{1-t_C}{1-t_x}R - (\kappa(l_B) + \gamma(l_B))\, R'R \\ &= \left(\tilde{P}(l_B, t) - \frac{1}{r} - \kappa(l_B) - \gamma(l_B)\right) R'R\,. \end{aligned}$$

Finally, we substitute P for $\tilde{P}(l_B, t)$ (by Lemma 1.3). □

Lemma 1.10. *Let Assumption 1.1 be true. If A is an equilibrium allocation for t then public consumption $c_G = \alpha h_G + x_G$ is given by*

$$c_G = \alpha\,(P - (1/\alpha + \gamma + \kappa))\, x_H^*(P)$$

with $P = \tilde{P}(t)$. If Assumption 1.3 is true as well, then c_G is strictly increasing in P.

Proof. Notice first that $(P - (1/\alpha + \gamma + \kappa))$ is positive for $P = \tilde{P}(t)$ as given in (1.11), because it is straightforward to show that $\tilde{P}(t) \geq (1/\alpha + \gamma + \kappa)$ when the tax system does not allow for subsidies, which we do require. Now, we check the two claims of Lemma 1.10.

Claim 1: In any equilibrium $c_G = \alpha\,(P - (1/\alpha + \gamma + \kappa))\, x_H^*(P)$.
From Lemma 1.9 we know the equilibrium value of the sum $x_B + c_G$. Under Very Boring Banking, banking surplus is zero, i.e., $x_B = 0$. Also, $x_H = R'(l_B)l_B = x_H^*(\tilde{P}(t))$, as $\tilde{P}$ depends on t only. Hence, in the equation stated in Lemma 1.9 we can substitute $x_B = 0$, $R'(l_B)l_B = x_H^*(\tilde{P}(t))$, and $P = \tilde{P}(t)$. This yield the first claim of Lemma 1.10.

Claim 2: c_G is strictly increasing in P.
With some abuse of notation, let $x_H'(P) := \partial x_H^*(P)/\partial P$.
(a) For any P such that $x_H'(P)$ is positive, the claim is trivial.
(b) Now, suppose $x_H'(P) < 0$ and take the derivative of the expression in question.

$$\begin{aligned} \frac{\partial\,(P - (1/\alpha + \gamma + \kappa))\, x_H^*(P)}{\partial P} &= (P - (1/\alpha + \gamma + \kappa))\, x_H'(P) + x_H^*(P) \\ &= x_H^*(P)\left(1 + \frac{P x_H'(P)}{x_H^*(P)}\right) - (1/\alpha + \gamma + \kappa)x_H'(P) > 0 \end{aligned}$$

The inequality is true because, as $x_H'(P) < 0$, the second term is positive, and, under Assumption 1.3,

$$1 + Px_H'(P)/x_H^*(P) = 1 - |Px_H'(P)/x_H^*(P)| \geq 0,$$

so the first term is positive as well. □

Proof of Proposition 1.2. Let $(x_B, l_B, x_H, h_H, c_G)$ and $(x'_B, l'_B, x'_H, h'_H, c'_G)$ be the respective parts of the equilibrium allocations given t and t', resp. We need to show that $x_B = x'_B$, $u(x_H, h_H) = u(x'_H, h'_H)$, and $c_G = c'_G$.

First, let Assumption 1.1 be true. From Proposition 1.1, in equilibrium, $x_B = x'_B = 0$. Also, $u(x_H, h_H) = u(x^*_H(P), Px^*_H(P))$, so $u(x_H, h_H) = u(x'_H, h'_H)$ if $P = \tilde{P}(t) = \tilde{P}(t') = P'$. By Lemma 1.10, $c_G = \big(\alpha(P - \gamma - \kappa) - 1\big)\, x^*_H(P)$ so, again, $c_G = c'_G$ for $P = \tilde{P}(t) = \tilde{P}(t') = P'$.

Second, suppose Assumption 1.2 is true. Loan supply under t must satisfy equilibrium condition (1.10) given t. If $\tilde{P}(l_B, t) \equiv \tilde{P}(l_B, t')$, then, under t', condition (1.10) is identical to the condition under t. As (1.10) does not have multiple solutions (Lemma 1.4), $l_B = l'_B$. By Proposition 1.1, this implies identical household utility levels, as well as identical before-tax banking profit $[R(l_B) - R'(l_B)l_B]$; By Lemma 1.9, it implies that the sum $(x_B + c_G)$ is identical. Banking surplus x_B is $(1 - t_C)(1 - t_F - t_\pi)$ times the before-tax banking profit. Hence, banking surplus is identical given the premise. Finally, public consumption c_G must be identical as both x_B and the sum $(x_G + c_G)$ are identical. □

Proof of Proposition 1.3. A tax system is efficient if it maximizes household utility subject to $c_G \geq \bar{c}_G$. By Lemma 1.10, $c_G = \alpha\,(P - (1/\alpha + \gamma + \kappa))\, x^*_H(P))$, and $P = \tilde{P}(t)$. Thus, any solution must satisfy

$$\alpha\left(\tilde{P}(t) - (1/\alpha + \gamma + \kappa)\right) x^*_H(\tilde{P}(t)) \geq \bar{c}_G\,.$$

Because household utility is strictly decreasing in P, the constraint must be binding (if it was not binding, a change in taxes inducing a small decrease in P would be feasible and would enhance utility.) Hence, t is a solution only if it satisfies the equation in Proposition 1.3.

It remains to show that every t satisfying the equation, is a solution. As argued above, any such t is feasible. By contradiction, suppose t is not a solution to the maximization. Then there is some feasible t' such that $u(t') > u(t)$, where $u(t)$ and $u(t')$ are the respective utility levels induces by the tax system. Then $\tilde{P}(t') < \tilde{P}(t)$. But then $c_G(t') < c_G(t) = \bar{c}_G$, because, by Lemma 1.10, c_G is an increasing function of P. So t' is not feasible, a contradiction. □

Proof of Lemma 1.6. By construction. Fix some t. Let A be the allocation implemented by t. Then l_B satisfies $x^*_H(\tilde{P}(l_B, t)) = l_B R'(l_B)$ (Lemma 1.3). Let (t'_h, t'_π) be an earnings tax system such that

$$\frac{\tilde{P}(l_B, t)}{1/R'(l_B) + \kappa(l_B) + \gamma(l_B)} = \frac{1}{1 - t'_h}$$

This way, $x_H^*(\tilde{P}(l_B, (t'_h, t'_\pi))) = l_B R'(l_B)$, so (t'_h, t'_π) induces the very same l_B and P as the original tax system t. Given Proposition 1.1 this implies identical x_H and h_H, hence identical utility u. Furthermore the sum $x_B + c_G$ is the same under the original system and the earnings tax system (see Lemma 1.9). Now, again by Proposition 1.1, $x_B = [R(l_B) - R'(l_B)l_B](1 - t_C)(1 - t_F - t_\pi)$. Pick t'_π such that $(1 - t_C)(1 - t_F - t_\pi) = (1 - t'_\pi)$ then $x_B = x'_B$ where x'_B is the banking surplus under the earnings tax system. Finally, $c_G = c'_G$ because $x'_B + c'_G = x_B + c_G$ and $x_B = x'_B$. □

Proof of Lemma 1.7. Let P and P' be the equilibrium price ratios for A and A', resp. Then, by Lemma 1.3,

$$x_H^*(P) = l_B R'(l_B)\,, \qquad\qquad x_H^*(P') = l'_B R'(l'_B)\,.$$

The Lemma follows from an examination of the relation of these two equations. First we show that a greater loan volume implies higher utility. If $l_B > l'_B$, then $l_B R'(l_B) > l'_B R'(l'_B)$. Under Very Boring Banking this is obvious because $R' \equiv \alpha$. Under Boring Banking, it follows from the assumption made upon the elasticity of loan supply (for details see the proof of Lemma 1.4). Consequently, $x_H^*(P) > x_H^*(P')$. As x_H^* is strictly increasing, $P < P'$. As argued in the main text, equilibrium household utility is a function of price ratio P and decreases in P (cf. equation (1.9)). Thus, $u(x_H, h_H) > u(x'_H, h'_H)$. Now we show that higher utility imply a higher loan volume. If $u(x_H, h_H) > u(x'_H, h'_H)$, then $P < P'$ because, again, household utility is strictly decreasing in P. Then $x_H^*(P) > x_H^*(P')$, hence $l_B R'(l_B) > l'_B R'(l'_B)$. Finally, $l_B > l'_B$ as $l_B R'(l_B)$ is strictly increasing in l_B. □

Proof of Proposition 1.4. We proof the three claims sequentially.

1. If the claim is false, then $l'_B \leq l_B$. Then (x_B, l_B) Pareto-dominates (x'_B, l'_B), then (x'_B, l'_B) cannot be on the Pareto-frontier, a contradiction.

2. If the claim is false, then $t'_h \geq t_h$. Then $1/(1 - t'_h) \geq 1/(1 - t_h)$. We know from Claim 1 that $l'_B > l_B$, implying $1/R'(l'_B) > 1/R'(l_B)$, given Assumption 1.2. Using Definition 1.1, we find that

$$\frac{1}{(1 - t_h)}\left[\frac{(1 + \kappa_l + \gamma_l)}{R'(l_B)} + \kappa_d + \gamma_d\right] = \tilde{P}(l_B, (t_h, t_\pi)) < \tilde{P}(l'_B, (t'_h, t'_\pi))\,.$$

As we require x_H^* to be decreasing, we have $x_H^*(\tilde{P}(l_B, (t_h, t_\pi))) > x_H^*(\tilde{P}(l'_B, (t'_h, t'_\pi)))$. In any equilibrium, $x_H^* = l_B R'(l_B)$ (Lemma 1.3). The LHS of this equation is increasing given Assumption 1.2 (see the proof of Lemma 1.4). Hence, $l_B > l'_B$. This contradicts Claim 1.

3. We know from Claim 1 that $l'_B > l_B$, implying $[R(l'_B) - R'(l'_B)l'_B] > [R(l_B) - R'(l_B)l_B]$ (to see why take the first derivatives to obtain $-l_B R''(l_B)$, which is positive under Assumption 1.2). Hence, before-tax profits are higher under tax system (t'_h, t'_π). Yet, by premise, banking surplus is higher under system (t_h, t_π). This implies $t'_\pi > t_\pi$ because $x_B = (1 - t_\pi)[R(l_B) - R'(l_B)l_B]$. □

1.A.3 Proofs for Section 1.5

Proof of Proposition 1.6. The proof relies on the function $\tilde{P}(l_B, t)$ and equation (1.10). Let t^{VATcf} be the VATcf tax system, let t^{VATex} be the VATex tax system, and let $t_V = t_{VATcf} = t_{VATex} > 0$ be the respective value added or cash flow tax rate. Then, for all l_B,

$$\begin{aligned} \tilde{P}\left(l_B, t^{VATcf}\right) &= \frac{\kappa(l_B) + \gamma(l_B) + 1/R'(l_B)}{1 - t_V} \\ &\geq \frac{(1 - t_V)\kappa(l_B) + \gamma(l_B) + 1/R'(l_B)}{1 - t_V} = \tilde{P}\left(l_B, t^{VATex}\right), \end{aligned}$$

with a strict inequality if $\kappa(l_B) > 0$ or $\kappa_d + \kappa_l > 0$. On the other hand, we have equality, i.e., $\tilde{P}\left(l_B, t^{VATex}\right) = \tilde{P}\left(l_B, t^{VATcf}\right)$ if $\kappa(l_B) = 0$ or $\kappa_d = \kappa_l = 0$. Let P^{VATcf} and P^{VATex} be the respective equilibrium values of price ratio P, and let l_B^{VATcf}, l_B^{VATex} be the loan supplies in the two tax regimes. Using Lemma 1.3 and Lemma 1.4 (uniqueness), we find

$$\begin{aligned} &\kappa_d = \kappa_l = 0 \Rightarrow P^{VATcf} = P^{VATex},\ l_B^{VATcf} = l_B^{VATex} \\ &\kappa_d + \kappa_l > 0 \Rightarrow P^{VATcf} > P^{VATex} \end{aligned}$$

The former implication is rather obvious; equation (1.10), which determines l_B is identical for both tax systems under the premise, thus $l_B^{VATcf} = l_B^{VATex}$. As a consequence, also $P^{VATcf} = P^{VATex}$. The latter implication is trivial under Assumption 1.1. Under Assumption 1.2, we need to make an argument. Suppose, by contradiction, $P^{VATcf} \leq P^{VATex}$, then $x_H^*(P^{VATcf}) \geq x_H^*(P^{VATex})$, then $l_B^{VATcf} \geq l_B^{VATex}$ (using the fact that the LHS of (1.10) is increasing.). Then $R'(l_B^{VATcf}) \leq R'(l_B^{VATex})$, implying $P^{VATcf} > P^{VATex}$, a contradiction. When P is lower under VATex, the loan volume must be higher (again, see (1.10)). Hence we also have

$$\kappa_d + \kappa_l > 0 \Rightarrow l_B^{VATcf} < l_B^{VATex}.$$

Now we can check the claims sequentially:

1. By an envelope argument, equilibrium household utility is a strictly decreas-

ing function of price ratio P.

2. Under Very Boring Banking, the intermediation technology exhibits CRS, thus banking surplus is zero under any tax system. Under Boring Banking, banks' before-tax profit, $[R(l_B) - R'(l_B)l_B]$, is lower with cash flow taxation. This is because $l_B^{VATcf} \leq l_B^{VATex}$ and before-tax profit are increasing in l_B (as is easy to show for $R'' < 0$). The after-tax profit, i.e, banking surplus x_B, equals $(1 - t_C)$ times the before-tax profit. As the tax factor is $(1 - t_C) = 1$ under VATex but $(1 - t_C) = (1 - t_V) < 1$ under VATcf, banking surplus is strictly lower with cash flow taxation.

3. First, consider Boring Banking and $\kappa_l = \kappa_d = 0$, then l_B and P remain unchanged. So does the sum $x_B + c_G$ (by Lemma 1.9). Banking surplus x_G, however, decreases, as already argued. Hence, c_G must increase. Second, for Very Boring Banking we can refer to Lemma 1.10, which states that the direction of change of c_G equals the direction of change in P.

4. To make a valid price comparison we normalize the general price level by fixing p_h. Then $p_d = Pp_h(1 - t_h)$ by the definition of P. As $t_h = 0$ in both regimes, p_d simply follows the change in P.

5. Again, we fix p_h to normalize the price system. Then, by Lemma 1.1, $p_l = p_h/(1 - t_x)$, which equals $p_h/(1 - t_V)$ under both tax regimes. □

Proof of Lemma 1.8. Let P^{VATcf} and P^{VATex} be the equilibrium values of price ratio P induced by the respective tax regimes. By Lemma 1.10, public consumption is strictly increasing in P, so the premise implies $P^{VATcf} = P^{VATex}$. Then utility must be the same as well. Banking surplus is always zero under Assumption 1.1. □

Proof of Proposition 1.7. The proof relies on the function $\tilde{P}(l_B, t)$ and equation (1.10). Let $t^{VATzero}$ be the VATzero tax system, let t^{VATex} be the VATex tax system, and, to save on notation, let $t_V = t_{VATex} > 0$ be the value added tax rate. Then, for all l_B,

$$\tilde{P}\left(l_B, t^{VATzero}\right) = \frac{(1 - t_V)\kappa(l_B) + (1 - t_V)\gamma(l_B) + 1/R'(l_B)}{1 - t_V}$$
$$< \frac{(1 - t_V)\kappa(l_B) + \gamma(l_B) + 1/R'(l_B)}{1 - t_V} = \tilde{P}\left(l_B, t^{VATex}\right),$$

Let $P^{VATzero}$ and P^{VATex} be the respective equilibrium values of price ratio P, and let $l_B^{VATzero}$, l_B^{VATex} be the loan supplies in the two tax regimes.

Under Assumption 1.1, $P^{VATzero}$ is strictly less than P^{VATex}, independent of l_B. Thus household utility increases due to a more favorable price ratio, and public consumption decreases by Lemma 1.10.

To check the statement for Assumption 1.2, we consider equilibrium condition (1.10). By contradiction, suppose $P^{VATzero} \geq P^{VATex}$, then $x_H^*(P^{VATzero}) \leq x_H^*(P^{VATex})$, so the LHS of (1.10) decreases. Then the RHS must also decrease, implying $l_B^{VATzero} \leq l_B^{VATex}$, because $l_B R'(l_B)$ is increasing in l_B under Assumption 1.2 (see the proof of Lemma 1.4). When we plug in $l_B^{VATzero} \leq l_B^{VATex}$ into $\tilde{P}$, we find $\tilde{P}(l_B^{VATzero}, t^{VATzero}) < \tilde{P}(l_B^{VATex}, t^{VATex})$, a contradiction.

Hence, $P^{VATzero} < P^{VATex}$, and, again using (1.10), $l_B^{VATzero} > l_B^{VATex}$. The former inequality implies increasing household utility. The latter inequality implies that banks' before-tax profit, $[R(l_B) - R'(l_B)l_B]$, increases (as is easy to show given $R'' \leq 0$). To complete the argument, note that neither VATex nor VATzero imposes taxes on banking profits, so banking surplus increases. □

Proof of Proposition 1.8. Let $l_B^{VATzero}, P^{VATzero}$ be the VATzero equilibrium values. Choose VAT rate t_{VATex} of the VATex system such that[26]

$$\begin{aligned} \tilde{P}\left(l_B^{VATzero}, t^{VATex}\right) &= \frac{(1 - t_{VATex})\kappa(l_B^{VATzero}) + \gamma(l_B^{VATzero}) + 1/R'(l_B^{VATzero})}{1 - t_{VATex}} \\ &\stackrel{!}{=} P^{VATzero} \end{aligned}$$

As $l_B^{VATzero}, P^{VATzero}$ satisfy equilibrium condition $x_H^*(P^{VATzero}) = l_B^{VATzero} R'(l_B^{VATzero})$, it is then also true that

$$x_H^*\left(\tilde{P}\left(l_B^{VATzero}, t^{VATex}\right)\right) = l_B^{VATzero} R'(l_B^{VATzero}).$$

Consequently, $l_B^{VATzero}$ solves equilibrium condition (1.10) for the VATex tax system and hence, $l_B^{VATzero}, P^{VATzero}$ are equilibrium values not only for the VATzero system but also for the VATex system. Then the corresponding surplus allocations are the same: 1. utility is the same as P is the same. 2. Before-tax banking profits are the same as l_B is the same, then banking surplus x_B is the same because neither tax system imposes taxes on banking profits. 3. The sum $c_G + x_B$ is the same by Lemma 1.9. Then public consumption must be identical because banking surplus is the same. □

Proof of Proposition 1.9. The proof is analogous to the proof of Proposition 1.6 and relies on the function $\tilde{P}(l_B, t)$ and equation (1.10). Let t^{FAT} be the FAT tax

[26] Given the expression for $\tilde{P}\left(l_B^{VATzero}, t^{VATzero}\right)$, it is straightforward to check that a suitable t_{VATex} exists.

system, let t^{VATex} be the VATex tax system, and let $t_V = t_z = t_x > 0$ be the respective value added tax rate under both regimes. Then, for all l_B,

$$\begin{aligned}\tilde{P}\left(l_B, t^{FAT}\right) &= \frac{1/R'(l_B)}{1-t_V} + (1+t_F)\kappa(l_B) + \frac{\gamma(l_B)}{1-t_V} \\ &\geq \frac{1/R'(l_B)}{1-t_V} + \kappa(l_B) + \frac{\gamma(l_B)}{1-t_V} = \tilde{P}\left(l_B, t^{VATex}\right),\end{aligned}$$

with a strict inequality if $\kappa(l_B) > 0$ or $\kappa_d + \kappa_l > 0$. On the other hand, we have equality, i.e., $\tilde{P}\left(l_B, t^{FAT}\right) = \tilde{P}\left(l_B, t^{VATex}\right)$ if $\kappa(l_B) = 0$ or $\kappa_d = \kappa_l = 0$. Let P^{FAT} and P^{VATex} be the respective equilibrium values of price ratio P, and let l_B^{FAT}, l_B^{VATex} be the loan supplies in the two tax regimes.

Analogous to in the proof of Proposition 1.6, we find that

$$\begin{aligned}\kappa_d = \kappa_l = 0 &\Rightarrow P^{FAT} = P^{VATex},\ l_B^{FAT} = l_B^{VATex} \\ \kappa_d + \kappa_l > 0 &\Rightarrow P^{FAT} > P^{VATex},\ l_B^{FAT} < l_B^{VATex}.\end{aligned}$$

For the concrete proofs of claims 1, 2, and 3, we can now invoke the very same arguments we made for claims 1, 2, and 3 in the proof of the Proposition 1.6. Only for claim 2 we do need a slight a modification; the after-tax profit, i.e., banking surplus, is 1 times before-tax profits under VATex, while it is $(1 - t_F) < 1$ times before-tax profits under a FAT. □

As spanish hauliers and French fishermen have shouted out for all the world to hear, higher fuel prices are not popular. This is uncomfortable for those – including this newspaper – who see increased taxation as a way of fighting global warming. Green taxes tend to fall hardest on the poor.

The Economist, June 2008

2 Environmental Taxation and Redistribution Concerns

2.1 Introduction

In June 2008, *The Economist* published an article discussing the pros and cons of a preceding oil price boost. On the one hand, the rise might be considered as "a gigantic carbon tax" that helped fighting global warming. On the other hand, it particularly hurt the poor who spent a considerably higher proportion of their income on fuel than the rich. Financial compensations for the core energy demand could help to solve the issue. However "it seems odd to try to prevent energy use with higher taxes ... and then to subsidise it" (The Economist, 2008). The article raised the question how to design green taxes optimally while accounting for distributive concerns. Rising awareness for global environmental problems under persisting inequality has increased the salience of that question. I propose an answer with a focus on the optimal level of green taxes and the relation to redistribution. Despite a huge theoretical literature on environmental taxation and quite some empirical interest on its impact on poor households, the normative question of the optimal response to inequality concerns in the environmental tax design has not had that much of attention.

I employ a simple Mirrlees (1971) income taxation framework which I extend by consumption externalities as proposed by Cremer, Gahvari, and Ladoux (1998). Within this framework, a welfare-optimising government uses non-linear income taxes to redistribute as well as Pigouvian taxation to reduce negative externalities (Pigou, 1932). I show that the two tax design problems are interconnected. In particular, the higher the level of redistribution, the lower the optimal level of environmental taxation. The optimal level has two determinants. First, the marginal social damage caused by the externality. Second, the cost of public funds, defined as the immediate marginal welfare losses associated with income tax collection.[1] If the government puts more weight on redistribution, it will have to accept a higher cost of public funds. Marginal revenues from the environmental tax are then more valuable from the government's point of view. Contrary to naive intuition, this calls for a lower environmental tax rate. The reason is that the tax rate is at its efficient level if the marginal revenues exactly compensate society for the marginal external harm. The more valuable the marginal revenues are, the less one needs to compensate for the marginal externality. To put it another way, consider the Pigouvian tax as a bribe that consumers pay the authorities in order to get allowance for pollution. The government is willing to accept a lower bribe if its utility per dollar is higher. Exactly this is the case if the cost of public funds is higher.

I measure the level of redistribution by a parameter that corresponds to the weight of less productive agents in a social-welfare function. As explained, Pigouvian taxation needs to decrease if the parameter increases. When first-best instruments are available, however, the result reverses. Without distortions, the cost of public funds actually decreases in the parameter, as the disutility of the hard working high productive agents receives less weight in the welfare function. Hence the first-best level of Pigouvian taxation increases with the level of redistribution.

My main contribution with respect to the existing literature is to draw attention to the *level* of Pigouvian taxation. Most of the respective literature focuses on tax *rules* and concludes that the distortions caused by second-best instruments do not alter these rules compared to first best. I show that, despite the first-best shape of these rules, the second-best level of Pigouvian taxation in fact depends on the distortions and the available income tax instruments.

This chapter also contributes to a branch of the literature that uses linear tax schemes to analyse the double-dividend hypothesis. Major insights from the linear model carry over to my setting with incentive constraints and optimal taxes. In particular, the optimal environmental tax is lower in second- than in first-best.

The remainder of the chapter is organised as follows. Section 2.2 discusses related literature. Section 2.3 presents the model. Section 2.4 states the rule for optimal internalisation. Section 2.5 introduces tax systems. Section 2.6 analyses

[1] In formal terms, the cost of public funds is the Lagrangian multiplier of the resource constraint.

optimal environmental taxes and provides the main results. Section 2.7 concludes. The appendices hold proofs and formal results. They also characterise optimal allocations and discusse corner solutions.

2.2 Related Literature

This chapter is part of a literature in which Pigouvian taxation meets non-linear income taxes under asymmetric information (Mirrlees, 1971). Cremer, Gahvari, and Ladoux (1998) show that under the separability assumptions from Atkinson and Stiglitz (1976) the optimal Pigouvian tax rate is uniform, i.e., it does not discriminate between agents. Gauthier and Laroque (2009) generalise the insight: a certain part of the second-best problem can be separated such that first-best rules apply for that part of the problem. Examples include Pigouvian taxation and the Samuelson Rule. Hellwig (2010) presents a similar result.

Kopczuk (2003), Pirttilä and Tuomala (1997), Jacobs and de Mooij (2011), and Kaplow (2012) explicitly centre on externality taxation within a general (income) taxation problem. In terms of questions posed their contributions are close to mine. Their answers have a different focus, though.

Kopczuk (2003) proposes to decompose the general taxation problem with externalities into two parts: "First, calculate the appropriate Pigouvian tax necessary to correct the externality. Then, with the externality accounted for, the usual second-best problem can be solved using standard formulae." (p. 84) His result holds for a variety of specifications (including the model presented here) and generalises the 'principle of targeting' (Dixit, 1985). Kopczuk (2003) also points out, though, that actually the two parts are interrelated: the Pigouvian tax rate might only be known after the whole problem is solved. My comparative statics analysis characterises this interrelation.

Kaplow (2012) summarises his findings by stating "that simple first-best rules – unmodified for labor supply distortion or distribution – are correct in the model examined." My analysis highlights that distribution and distortions have a significant influence on environmental policy with respect to tax *levels*, though.

Jacobs and de Mooij (2011, p. 2) find that the "optimal second-best tax on an externality-generating good should not be corrected for the marginal cost of public funds". However, they use a non-standard definition for the cost of public funds. Our formal analyses are consistent but focus on different interpretations.

An earlier branch of the literature, dating back to Sandmo (1975), examines environmental taxation as part of linear tax systems. Starting with Bovenberg and de Mooij (1994), the linear-taxation model was a primer workhorse model in the discussion of the double-dividend hypothesis.[2] By a central result of this literature, the second-best environmental tax is below the first-best one (e.g. Orosel and Schöb, 1996). As I show, these insights carry over to the case of optimal/non-linear income taxation. Metcalf (2003) uses the linear model to carry out a comparative static analysis with a focus on environmental quality.

My analysis also relates to the literature on comparative static properties of non-linear taxation, with and without public goods (Bierbrauer and Boyer, 2010; Brett and Weymark, 2008; Weymark, 1987), and to applied analyses of the question how to overcome negative distributional effects of environmental taxes (like Ekins and Dresner, 2004; Metcalf, 1999; West, 2005). Rausch, Metcalf, and Reilly (2011) recently studied the U.S. economy, Kosonen (2012) did so for the European Union. The empirical papers investigate the relationship between household income and emission-heavy consumption like driving or heating in order to check whether environmental taxes are regressive. They also discuss distributional impacts of environmental taxes and policies to support the poorest household. I add insights from normative theory to the discussion. In particular, I show that (a) emission-heavy consumption should not be subsidised for poor households and (b) whether or not environmental taxes are regressive is not per se relevant for their optimal level.

2.3 Model

The model is based on work by Cremer, Gahvari, and Ladoux (1998). There are three different goods. First, an intermediate good called *output*, denoted by Y. It serves as the numéraire and may be interpreted as money. Second, a *clean*, completely private consumption good, C, and third, a *dirty* consumption good, D. The intermediate good can be transformed into the consumption goods at fixed rates of transformation equal to p_C and p_D, respectively. Parameters p_C and p_D may be interpreted as the producer prices of C and D. The intermediate good itself can be produced with a linear technology using labour as the single input good (but labour is not modelled explicitly). The rate of transformation between labour and the intermediate good mirrors productivity and is denoted by w. It may be interpreted as the wage rate.

Households and Allocations There is a continuum of measure one of agents. They differ in exactly one dimension, namely their productivity, which can be either low or high. An agent's type is denoted by $\theta \in \{L, H\}$. Their respective

[2] See Goulder (1995), Schöb (1997, 2005), Bovenberg (1999) or Bovenberg and Goulder (2002) for more details and surveys on the double-dividend discussion.

productivity is $w_\theta \in \{w_L, w_H\}$. The fraction of low-type agents is denoted by $\gamma \in (0,1)$. An *allocation* A specifies levels of (C, D, Y) for both generic types, i.e., $A \equiv (C_L, D_L, Y_L, C_H, D_H, Y_H)$. For a given allocation the utility of an agent of type θ is

$$U_\theta(A) = u(C_\theta, D_\theta) - \frac{Y_\theta}{w_\theta} - (\gamma D_L + (1-\gamma)D_H)e. \tag{2.1}$$

Function u is continuously differentiable three times, strictly increasing, strictly concave, has nonnegative cross derivatives, and satisfies the Inada conditions.[3] It represents private consumption utility. The function satisfies standard Inada-conditions. In order to produce Y_θ units of output, an agent has to provide Y_θ / w_θ units of labour. This provision is associated with a linear disutility. The last term in the utility function reflects the externality. Independently of his type, every agent suffers from the overall consumption of dirty goods, $\gamma D_L + (1-\gamma)D_H$. The social harm is proportional to total dirty good consumption, and $e > 0$. From an agent's point of view, own consumption has no negative effect on own utility as a single contribution is negligible in comparison to the large contribution of others. Individual contributions are in fact zero due to the assumption of a continuum of agents.[4]

Notice that all agents in society have quite similar preferences. In particular, their consumption choice for a given budget is identical. Also, they suffer from the externality in exactly the same way. This is not only a simplification but rather a design choice. If agents had different tastes for environmental protection, then the optimal policy would obviously depend on distributional considerations. The homogeneity in agents' preference allows to isolate the more subtle relations between equity and environmental policy.

Social Welfare This chapter takes a normative perspective by examining what a social planner (SP) would do in order to maximise the social welfare function W, defined as

$$W(A) = \alpha U_L(A) + (1-\alpha)U_H(A), \ \alpha \in (0,1),$$

where A is the allocation. The welfare function is a weighted sum of the generic

[3]Formally, $u_{CD} \geq 0$ as well as $u_K \to \infty$ as $K \to 0$, and $u_K \to 0$ as $K \to \infty$, for $K \in \{C, D\}$. The Inada-conditions are imposed in order to guarantee strictly positive optimal consumption levels. Strict concavity guarantees unique solutions.

[4]Externalities of this type were termed "atmospheric" by Meade (1952). A different way to interpret the mechanism is to consider a public good that is provided by nature (like "fresh air" or "nice atmosphere"). Dirty good consumption reduces the level or quality of the public good, whereat only total consumption matters. The presented model wold fit this interpretation, with the initial amount of this public good normalised to zero.

types' utilities. The parameter α measures the weight SP puts on a generic low-type agent. If $\alpha = \gamma$, then W is the utilitarian welfare function. For $\alpha = 1$, W would be the Rawlsian welfare function.

Overall, the economy cannot consume more than it produces in terms of output. Furthermore, an exogenous revenue requirement r has to be met. The social planner thus faces a resource constraint given by

$$\gamma(Y_L - p_C C_L - p_D D_L) + (1-\gamma)(Y_H - p_C C_H - p_D D_H) - r \geq 0. \tag{2.2}$$

If (2.2) holds and $A \geq 0$, then A is *feasible*. An allocation that maximises W among all feasible allocations is a *first-best allocation*.

If the social planner does not observe an agent's type, not all feasible allocations are implementable. If, for instance, an allocation disadvantages the high-type agents, they might have an incentive to pretend to be low-types, making it impossible to implement this allocation. As a consequence, under asymmetric information, SP has to ensure that agents do not want to misrepresent their type. This is the case if the following *incentive-compatibility* constraints hold.

$$u(C_L, D_L) - \frac{Y_L}{w_L} \geq u(C_H, D_H) - \frac{Y_H}{w_L}, \tag{2.3}$$

$$u(C_H, D_H) - \frac{Y_H}{w_H} \geq u(C_L, D_L) - \frac{Y_L}{w_H}. \tag{2.4}$$

An allocation that maximises welfare among all feasible, incentive-compatible allocations is a *second-best allocation*. The underlying idea about the relation between incentive compatibility and decentral implementation, i.e. taxation, is known as the 'Taxation Principle' (Guesnerie, 1998; Hammond, 1979).[5]

By means of the following assumption, I restrict the analysis to the cases in which SP likes to redistribute from high-type agents to low-type agents.

Assumption 2.1. $\alpha(1-\gamma)w_H > (1-\alpha)\gamma w_L$.

The assumption generally holds if SP puts a sufficiently high welfare weight on low-type agents. The lower w_L is relative to w_H, the lower α may be, because a large difference in productivity provides an efficiency argument for making high-types work more than low-types. A low population share γ of low-types makes redistribution in their favour very cheap, hence it also allows for a low α.

Given the shape of u, it is efficient to produce strictly positive amounts of the consumption goods rather than abstain from economic activity. In turn, agents have to provide output. A look at Assumption 2.1 and the definition of W shows

[5]For a formal argument see Appendix 2.A.

that, in terms of welfare, it is always better to let the high- rather than the low-type agents produce an output unit. Consequently, high-type agents should produce *all* output. In first-best, this is indeed the case.[6] In second-best, this might be out of reach, as incentive constraints have to be satisfied. It is then ambiguous whether low-type agents work. My main analysis focuses on the cases in which they do work, i.e. $Y_L > 0$. In these cases a reallocation of output provision from low- to high-type agents improves welfare, but is possible only if high-types' incentive constraint (2.4) is slack. Consequently, at an interior second best allocation, (2.4) needs to bind. As Assumption 2.1 favours the low-type agents, their incentive constraint (2.3) is always slack.[7]

For a discussion of existence of second-best corner solutions and it properties, see Appendix 2.B.

2.4 Optimal Internalisation

This section provides a general property of Pareto-optimal allocations, with respect to the externality. At first sight, the presented rule is identical for first- and second-best allocations. This is a reason why redistribution and distortions are sometimes considered to have no structural influence on Pigouvian taxation. In the next step, I show, however, in what way the first- and second-best rules are in fact different.

To shorten exposition, I use the following notation for $\mathcal{J}, K \in \{C, D\}$. $u^L := u(C_L, D_L)$, $u^L_{\mathcal{J}} := \partial u(C_L, D_L)/\partial \mathcal{J}_L$, $u^L_{K\mathcal{J}} := \partial^2 u(C_L, D_L)/(\partial K_L \partial \mathcal{J}_L)$. Analogous definitions apply to $u^H := u(C_H, D_H)$. The Lagrangian multiplier of the resource constraint is denoted by λ. All results in the current section are derived in Appendix 2.A.

A Rule for Optimal Internalisation Both first- and second-best allocation feature the property that the marginal rates of substitution (MRS) between the two consumption goods are the same for both types of agents. Rather than being equal to the rate of transformation (namely, producer-price ratio), as would be the case in an unregulated market, the MRS is equal to

$$MRS = \frac{u^L_D}{u^L_C} = \frac{u^H_D}{u^H_C} = \frac{p_D}{p_C} + \frac{e}{\lambda p_C}. \tag{2.5}$$

[6] If Y_H could be negative, (first-best) welfare would be unbounded. Obviously, that is not an option. Accordingly, the nonnegativity constraint for Y_L binds at the first-best allocation.

[7] See Lemmas 2.1, 2.2, 2.3 in Appendix 2.A for the formal arguments.

This is a standard result in the literature. It follows, for instance, from the more general analysis by Hellwig (2010). It is driven by the separability feature of the utility functions. Cremer, Gahvari, and Ladoux (1998) point out the relation to the famous result in Atkinson and Stiglitz (1976), namely that, under the given assumptions, commodity prices should not be distorted, and all redistribution can be done within the labour market. The intuition of the Atkinson/Stiglitz result is as follows. By assumption, all agents have the same consumption pattern.[8] Therefore the commodity demand cannot be used to screen types and commodity taxation cannot contribute to relax the equity-efficiency trade-off. Hence there is no point in distorting them.

The intuition carries over partially to the case where an externality is introduced. In fact, as agents are equal in terms of their consumption preferences and their exposure to the externality, there is no point in treating them differently in this respect. Yet, it is no longer true that optimal redistribution only affects the labour market. Optimal consumption now depends on multiplier λ. The multiplier is crucially related to redistribution. Also, while at first sight the above formula is the same for both first- and second best allocation, λ is different in first- and second-best. This has significant consequences for the relation between the degree of redistribution and the degree of intervention in the commodity market.

2.4.1 The Cost of Public Funds

There is no universal definition for the (marginal) cost of public funds in the literature. Jacobs (2012), for instance, recently suggested a definition which implies a marginal cost of 1 for typical optimal taxation schedules. In this chapter I stick to the classical definition, also used in the textbook by Dahlby (2008); the (marginal) cost of public funds measures the loss in welfare associated with raising tax revenues. Being a cost, the concept does not account for potential benefits from the revenues. It just tells how (welfare-)costly it is to raise a (marginal) tax dollar.

As is well known, the so defined (marginal) cost of public funds is equal to λ, the Lagrangian multiplier for the resource constraint (2.2). Formally, $\lambda = -\partial \hat{W}/\partial r$, where $\hat{W}$ is the optimised value of the welfare function. Throughout the chapter I normally drop the explicit "marginal" when referring to the cost of public funds – relying on the fact that the concept is per se a marginal one. Also, "multiplier" interchangeably refers to λ, i.e., the cost of public funds.

The quasi-linearity in labour allows for closed-form solutions for the multiplier and plainly reveals the dissimilarity between first- and second-best. It also shows the dependency on the underlying parameters α and γ.

[8] More precisely, for a given amount of total consumption spending, all agents consume the same commodity bundle.

Second-best The value of the multiplier at an interior second-best allocation is

$$\lambda^* := \frac{\alpha}{w_L} + \frac{1-\alpha}{w_H}. \tag{2.6}$$

To grasp the intuition, note that agents do not benefit from r, so an increase is pure burden. A way to finance the additional requirement is to increase output. As the incentive constraint for the high-type agents is binding, their output may only be increased if the low-type's output is increased as well. The weighted welfare loss of such an increase is equal to α/w_L for the generic low-type and $(1-\alpha)/w_H$ for the generic high-type. Notice that the multiplier does not depend on the population shares. The reason is that a higher revenue requirement has to be produced by all agents (independently of their type) in order to sustain incentive compatibility.

First-best The multiplier at the first-best allocation is

$$\lambda^F := \frac{1-\alpha}{w_H(1-\gamma)}. \tag{2.7}$$

Because only high-types work in first-best, only parameters related to them matter for λ^F. If SP needs an additional unit of revenue, he will make high-type agents work more. As there are only $1-\gamma$ high-type agents, the generic high type has to provide $1/(1-\gamma)$ (marginal) units of output and needs to work $1/(w_H(1-\gamma))$ additional hours. The incurred marginal disutility is weighted by $1-\alpha$.

The multipliers are not only different in size, but also with respect to their directions of change in the parameters α and γ. The welfare weight has an impact on the optimal tax design with respect to the externality. If interpreted naively, the optimal rule (2.5) itself hides this fact.

2.5 Taxation

The current section adapts the interpretation of output being money. In this interpretation, Y denotes gross income, w corresponds to the wage rate, and p_C, p_D are producer prices. A *tax system* $\tau = (t_C, t_D, T)$ consists of an income tax function T and specific commodity taxes $t_C, t_D \in \mathbb{R}$. Consumer prices are $q_k := p_k + t_k$ for $k \in \{C, D\}$. Consumption may be subsidised through negative commodity taxes. T may be negative as well, in which case it is a transfer to the agent.

For any type θ, let $(C_\theta(\tau), D_\theta(\tau), Y_\theta(\tau))$ be the maximisers of individual utility, given τ:

$$
\begin{aligned}
(C_\theta(\tau), D_\theta(\tau), Y_\theta(\tau)) \in \underset{(C,D,Y)}{\operatorname{argmax}} \Bigg(& u(C,D) - \frac{Y}{w_\theta} - (\gamma D_L + (1-\gamma) D_H) e \\
& \text{subject to } \; q_C C + q_D D \leq Y - T(\cdot) \Bigg) .
\end{aligned}
\tag{2.8}
$$

Households take D_L and D_H as given, so the externality is not relevant for their decision.

When choosing a tax system, the social planner takes individual optimisation into account and needs to respect the following fiscal budget constraint, which is equivalent to resource constraint (2.2).

$$
\gamma \Big(T_L(\tau) + t_C C_L(\tau) + t_D D_L(\tau) \Big) + (1-\gamma) \Big(T_H(\tau) + t_C C_H(\tau) + t_D D_H(\tau) \Big) \geq r \tag{2.9}
$$

Here, $T_L(\tau)$ and $T_H(\tau)$ amount to the respective total income tax payments of low- and high-type agents.

If (2.9) is satisfied for some tax system τ, then τ is said to *implement* allocation A with $A = (C_L(\tau), D_L(\tau), Y_L(\tau), C_H(\tau), D_H(\tau), Y_H(\tau))$ as defined by (2.8). The set of available tax systems to choose from depends on the informational constraints. When the social planner can observe an agent's type, the income tax may be contingent on the type. Under asymmetric information it can only be contingent on observed gross income. In fact, with $T : (w, Y) \mapsto T(w, Y)$, it is possible to find a system τ that implements the first-best allocation. With $T : Y \mapsto T(Y)$, it is possible to find a system τ that implements the second-best allocation. This insight allows to restrict attention to the chosen tax structure albeit the linearity in commodity taxation.[9]

2.5.1 Normalisation

As usual in these type of models, there is a degree of freedom in the taxation choice. A common way to deal with this is to normalise the tax system and often it is innocuous to do so. Yet, when properties of the tax system, like a particular tax level, are the object of interest rather than the real allocation, one has to be careful with normalisations.

This was a major issue in the double-dividend discussion between Bovenberg and de Mooij (1994, 1997), Fullerton (1997) and others. The discussion centres on the comparison of the second-best pollution tax and the first-best Pigouvian

[9] The underlying arguments are standard. For details, see Appendix 2.A.

tax (the marginal social harm). The actual tax level obviously depends on the chosen normalisation and a priori it is unclear which normalisation is "correct". In a related contribution, Schöb (1997) focuses on the normalisation choice and shows that also "the difference of the first-best and second-best optimal tax on the polluting good depends on the normalization chosen." (p. 174) He concludes that "such a comparison provides an inappropriate indicator for the existence of a second dividend." (ibid.)

To obtain valid results on comparative static properties of environmental taxation and the relation between first- and second-best level, it is important to avoid the "normalization trap" (ibid.). Orosel and Schöb (1996) propose to study an object called the *second-best internalization tax*. Unlike an actual tax rate, it is a "real" variable, derived directly from the underlying allocation, and independent of the normalisation.[10] Using their concept, the authors find a particular normalisation to be correct for doing the comparison of actual first- vs. second-best tax rates.

The aforementioned contributions feature linear labour and commodity taxation, and do not model distributive issues.[11] Their insights on normalisations carry over to my model, though. Here the "real" object of interest is the so called *greenness* g. For any tax system $\tau = (t_C, t_D, T)$, the greenness is defined by

$$g := t_D - t_C \frac{q_D}{q_C}.$$

Similar to the second-best internalisation tax as proposed by Orosel and Schöb (ibid.), the greenness is a real variable independent of the normalisation. More precisely, g is unique in the sense that it is the same for any tax system which implements the second-best allocation. The same is true with respect to the first-best allocation.

For an intuitive understanding of the greenness consider an agent who faces some tax system τ and decides to purchase an additional (marginal) unit of D, while keeping total spending constant. Then the greenness is the change in the agent's total tax payment. Thereby it quantifies the tax system's inherent incentives to shift consumption from D to C. Plainly put, it tells how "green" the system is. Notice that the a tax system could be green due to high t_D or due to low (potentially negative) t_C. The greenness covers both cases.

It turns out that the greenness equals the tax rate t_D on the dirty good iff the tax rate t_C on the clean good is normalised to zero: precisely the normalisation

[10]The definition of the second-best internalisation tax uses the observation that private marginal utility should equal social marginal welfare – a property of an allocation rather than a tax system.

[11]The papers on the double dividend normally have identical/representative consumers. Distributional concerns appear only indirectly as a motive for the unavailability of lump sum taxation.

identified as "correct" by Orosel and Schöb (1996) for the respective purpose. For this reason it is save to proceed the analysis with $t_C \equiv 0$. For further reference, I call such a tax system *normalised*.[12]

2.6 The Optimal Pigouvian Tax

Given $t_C \equiv 0$, how high should t_D be? An optimum is characterised by the fact that a marginal reallocation does not change welfare. In particular, keeping private spending constant, a marginal change in consumption levels must not change welfare. Consider a marginal shift from C to D (for all agents, taking account of differences in prices). This has three effects: 1. Consumption utility u is unchanged as agents are at their individual optimum. 2. External harm increases at rate e. 3. Tax revenues increase at rate t_D and relax the budget constraint of the social planner. Multiplier λ tells how welfare is affected from relaxing the public budget. Thus the marginal effect of tax revenues on welfare amounts to $t_D\lambda$. The overall marginal change in welfare is $-e + t_D\lambda$. For this change to be zero, t_D needs to be

$$t_D = \frac{e}{\lambda}\,. \tag{2.10}$$

2.6.1 The Comparative Statics of Pigouvian Taxation

The following Propositions essentially combine equation (2.10) with the findings from section 2.4.1. They state the main result of the chapter: comparative static properties of those tax systems that implement the first- and second-best allocation, respectively.

Proposition 2.1 (First-best Pigouvian taxation). *If a normalised tax system $\tau^F = \left(0, t_D^F, T^F\right)$ implements the first-best allocation A^F, then*

$$t_D^F = \frac{e}{\lambda^F}\,.$$

Furthermore,

$$\frac{\partial t_D^F}{\partial \alpha} > 0, \qquad \frac{\partial t_D^F}{\partial \gamma} < 0.$$

Despite the lack of distortions, distributive concerns influence the environmental tax; more redistribution calls for a higher first-best Pigouvian tax t_D^F. The rela-

[12] A detailed discussion of the greenness and the respective proofs is contained in an earlier working paper version of Aigner (forthcoming), which is available upon request.

tion reverses completely if first-best instruments are not available and the labour market is distorted.

Proposition 2.2 (Second-best Pigouvian taxation). *If normalised tax system $\tau^* = (0, t_D^*, T^*)$ implements an interior second-best allocation A^*, then*

$$t_D^* = \frac{e}{\lambda^*} .$$

Furthermore,

$$\frac{\partial t_D^*}{\partial \alpha} < 0, \qquad \frac{\partial t_D^*}{\partial \gamma} = 0.$$

In a nutshell, higher labour market distortions coming from increased redistribution imply a lower optimal Pigouvian tax level. To develop a detailed intuition for the results, decompose the comparative statics into two aspects. (1) In first- as well as in second-best t_D is inversely proportional to λ. (2) The reaction of λ differs for first- and second-best. The first aspect is not new. It is already well established for models of linear labour/commodity taxation. As shown, it carries over to a world with incentive constraints. The second aspect has not drawn that much of attention in the literature but is crucial as it drives the reversed results. I discuss the two aspects in turn.

The inverse relation of environmental taxes and the cost of public funds

To grasp the intuition behind the inverse relation, consider the purpose of Pigouvian taxation: its (only) goal is to restore the efficient level of dirty-good consumption. From a welfare perspective, a unit of the dirty-good should be consumed if and only if consumption is not only individually optimal, but private benefits also outweigh social harm. Consequently, dirty-good consumption is at its socially optimal level only if marginal private (net) benefits exactly equal marginal social harm. To measure and compare these two objects, it is useful to quantify them in terms of money.

(1) The optimising agent is willing to pay t_D units of additional taxes for the right to consume her last unit of D rather than spending the respective money on C. So t_D is a good measure of (net) private benefits of the marginal unit of dirty-good consumption.

(2) Now consider the social planner. If D increases by one unit, welfare decreases by e. If SP receives exactly e/λ units of money to relax the budget constraint, welfare increases by $(e/\lambda)\lambda = e$. Thus the marginal social harm measured in money is equal to e/λ. It is the exact amount of money that society needs as a

compensation for additional dirty-good consumption. The amount is lower if the received money is more useful in the sense that the cost of public funds is higher. Putting together (1) and (2) shows that if $t_D = e/\lambda$, then individual maximisation leads to an allocation in which, at the margin, private (net) benefits equal social harm.

A more naive view, which evaluates Pigouvian taxes in a partial or isolated manner rather than viewing it as part of a whole tax system, could reason that 'Pigouvian taxes do two things: reduce pollution and create revenue. So they should be high if pollution is severe or if revenues are very valuable to the state.' Naive intuition would thus suggest that higher cost of public funds (associated with marginal tax revenue being more valuable) should lead to higher Pigouvian tax rates. In fact, this 'rationale' would provide a straightforward argument for the double dividend hypothesis, which by now has been mostly falsified (e.g., Fullerton and Metcalf, 1998). The strong form of the double-dividend hypothesis states that a revenue-neutral introduction of green taxes is desirable even if environmental benefits are not taken into account (Goulder, 1995). In the model that I propose this fails clearly: Pigouvian taxation, namely $t_D > 0$, is optimal only if an externality is present, i.e., if $e > 0$. Among others, Bovenberg (1999) gives the same argument, albeit for a model with linear taxation. Empirical investigations by Goulder (1995) tend to reject the hypothesis as well.

The intuition that rejects the double dividend hypothesis is also central to the comparative static analysis. The more valuable the marginal tax revenues, the less is needed to compensate for the marginal externality, and — because the one and only purpose of Pigouvian taxation is to induce alignment of private benefits and social harm at the margin — a *lower* Pigouvian tax rate is asked for. Various authors have noticed the underlying rationale in their respective settings, so it applies quite universally (e.g., Schöb, 1997). As Bovenberg and de Mooij (1994) put it, "each unit of pollution does not have to yield as much public revenue to offset the environmental damage if this revenue becomes more valuable" (p. 361).

In a recent contribution, Jacobs and de Mooij (2011) make the seemingly contradictory statement that the optimal second-best environmental tax is not sensitive to the cost of public funds at all. Their conclusion follows from a their newly proposed definition of the cost of public funds. So the difference in conclusion is one of interpretations rather than formal results. Their interpretation suggests that tax distortions do not play a role for optimal environmental taxes, which clearly is at odds with my analysis. Indeed, Jacobs and de Mooij (ibid., p. 13) qualify their interpretation themselves: "The optimal second-best environmental tax does require a correction for distributional concerns and interactions with labor supply, but not for pre-existing tax distortions." The comparative statics results fill the gap of specifying the "correction for distributional concerns" but also broaden the existing insights by highlighting that even without distortions, distribution concerns influence the optimal environmental tax level.

I should highlight that the preceding discussion is about *marginal* rather than *total* revenues. The difference is crucial: total revenues from Pigouvian taxation do *not* compensate for the overall external harm. Although the two figures coincide in the linear specification, they generally differ. More to the point, Pigouvian revenues should not be used to compensate the harmed people; it is not its purpose, and it might reduce incentives to avoid exposure to an externality in the first place (W. Oates, 1995).

How the cost of public funds changes in parameters

The changes of t_D^F and t_D^* with respect to welfare weight α have different signs. This point is worth stressing again as previous contributions with non-linear income taxes tend to highlight the similarities rather than the differences of first- and second-best Pigouvian taxes. This focus comes naturally when examining the optimal rules, which are – almost – identical for first- and second-best.

For the version of their model that resembles the one of this chapter, Cremer, Gahvari, and Ladoux (1998) conclude that "the optimal tax on the externality generating good is strictly Pigouvian" (p. 345; Proposition 1), where the term 'Pigouvian' is based on the first-best tax on the dirty good (Definition 1). Likewise Gauthier and Laroque (2009) show that first-best rules quite often hold also at second-best allocations if utility is separable. With respect to externalities they find that "a non-satiated second best allocation can be supported with a first best Pigovian tax" (Remark 4).[13] Kopczuk (2003) and Kaplow (2012) make similar observations.

While all of these findings are correct, they suggest (quite explicitly in some cases) that distortions are not that relevant for the second-best tax. Proposition 2.2 highlights the opposite. Also, these results might distract from the considerable differences between first- and second-best when it comes to tax *levels* rather than tax *rules.* In fact, the optimal rule for the model at hand is given in (2.5) – for both first- and second best. Only an inspection of the respective multipliers reveals the differences between them.

First-best multiplier Recall that only high types work at a first-best allocation and that λ^F is derived from an output increase of high-types. Now, if α increases,

[13] Gauthier and Laroque (2009) do point out, though, that the whole second-best problem must be solved to obtain the actual Pigouvian tax.

SP cares less about high-type agents working more, thus the cost of public fund *decrease*. Tax revenues generated by t_D are less valuable per unit so more (marginal) revenues need to be collected at the optimum. If γ is increased, the generic high-type has to work more for an higher overall output requirement and the cost of public funds increases. Marginal revenues generated by t_D now have higher value per unit and less is needed to satisfy optimality condition (2.10).

Second-best multiplier From (2.6) the cost of public funds at a second-best allocation is $\lambda^* = \alpha/w_L + (1-\alpha)/w_H$. As λ^* does not depend on γ, neither does t_D^*. The higher the welfare weight of low-type agents, the more redistribution is asked for and the more distortions are accepted. Higher distortions imply higher excess burden of taxation and thereby *higher* cost of public funds. Marginal revenues from Pigouvian taxation are then more valuable and less marginal revenue is needed to satisfy optimality condition (2.10).

First vs. second best As argued, the different directions of change of the Pigouvian tax with respect to the welfare weight derive from the different reactions of the cost of public funds. In the one case, the social planer cares less about those who work, i.e., environmental revenues decrease in value, in the other case, income tax distortion increase and environmental revenues increase in value.

Distribution and taxes Summing up, this section highlights a link between Pigouvian taxation and the degree of redistribution, measured by welfare weight α. Welfare optimising societies with different opinions about equity need to have different levels of Pigouvian taxation, even if first-best instruments are feasible. This is not entirely obvious, because in partial equilibrium models, the level of Pigouvian taxation is typically pinned down solely by Pareto efficiency. Asymmetric information proofs to be a crucial determinant of the link between environmental taxation and redistribution: the sign of the dependence changes when going from first- to second-best.

Corner solutions So far, the second-best comparative statics assumed an interior solution. Corollary 2.2 in Appendix 2.B shows that for corner solution $\partial t_D/\partial\alpha = 0$. In fact, if $Y_L^* = 0$ for some α, then a further increase in α cannot change the optimal allocation at all: it is neither possible to decrease Y_L nor to increase low-type agents' consumption without violating incentive constraints. Consequently, the optimal tax remains constant as well.[14]

[14] The result is shown to hold for $\alpha \geq \gamma$. I expect it to hold for a broader range of parameters, though.

2.6.2 First- vs Second-best Pigouvian Tax Level

Bovenberg and de Mooij (1994) examine the double dividend hypothesis by comparing first- and second-best environmental tax. Their contribution led to quite some follow-up papers on the subject. The workhorse model of this literature is a representative household model with linear income and commodity taxes (e.g., Bovenberg, 1999). Lump sum taxes are allowed or disallowed for exogenous reasons. In these settings, the second-best environmental tax falls short of the first-best one.[15] The result carries over to my setting with incentive-constraint redistribution.

Proposition 2.3. *Fix parameters and consider two normalised tax systems, τ^F and τ^*, which implement the first- and second-best allocations, respectively. Then*

$$t_D^* < t_D^F.$$

Proof. (2.7) and (2.6) imply $\lambda^* - \lambda^F > 0$ due to Assumption 2.1. $t_D^* < t_D^F$ follows from (2.10). □

Following the discussion of the comparative statics, the intuition for the result should be clear. The second-best tax system distorts the labour market which increases the cost of public funds compared to a first-best system. The difference in the cost of public funds causes the difference in the tax levels.

2.6.3 The Role of Regressivity

Environmental taxes are *regressive* if total tax payments in proportion to total consumption spending decrease in income levels. If so, these taxes impose a disproportionate burden on low-income households. Energy-intensive goods like electricity and heating are often considered to feature regressive consumption pattern. Taxes on these goods might then indeed be regressive. The Economist (2008), for instance, is concerned about this possibility in the article I cited in the Introduction. The presumption also gave rise to applied studies on the impact of green tax reforms on low-income households (e.g., Ekins and Dresner, 2004; Metcalf, 1999; West, 2005). In a recent empirical study focusing on the European Union, Kosonen (2012) finds that electricity and heating tend to be regressive. For transport fuel and vehicles it is the other way around, though; they seem to be progressive. Also, there are considerable differences between countries. Overall, the evidence for regressive spending patterns is quite mixed.

[15] See also Schöb (2005). As detailed in section 2.5.1 and the references mentioned there, the comparison hinges on the “correct” normalisation choice, which by now is well-understood (Orosel and Schöb, 1996).

The analysis in this chapter contributes to these considerations by showing that the question of regressivity might not be that important after all. To make this point, compare two examplary utility functions, U_θ^1 and U_θ^2. In line with (2.1), let

$$U_\theta^1(A) = u^1(C_\theta, D_\theta) - \frac{Y_\theta}{w_\theta} - (\gamma D_L + (1-\gamma) D_H)e,$$
$$U_\theta^2(A) = u^2(C_\theta, D_\theta) - \frac{Y_\theta}{w_\theta} - (\gamma D_L + (1-\gamma) D_H)e.$$

The utility functions differ only in consumption utility u, which is u^1 or u^2, resp., but are identical with respect to disutility of labour and with respect to the externality. To be concrete, suppose

$$u^1(C,D) = 2(CD)^{1/2}, \qquad u^2(C,D) = 2(2D^{1/2} + C)^{1/2}.$$

As defined in (2.8), households maximise utility facing consumer prices q_D and q_C. Let m^* be the individually optimal amount of after-tax income. Then dirty good demand is $D^1 = m^*/(2q_D)$ under utility U_θ^1, and is $D^2 = (q_C/q_D)^2$ under utility U_θ^2 (provided m^* not too small). Hence in the former case, dirty good taxation is not regressive at all, because dirty good consumption is proportional to after-tax income. In the latter case, though, it is highly regressive: rich and poor households typically spend the exact same amount on dirty good taxes, irrespective of after-tax income.[16]

How do these different degrees of regressivity translate into differences in optimal environmental taxation? The respective analyses show that the optimal tax on the dirty good is either $t_D^1 = e/\lambda^1$ or $t_D^2 = e/\lambda^2$. Yet, multiplier λ is independent of consumption utility u, as discussed in section 2.4.1. So $\lambda^1 = \lambda^2$, because the two version of overall utility U are identical except for u. Hence,

$$t_D^1 = t_D^2.$$

This example shows that two economies can actually have the very same optimal level of environmental taxation, despite the fact that the dirty good tax is regressive in one of them and is not regressive in the other one.[17] Of course, this neutrality observation would be diluted without quasilinear utility U. However,

[16]Consumption utility u^1 is a standard Cobb-Douglas utility and as such satisfies the conditions imposed upon u in the model description. Consumption utility u^2, on the other hand, does not satisfy all these condition. Nevertheless, I use it in this example because the quasi-linear structure is a familiar form, which is well-known to induce regressive demand patterns. Though some parts of the formal analysis in the Appendix do not (necessarily) apply to U^2, the first-orders condition relevant for the example are not affected.

[17]The income tax T generally differs, though. If marginal income taxes were different as well, the

regressivity apparently is not relevant for the Pigouvian tax per se – otherwise this should appear in the example, which isolates a pure regressivity effect.

2.7 Conclusion

This chapter looks at the interdependence of distributive and environmental policies from a normative perspective. It reveals a qualitative difference between first- and second-best. Distributive goals and environmental policies are linked by the cost of public funds. On the one hand they influence the optimal environmental tax level, on the other hand they are a function of distribution policies. I find that if society wants more redistribution, the second-best environmental tax is lower, whereas the first-best environmental tax is higher.

The results also clarify some aspects of the literature on Pigouvian taxation. First, it is important to distinguish optimal rules from optimal levels. Former contributions on second-best environmental taxation with non-linear income taxes tend to focus on the optimal tax rule and point out their "first best flavor"[18], emphasising the similarity of first- and second-best with respect to environmental taxes. My focus on tax level shows significant differences in the level and the parameter dependence. Income tax distortions do play a substantial role for optimal environmental taxes.

Second, insights gained from models with linear income/commodity taxation carry over to settings with non-linear income taxation and incentive constraints. This holds true for the role of distortions as well as the result that the second-best environmental tax falls short of the first-best one.

What can be learned in terms of policy implications? First, the view that the two goals of redistribution and environmental protection can be addressed independently by means of two different instruments (income tax and Pigouvian taxation) needs to be reconsidered. In particular, the designer of environmental taxes has to account for the value in terms of welfare that is created by the tax revenues. This value is a function of the income tax schedule and depends on the set of available instruments as well as on informational constraints. The optimal tax level then derives from the trade off between external harm and useful tax revenues. Importantly, it is the marginal effect that counts. Total revenues are irrelevant for the optimal level of Pigouvian taxation. So are the total environmental taxes paid by the households.

incentive to pollute would change despite constant t_D. An analysis of the optimal allocation shows, however, that this is not the case; the distortions induced by redistribution do not hinge on the particularities of u (see Appendix 2.A, Proposition 2.5).

[18] Gauthier and Laroque (2009, p. 1168)

Second, the intricate empirical question of regressivity is not too relevant for tax designers. In fact, whether or not environmental taxes are regressive should not influence their level. In particular, tax rates should not be reduced for poor households in an attempt to compensate for any disproportionate burden from environmental taxes. This would reduce incentives and provide an inefficient means of redistribution. Instead, one might raise the transfers to those households.

Appendix 2.A Optimal Allocation and Proofs

In this appendix, I characterise the first- and second-best allocation and give the proofs of the results on optimal taxation.

Lemma 2.1. *If A^F is a first-best allocation under Assumption 2.1, then $Y_L^F = 0$ and $Y_H^F > 0$.*

Proof. Suppose $Y_L^F > 0$. If the total output of all low-type agents is lowered by $\Delta \in (0, \gamma Y_L^F)$, every low-type individual may reduce his own output by Δ/γ. The immediate welfare gain is $\alpha\Delta/(\gamma w_L)$. To finance the output reduction high types have to increase their total output by Δ, resp. their individual output by $\Delta/(1-\gamma)$. The immediate welfare loss is $(1-\alpha)\Delta/((1-\gamma)w_H)$. The net effect of the alteration is strictly positive given Assumption 2.1, a contradiction.

Hence $Y_L^F = 0$. $Y_H^F > 0$ needs to holds given the Inada-conditions on u. □

Lemma 2.2. *If A is a second-best allocation, then*

1. *At most one incentive compatibility constraint is binding.*
2. *$Y_H > Y_L$ and $u^H > u^L$.*

Proof. 1. Suppose the contrary. Summation of both ICs yields $Y_L = Y_H$ and $u(C_L, D_L) = u(C_H, D_H)$. Given the shape of u, this can be optimal only if $(C_L, D_L) = (C_H, D_H)$. To complete the argument, it suffices to show that such a bunching allocation is dominated by a *constrained laissez-faire* allocation. Fix any feasible bunching allocation $A^b = (C^b, D^b, Y^b, C^b, D^b, Y^b)$ and define for any type θ, $(C_\theta^{lf}, Y_\theta^{lf}) :=$ argmax $_{C,Y}\{u(C, D^b) - Y/w_\theta$ s.t. $Y \geq p_C C + p_D D^b + r\}$. Then, in particular, $u(C_\theta^{lf}, D^b) - Y^{lf}/w_\theta \geq u(C^b, D^b) - Y^b/w_\theta$. Furthermore, maximisers are unique and $(C_L^{lf}, Y_L^{lf}) \neq (C_H^{lf}, Y_H^{lf})$. Thus there exist θ such that $u(C_\theta^{lf}, D^b) - Y^{lf}/w_\theta > u(C^b, D^b) - Y^b/w_\theta$. The constraint laissez-faire allocation A^{lf} thereby Pareto-dominates the bunching allocation A^b. A^{lf} is also incentive compatible and feasible. Hence, A^b cannot be a solution to the second-best problem and the contradiction is completed. (The argument builds on Bierbrauer and Boyer, 2010, Lemma 1)

2. Add both ICs to obtain $Y_H \geq Y_L$. Equality would imply a bunching allocation which is not optimal as shown above. Hence $Y_H > Y_L$. High types' incentive constraint then implies $u^H > u^L$. □

Lemma 2.3. *If A^* is an interior second-best allocation under Assumption 2.1, then high types' incentive constraint* (2.4) *is binding, low types' incentive constraint* (2.3) *is slack.*

Proof. Suppose by contradiction that (2.4) was slack, i.e., $u^H - Y_H/w_H > u^L - Y_L/w_H$. Then there exits an $\varepsilon > 0$ such that also $u_H - (Y_H + \varepsilon)/w_H > u_L - (Y_L - \varepsilon(1-\gamma)/\gamma)/w_H$. The ε-perturbed allocation is constructed in a way to keep total output constant. Incentive compatibility is sustained, too. The welfare effect of the perturbation is

$$dW = \alpha \frac{1-\gamma}{\gamma w_L}\varepsilon - \frac{(1-\alpha)\varepsilon}{w_H} > 0 \Leftrightarrow \frac{\alpha}{\gamma w_L} > \frac{(1-\alpha)}{(1-\gamma) w_H}$$

dW is strictly positive precisely under Assumption 2.1, hence a contradiction.

If IC_H is binding then IC_L must be slack by Lemma 2.2. □

First-order conditions Considering the lemmas, an appropriate Lagrangian for an optimal allocation is

$$\begin{aligned}\mathcal{L} = {} & \alpha\left[u(C_L, D_L) - Y_L/w_L - (\gamma D_L + (1-\gamma)D_H)e\right] \\ & + (1-\alpha)\left[u(C_H, D_H) - Y_H/w_H - (\gamma D_L + (1-\gamma)D_H)e\right] \\ & + \lambda(\gamma(Y_L - p_C C_L - p_D D_L) + (1-\gamma)(Y_H - p_C C_H - p_D D_H) - r) \\ & + \mu(u(C_H, D_H) - Y_H/w_H - u(C_L, D_L) + Y_L/w_H) \\ & + \delta Y_L.\end{aligned} \tag{2.11}$$

Next, set the partial derivatives to zero:

$$\alpha u_C^L - \gamma\lambda p_C - \mu u_C^L = 0 \Rightarrow (\alpha - \mu) u_C^L = \lambda\gamma p_C \tag{2.12}$$

$$\alpha u_D^L - \gamma\lambda p_D - \mu u_D^L - \gamma e = 0 \Rightarrow (\alpha - \mu) u_D^L = \lambda\gamma p_D + \gamma e \tag{2.13}$$

$$-\alpha/w_L + \gamma\lambda + \mu/w_H + \delta = 0 \Rightarrow \gamma\lambda = \alpha/w_L - \mu/w_H - \delta \tag{2.14}$$

and

$$\begin{aligned}&(1-\alpha)u_C^H - (1-\gamma)\lambda p_C + \mu u_C^H = 0 \\ &\Rightarrow (1-\alpha+\mu)u_C^H = (1-\gamma)\lambda p_C\end{aligned} \tag{2.15}$$

$$\begin{aligned}&(1-\alpha)u_D^H - (1-\gamma)\lambda p_D + \mu u_D^H - (1-\gamma)e = 0 \\ &\Rightarrow (1-\alpha+\mu)u_D^H = (1-\gamma)(\lambda p_D + e)\end{aligned} \tag{2.16}$$

$$\begin{aligned}&-(1-\alpha)/w_H + (1-\gamma)\lambda - \mu/w_H = 0 \\ &\Rightarrow (1-\alpha+\mu)/w_H = (1-\gamma)\lambda\end{aligned} \tag{2.17}$$

It follows that

$$\mu = w_H\left(\frac{(1-\gamma)\alpha}{w_L} - \frac{\gamma(1-\alpha)}{w_H}\right) - w_H(1-\gamma)\delta \tag{2.18}$$
$$\lambda = \frac{\alpha}{w_L} + \frac{1-\alpha}{w_H} - \delta$$

For a first-best allocation, set $\mu = 0$, for an interior second-best allocation, set $\delta = 0$.

Proposition 2.4 (First-best allocation). *Given Assumption 2.1, allocation A^F is a first-best allocation if and only if it satisfies the following system of equations.*

$$u_C^L = \frac{p_C}{w_H}\frac{1-\alpha}{\alpha}\frac{\gamma}{1-\gamma}, \qquad u_D^L = \frac{p_D + e/\lambda^F}{w_H}\frac{1-\alpha}{\alpha}\frac{\gamma}{1-\gamma}, \tag{2.19}$$
$$u_C^H = \frac{p_C}{w_H}, \qquad u_D^H = \frac{p_D + e/\lambda^F}{w_H}, \tag{2.20}$$

$$Y_L^F = 0, \quad Y_H^F = \frac{\gamma}{1-\gamma}\left(p_C C_L^F + p_D D_L^F\right) + p_C C_H^F + p_D D_H^F + \frac{r}{1-\gamma},$$
$$\lambda^F = \frac{1-\alpha}{w_H(1-\gamma)}.$$

Proof. With $\mu = 0$, the Lagrange function (2.11) is concave and the first order conditions are necessary and sufficient for a solution. Consider conditions (2.14) and (2.17) with $\mu = 0$. Then

$$\lambda^F = \frac{1-\alpha}{w_H(1-\gamma)}, \quad \delta^F = \frac{\alpha}{w_L} - \gamma\lambda > 0 \Rightarrow Y_L^F = 0.$$

Notice that the inequality is satisfied if and only if Assumption 2.1 holds. The statement of the Proposition now follows from conditions (2.12), (2.13), (2.15), (2.16), and the binding resource constraint (2.2). □

According to Proposition 2.4, low-type agents do not work at all. Due to linear disutility from working, Assumption 2.1 implies that any given amount of output requirement fosters lower aggregated disutility if it is provided solely by high types rather than low types. If Y_L could be negative, welfare would be unbounded.

For a moment, ignore the Lagrangian multiplier of the resource constraint λ^F. Then consumption of high types is independent of the welfare weight and the population shares, and is just determined by efficiency considerations. It departs from standard results only through a corrective element that takes care of the external effects of dirty-good consumption. The consumption levels of the low-type agents, though, heavily depend on welfare weights as well as the population shares. The underlying trade-off lies between consumption utility of low-types and disutility of high types, who have to work for the provision of low-type consumption. Low-type productivity w_L is irrelevant for the allocation given that they do not work.

Lemma 2.4. *Let $u(C, D)$ be strictly concave and continuously differentiable and let k^C, k^D be two constants such that the system $u_C(C, D) = k^C,\ u_D(C, D) = k^D$ has a solution. Then the solution is unique.*

Proof. Consider the three-dimensional space. Let $s = (s^C, s^D)$ be a solution. The tangential plane at $S = (s^C, s^D, u(s))$ is spanned by the directions of the two partial derivatives at S. As u is strictly concave, the whole range of u – except $u(s)$ – lies below that plane. Now consider a point s' that also solves the above system but is different from s. The tangential plane at s' is parallel to the one at s, yet one of the planes is higher than the other. But than it is no longer possible that the whole range of u lies below the lower plane. This creates a contradiction. □

Proposition 2.5 (Interior second-best allocation). *If A^* is an interior second-best allocation under Assumption 2.1, then it is unique and solves the following system of equations with $d = \mu^*(1 - w_L/w_H)/(\alpha - \mu^*)$.*

$$u_C^L = \frac{p_C}{w_L}(1 + d) \qquad u_D^L = \frac{p_D + e/\lambda^*}{w_L}(1 + d)$$

$$u_C^H = \frac{p_C}{w_H} \qquad u_D^H = \frac{p_D + e/\lambda^*}{w_H}$$

$$Y_L^* = r + \gamma(p_C C_L^* + p_D D_L^*) + (1 - \gamma)(p_C C_H^* + p_D D_H^*) - w_H(u^H - u^L)(1 - \gamma)$$
$$Y_H^* = r + \gamma(p_C C_L^* + p_D D_L^*) + (1 - \gamma)(p_C C_H^* + p_D D_H^*) + w_H(u^H - u^L)\gamma$$

$$\lambda^* = \frac{\alpha}{w_L} + \frac{1 - \alpha}{w_H}, \qquad \mu^* = \alpha(1 - \gamma)\frac{w_H}{w_L} - (1 - \alpha)\gamma. \tag{2.21}$$

Proof. If A^* is an interior second-best allocation, then it satisfies conditions (2.14) and (2.17) with δ set to zero. Then λ^* and μ^* are uniquely determined and strictly positive. For given values of λ^* and μ^*, (2.12), (2.15), (2.13), and (2.16) uniquely determine the consumption levels (uniqueness is established by Lemma 2.4). Output requirements follow from the binding resource constraint (2.2) combined with the binding incentive constraint (2.4). □

The conditions for high-type consumption levels are almost identical to the corresponding first-best conditions (2.19) and (2.20). The subtle but important difference lies in the Lagrangian multiplier λ, which is different in first- and second-best and, most importantly, features different comparative statics properties.

The consumption levels of low types are distorted downwards, i.e., the labour choice is distorted in favour of leisure. The distortion is captured by d and is higher if μ^* is higher or the difference in productivities is larger.

Taxation

Recall that a tax system $\tau = (t_C, t_D, T)$ consists of an income tax function T and a linear commodity tax rates, t_C and t_D. Tax system τ implements allocation A if

1. For any θ, $(C_\theta, D_\theta, Y_\theta) \in \operatorname{argmax}_{(C,D,Y)}\{u(C,D) - Y/w_\theta \text{ s.t. } q_C C + q_D D \leq Y - T\}$
2. $\gamma(Y_L - p_C C_L - p_D D_L) + (1-\gamma)(Y_H - p_C C_H - p_D D_H) - r \geq 0$.

It is straightforward to see that a tax system with linear commodity taxes can only implement allocations in which the marginal rates of substitution between the two consumption good are equal across all agents.

Lemma 2.5. *If tax system τ implements allocation A, then*

$$\frac{q_D}{q_C} = MRS = \frac{u_D(C_L, D_L)}{u_C(C_L, D_L)} = \frac{u_D(C_H, D_H)}{u_C(C_H, D_H)}$$

Proof. Individual choices, as defined in (2.8), need to satisfy $u_D^\theta / u_C^\theta = q_D/q_C$. □

The linearity of commodity taxation is generally restrictive. But the following two lemmas show that, in the model at hand, more general commodity tax system could not improve upon the linear systems; both the first- and the second-best allocation are implementable with tax systems as defined above.

Lemma 2.6. *There exists a tax system $\tau = (t_C, t_D, T)$ with $T : (w, Y) \mapsto T(w, Y)$ that implements the first-best allocation $A^F = (C_L^F, D_L^F, Y_L^F, C_H^F, D_H^F, Y_H^F)$.*

Proof. Fix some t_C, t_D such that $q_D/q_C = (p_D + e/\lambda^F)/p_C$. Define the income tax function for $w \in \{w_L, w_H\}$ to be

$$T(w,Y) = \begin{cases} Y, & w = w_L \text{ and } Y \neq 0 \\ -(q_C C_L^F + q_D D_L^F), & w = w_L \text{ and } Y = 0 \\ Y, & w = w_H \text{ and } Y \neq Y_H^F \\ Y_H^F - (q_C C_H^F + q_D D_H^F), & w = w_H \text{ and } Y = Y_H^F \end{cases}$$

It remains to show that an agent who solves the individual maximisation problem (2.8) chooses exactly the bundle that is intended for her type.

(a) Low types. For low-type agents it is never optimal to choose $Y \neq 0$. So their problem reduces to $\max_{C,D} u(C,D)$ s.t $q_C C + q_D D \leq q_C C_L^F + q_D D_L^F$. Necessary and sufficient conditions for a solution are $u_D/u_C = q_D/q_C$ and $q_C C + q_D D = q_C C_L^F + q_D D_L^F$. By construction, the first-best consumption levels (C_L^F, D_L^F) satisfy these conditions.

(b) High types. For high-type agents, it is never optimal to choose $Y \neq Y_H^F$. So their problem reduces to $\max_{C,D} u(C,D)$ s.t $q_C C + q_D D \leq q_C C_H^F + q_D D_H^F$. Necessary and sufficient conditions for a solution are $u_D/u_C = q_D/q_C$ and $q_C C + q_D D = q_C C_H^F + q_D D_H^F$. By construction the first best consumption levels (C_H^F, D_H^F) satisfy these conditions. □

Lemma 2.7. *There exists a tax system with $T : Y \mapsto T(Y)$ that implements the second-best allocation $A^* = (C_L^*, D_L^*, Y_L^*, C_H^*, D_H^*, Y_H^*)$.*

Proof. Fix some t_C, t_D such that $q_D/q_C = (p_D + e/\lambda^*)/p_C$. Let $T_\theta := Y_\theta - q_C C_\theta^* - q_D D_\theta^*$, $\theta \in \{L, H\}$. Define the income tax function $T(Y)$ as

$$T(Y) = \begin{cases} T_L, & Y = Y_L^* \\ T_H, & Y = Y_H^* \\ Y, & \text{otherwise.} \end{cases}$$

An agent who faces the individual maximisation problem defined in (2.8) will never choose $Y \notin \{Y_L^*, Y_H^*\}$. Thus the agents' problem can be decomposed into two steps: Step 1 is to solve for $\hat{u}(B) := \max_{C,D} u(C,D)$ s.t. $q_C C + q_D D = B$. Step 2 is to solve for $\max_Y \hat{u}(B) - Y/w$ subject to $B = (Y_L^* - T_L)$ if $Y = Y_L^*$, $B = (Y_H^* - T_H)$ if $Y = Y_H^*$, and $B = 0$ otherwise.

Necessary and sufficient conditions for a solution of step 1 are $u_D/u_C = q_D/q_C$ and $q_C C + q_D D = B$. By construction the conditions are met by (C_L^*, D_L^*) if $B = Y_L^* - T_L$ and by (C_H^*, D_H^*) if $B = Y_H^* - T_H$. Consequently, step 2 is equivalent to choosing between (C_L^*, D_L^*, Y_L^*) and (C_H^*, D_H^*, Y_H^*). As the second-best allocation satisfies the incentive compatibility constraints (2.4) and (2.3), each agent does indeed choose the bundle intended for her type. □

Proof of Proposition 2.1. Under a normalised tax system, $q_D = p_D + t_D$ and $q_C = p_C$. Combining Proposition 2.4 and Lemma 2.5 shows that $MRS = (p_D + e/\lambda^F)/p_C = q_D/q_C = (p_D + t_D^F)/p_C$. It follows that $t_D^F = e/\lambda^F$. From (2.7), $\lambda^F = (1-\alpha)/(w_H(1-\gamma))$. The comparative statics results follow from first derivatives. □

Proof of Proposition 2.2. Under a normalised tax system, $q_D = p_D + t_D$ and $q_C = p_C$. Combining Proposition 2.5 and Lemma 2.5 shows that $MRS = (p_D +$

$e/\lambda^*)/p_C = q_D/q_C = (p_D + t_D^*)/p_C$. It follows that $t_D^* = e/\lambda^*$. From (2.21), $\lambda^* = \alpha/w_L + (1-\alpha)/w_H$. The comparative statics results follow from first derivatives. □

Appendix 2.B Second-Best Corner Solutions

A *second-best corner solution* is a second-best allocation with $Y_L = 0$. In this section I show that for some parameters this is the relevant case. I then claim that a corner solution does not change at all if α is increased (Proposition 2.6). Thereby I extend the comparative statics properties of Pigouvian taxation to instances of corner solutions (Corollary 2.2). Proposition 2.6 builds on a conjecture that generalises Lemma 2.3. Unfortunately, I can only partially verify that conjecture (Lemma 2.10).

Lemma 2.8. *If A is a second-best allocation and*

$$\alpha \geq \frac{\gamma}{1-\gamma} \frac{1}{\frac{w_H}{w_L} - 1}, \tag{2.22}$$

then A is a corner solution, i.e. $Y_L = 0$.

Proof. By contradiction, assume A is an interior solution. Then it satisfies conditions (2.12) to (2.17) with $\delta = 0$. Hence $\mu = (1-\gamma)\alpha w_H/w_L - \gamma(1-\alpha)$ by (2.18). But then the condition on parameters stated in the Lemma implies $\alpha - \mu \leq 0$. This, however, contradicts (2.12). □

Remark 2.1. Bierbrauer and Boyer (2010) exclude corner solutions in their comparative statics analysis by assuming $1 > (1-\gamma)w_H/w_L$. Their inequality always holds if (2.22) is not satisfied, but the converse is not true. Hence, I do not expect (2.22) to be a necessary condition for a corner solution.

To proceed, let me introduce some convenient notation.

Definition 2.1. Define $A(\alpha)$ to be a second-best allocation, in which the welfare weight is given by α and all other parameters are fixed. For $\alpha' < \alpha''$ and $\theta \in \{L, H\}$ define $U'_\theta := U_\theta(A(\alpha'))$, $U''_\theta := U_\theta(A(\alpha''))$, and $dU_\theta := U''_\theta - U'_\theta$.

Quite intuitively, if the taste for redistribution increases, low-type agents receive higher utility. At the same time, high-type agents have to receive lower utility because someone has to pay for the increase in U_L. The following lemma formalises this intuition. Notice that also a zero-change in utility is possible.

Lemma 2.9. *If α increases from α' to α'', then*

$$dU_L \geq 0 \geq dU_H.$$

Proof. From Definition 2.1 it follows, in particular,

$$\alpha' U'_L + (1-\alpha')U'_H \geq \alpha' U''_L + (1-\alpha')U''_H \Rightarrow 0 \geq \alpha' dU_L + (1-\alpha')dU_H \quad (2.23)$$
$$\alpha'' U''_L + (1-\alpha'')U''_H \geq \alpha'' U'_L + (1-\alpha'')U'_H \Rightarrow \alpha'' dU_L + (1-\alpha'')dU_H \geq 0$$

Summing up the two implied inequalities yields $dU_L \geq dU_H$. Next, suppose by contradiction that $dU_H > 0$, then $dU_L > 0$, but that contradicts (2.23). The proof of $dU_L \geq 0$ is analogous. □

Lemma 2.3 shows that the incentive constraint of high-type agents binds at interior second-best allocations, given Assumption 2.1. Its proof does not work for corner solutions, though. By contrast, the following lemma does hold for corner solutions, albeit under more restrictive conditions on parameters.

Lemma 2.10. *Suppose $\alpha \geq \gamma$. If A is a second-best allocation, then the incentive constraint* (2.4) *for the high-type agents is binding at A.*

Proof. I first show that marginal utility is lower for high- than for low-type agents. Then I show that a marginal redistribution of C from high- to low-type agents increases welfare and hence needs to be ruled out by a binding incentive constraint. Otherwise the allocation cannot be second-best. In term of notation, recall that $u^\theta = u(C_\theta, D_\theta)$.

Claim: $u_C(C_L, D_L) > u_C(C_H, D_H)$

Case 1, $D_H \leq D_L$: then $C_H > C_L$, because, by Lemma 2.2, $u(C_L, D_L) < u(C_H, D_H)$. Decreasing marginal utility and a positive cross derivative $u_{CD} \geq 0$ then imply $u_C(C_L, D_L) > u_C(C_H, D_L) \geq u_C(C_H, D_H)$, hence the claim holds.

Case 2A, $D_H > D_L$, $C_H \leq C_L$: then, similar to Case 1, $u_D(C_L, D_L) > u_D(C_L, D_H) \geq u_D(C_H, D_H)$. At an optimal allocation, $u^L_D/u^L_C = u^H_D/u^H_C$, hence $u^L_D > u^H_D$ implies $u^L_C > u^H_C$, as claimed.

Case 2B, $D_H > D_L$, $C_H > C_L$: as u is strictly concave,

$$u^H - u^L < (C_H - C_L)u^L_C + (D_H - D_L)u^L_D, \text{ and}$$
$$u^L - u^H < (C_L - C_H)u^H_C + (D_L - D_H)u^H_D$$

need to hold.[19] Rearranging the second inequality gives, in combination with the

[19]In general, if a continuously differentiable function f is strictly concave over an open, convex subset of $\mathbb{R}^n$, then $f(x) - f(x^0) < \sum_i f_{x_i}(x^0)(x_i - x_i^0)$, for all x, x^0 from that subset. See Sydsaeter et al. (2008, Theorem 2.4.1) for a textbook reference.

first,

$$u_D^H(D_H - D_L) + u_C^H(C_H - C_L) < u^H - u^L < (C_H - C_L)u_C^L + (C_H - C_L)u_D^L.$$

$(C_H - C_L), (D_H - D_L) > 0$, thus $u_D^H < u_D^L$ or $u_C^H < u_C^L$. If one of these two inequalities holds, the other one must hold as well, otherwise $u_D^L/u_C^L = u_D^H/u_C^H$ cannot be true. This completes the proof of the claim.

Now, suppose that the lemma is false, then $u(C_H, D_H) - Y_H/w_H > u(C_L, D_L) - Y_L/w_H$ at a second-best allocation. Then there exist dC_H, dC_L satisfying $dD_H = -dC_L\gamma/(1-\gamma) < 0$, such that the incentive constraint still holds, i.e. that $u(C_H + dC_H, D_H) - Y_H/w_H > u(C_L + dC_L, D_L) - Y_L/w_H$. The modified allocation $(C_L + dC_L, D_L, Y_L, C_H + dC_H, D_H, Y_H)$ is also feasible by construction (and still satisfies low-type agents' incentive constraint). For $dD_L \to 0$, the change in welfare is approximately

$$dW \approx \alpha u_C^L dC_L - (1-\alpha)u_C^H dC_L\gamma/(1-\gamma) = dC_L\left(\alpha(1-\gamma)u_C^L - (1-\alpha)\gamma u_C^H\right). \tag{2.24}$$

If $\alpha \geq \gamma$, and $u_C^L > u_C^H$ as claimed, then $\alpha(1-\gamma)u_C^L - (1-\alpha)\gamma u_C^H > 0$ and welfare increases. Hence, a contradiction. □

If the welfare function is utilitarian ($\alpha = \gamma$) or exhibits an even stronger tendency to redistribute in favour of the low-type agents, high-type agents incentive constraint must be binding. This is not a necessary condition, though. Equation (2.24) shows that even with $\alpha < \gamma$ a slack incentive constraint would be impossible, provided that u_C^L is sufficiently greater than u_C^H. In fact, I believe that the constraint is binding whenever Assumption 2.1 is satisfied.

Conjecture 2.1. If A is a second-best allocation under Assumption 2.1, then high types' incentive constraint is binding at A.

The following proposition is the main result of the current section. If conjecture 2.1 holds, the proposition and its two corollaries extend to all parameters satisfying Assumption 2.1.

Proposition 2.6. *Suppose $\alpha' \geq \gamma$. Let $A(\alpha')$ be a second-best corner solution. Then $A(\alpha'') = A(\alpha')$ for all $\alpha'' > \alpha'$.*

Proof. Claim 1: At a corner solution all agents have the same utility level.
By Lemma 2.10, high-type agents incentive constraint is binding. Adding $-(\gamma D_L + (1-\gamma)D_H)e$ to the binding incentive constraint gives $U_H(A(\alpha')) = u(C_H, D_H) - Y_H/w_H - (\gamma D_L + (1-\gamma)D_H)e = u(C_L, D_L) - (\gamma D_L + (1-\gamma)D_H)e = U_L(A(\alpha'))$.

Claim 2: $U_L(A(\alpha'')) = U_H(A(\alpha''))$.
From Lemma 2.9, $U_L(A(\alpha'')) - U_L(A(\alpha')) \geq U_H(A(\alpha'')) - U_H(A(\alpha'))$. Given Claim 1 this reduces to $U_L(A(\alpha'')) \geq U_H(A(\alpha''))$. The incentive constraint of high types implies, though, that $U_H(A(\alpha'')) \geq U_L(A(\alpha''))$. Hence $U_L(A(\alpha'')) = U_H(A(\alpha''))$.
Claim 3: $A(\alpha'') = A(\alpha')$.
Suppose the opposite, then $U_L(A(\alpha'')) > U_L(A(\alpha'))$ by Lemma 2.9 and the fact that solutions are unique (if they exist). But then Claims 1 and 2 imply that also $U_H(A(\alpha'')) > U_H(A(\alpha'))$. This contradicts Lemma 2.9. □

Increasing the welfare weight of low-type agents does not change the allocation if low-type agents already provide zero output. The only way to increase their utility is to increase their consumption. But then high-types incentive constraint can no longer be satisfied. Thus the limits of redistribution (under information constraint) are met, once all output is produced by high-type agents:

Corollary 2.1. *Suppose $\alpha' \geq \gamma$. Let $A(\alpha')$ be a second-best corner solution. Then the Rawlsian allocation $A^R = \lim_{\alpha \to 1} A(\alpha)$ is equal to $A(\alpha')$. Also, the second-best Pareto-frontier has a kink at $[U_L(A^R), U_H(A^R)]$ if $\alpha' < 1$.*

Notice that it is possible that the Rawlsian allocation is not a corner solution. Put differently, (second-best) redistribution can hit its very limit well before low-type agents provide zero output.

Yet, if for some $\alpha < 1$, low-type agents' output does equal zero, then, consequently, the comparative statics of Pigouvian taxation are also zero:

Corollary 2.2. *Suppose $\alpha \geq \gamma$. Let $A(\alpha)$ be a second best allocation with $Y_L = 0$. Let $t_D(\alpha)$ be the dirty good tax of a normalised tax system that implements $A(\alpha)$. Then $t_D(\alpha') = t_D(\alpha)$ for all $\alpha' \geq \alpha$.*

A rational voter first decides what party he believes will benefit him most; then he tries to estimate whether this party has any chance of winning.

Anthony Downs, An Economic Theory of Democracy, 1957

Proportional representation is usually not free from important restrictions – most notably a minimum vote threshold that a party must pass to become eligible for seats in parliament. Falling short of such a threshold means that a vote for a party is 'wasted' or 'lost' because it does not count toward the distribution of seats in parliament.

Meffert & Gschwend, 2011

3 Investing Your Vote – On the Emergence of Small Parties

3.1 Introduction

In parliamentary systems with proportional representation, the seat share of parties in parliament is approximately proportional to their vote share. But to enter parliament in the first place, a party has to receive some minimal amount of votes, called the election threshold. An *implicit* threshold is present in any system and amounts to the minimum vote share necessary to get a single seat. The implicit threshold generally depends on the number of seats in parliament, but also on institutional details and the distribution of votes. Beyond that, many countries have a more restrictive *explicit* threshold. It typically ranges from 2% of votes in Israel and Denmark to 5% in countries like Belgium or Germany. Russia (7%) and Turkey (10%) have particularly high thresholds.

Rational voters supposedly cast their ballot only for parties that they anticipate to enter parliament. With respect to a majoritarian system, Duverger (1951, 1954) was the first to point out that votes are often concentrated on two dominant parties.[1] Downs (1957, p. 48) notes that a rational voter needs to estimates a party's chances of winning before voting for it. The same reasoning applies to proportional systems with election thresholds. However, in that case, the relevant question is whether a party enters parliament and, if so, whether it joins a

[1] See Da Silva (2006) for the recent discussion of Duverger's Law in the political literature.

governing coalition. Following this logic, a vote for a party that fails the threshold is a wasted vote (Meffert and Gschwend, 2011, p. 3).

Despite this reasoning, small parties often do get a considerable amount of votes while failing the threshold. Extreme examples include the Russian legislative election in 1995, in which almost half of all votes were split among parties falling short of the threshold. Similarly, in the 2002 election in Turkey, about 45% of votes were unrepresented in parliament.[2]

Existing models of political competition typically neglect this issue. To fill this gap, we consider the following setting: Two incumbent parties split the majority of votes between themselves. A new party tries to enter parliament, but is expected to fall short of the threshold. Given this expectation, why should anyone vote for the new party?

Using a game-theoretic approach, we show that for some voters it is in fact strategically optimal to vote for a party that does not enter parliament at the current election. But rather than wasted, their vote is invested – trading off influence at the current election against future policy returns. On the other hand, some voters strategically vote for an incumbent party despite preferring the new party's platform.

Voters' preferences over policies depend on their type and there is aggregate uncertainty about the type distribution. Given their priors, citizens vote strategically to maximize their expected utility. Voting takes place in two consecutive periods. In the second period, voters update their beliefs about the type distribution, based on the first election result. This gives rise to (at least) two additional motives to vote for a party in period one. First, these votes may be a signal to other supporters of the small party to vote for the party in the next election, because its probability of entering parliament is higher under the revised prior than under the initial prior. Second, these votes may induce incumbent parties to change their platform due to updated beliefs about the number of voters that could be attracted by a given platform. In a related work, Castanheira (2003) focuses on the second motive. In our analysis, we explore the first one and take party positions as exogenously given. Notice that these two motives predict different voting patterns over time. In particular, adapting their platforms allows incumbent parties to keep entrant parties out of power; in the last period no voter supports the entrant parties. If the entrant party is successful in our model, its vote share is increasing over time and it might become part of the government.

We characterize an equilibrium in which supporters of the new party make their second-period ballot decision conditional on the results of the first election. In

[2]See www.bbc.co.uk/turkish/secim2002/election_results_en.shtml (retrieved March 2010) and White, Wyman, and S. Oates (1997, Table 2), respectively.

particular, they coordinate on the new party if the party's vote share has exceeded some endogenous threshold. Otherwise, the entrant party will receive no votes in the second period. Figure 3.1 outlines this *investing equilibrium*. Our findings rationalize a voting behavior that looks wasteful upon first sight, but turns out to be strategically optimal. Hence, there is no coordination failure in the sense of inconsistent beliefs about equilibrium play.

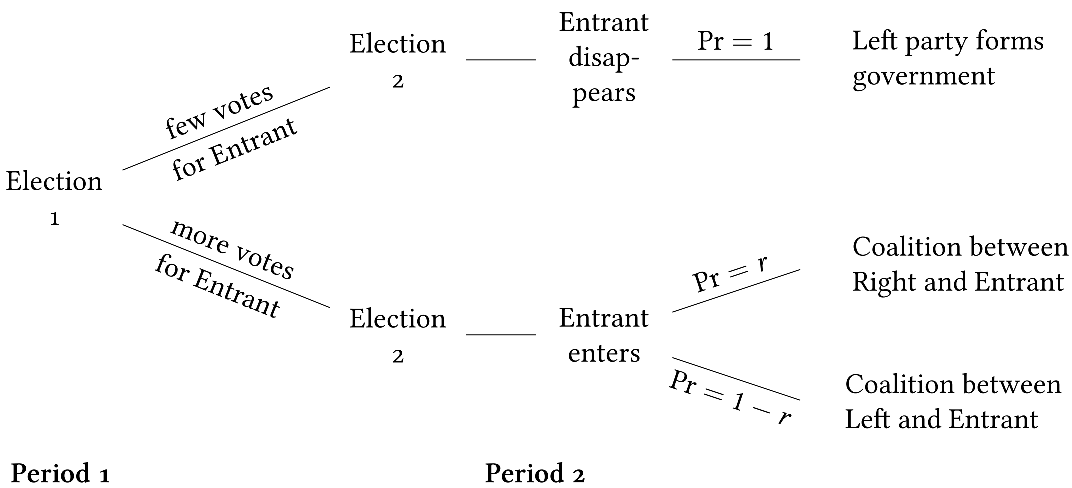

Figure 3.1: Three parties, Left, Right, and Entrant, compete in two consecutive elections. In period 1, the Entrant party fails the election threshold. Yet, if the Entrant's result in the first election is sufficiently strong, the party enters parliament in period 2. r denotes the fraction of supporters of the right party.

In the investing equilibrium, the first period serves as a costly coordination stage for supporters of the entrant party. Passionate voters reveal their type by voting for the entrant knowing that it will fail the election threshold. The costs arise from giving up influence on parliament's current composition.

Polls could offer a possibility to reveal types without these cost. Yet, the absence of such a cost is also a major drawback, as the signal becomes less credible. If polls were able to serve the signaling purpose nonetheless, the electoral patterns motivating our model are unlikely to occur. The costs of using the vote as a signal seem necessary in order to make the signal credible.

Our model is consistent with the paradigm of instrumental voting; voters care exclusively about the policy impact of their vote. This is in line with empirical findings summarized in Section 3.3 suggesting that voters do use their votes strategically. An alternative explanation goes under the heading of 'expressive voting': by voting for a particular party, a voter might want to express certain opinions regardless of the effects on election outcomes and gain utility just from the act of doing so.[3] Our model suggests that voting behavior that looks expressive at first sight could well be instrumental in the long run.

The rest of this chapter is organized as follows. Section 3.2 provides empirical motivation. Section 3.3 discusses related literature. Then, Section 3.4 sets up the model and Section 3.5 details on coalition formation and policy choice. Section 3.6 gives the main results of the chapter. Section 3.7 discusses sincere voting. We conclude in Section 3.8. All proofs are relegated to Appendix 3.A.

3.2 Empirical Motivation

In a survey two months before the 2009 German national elections, 6% of the respondents considered voting for the Pirate Party, which ran for the first time for national parliament. In addition, 58% reported not knowing the party.[4] The expected vote share for the Pirate Party was 2% according to polls, i.e., well below the German threshold of 5%. In the election, the Pirate Party received in fact 2% of the votes. Hence 2% of the electorate voted for the Pirate Party despite its low chance to enter parliament. On the other hand, the Pirate Party voters had reasons to believe that the number of supporters of the Pirate Party is in fact higher than 2%, namely 6% plus a potential fraction of people who did not know the party yet. We offer an explanation for why these 2% indeed had a strategic incentive to vote for the Pirate Party. Our model suggests that the early Pirate Party voters played an important role for the subsequent success of the Pirate Party in recent elections at state level.[5]

In Austria, a new right-wing party, the BZÖ, was founded in April 2005 by former members of the FPÖ, another right-wing party. According to polls for the general election 2006, the BZÖ was expected to receive a vote share of just around 4%, which is the election threshold in Austria.[6] So there was high uncertainty on the party's probability to enter parliament. With an election result of 4.1%, the BZÖ did succeed after all.[7] Two years later, in the general election of 2008, the

[3] Hamlin and Jennings (2011) provide an introduction to this line of thought.

[4] The survey was conducted by TNS Emnid for Cicero, a German magazine, and published on the 23rd of July 2009.

[5] The Pirates received 8.9% in the state of Berlin in 2011, and 7.4% in Saarland, another German state, in 2012. They entered parliament in both states. For a reference, see www.dw.de/cdu-holds-its-own-in-saarland/a-15837387-1, retrieved May 2012.

[6] A party enters parliament in Austria, if it wins more than 4% of the votes or if it crosses a particular threshold of the valid votes in one of the 43 regional electoral districts. See also Solsten (1994).

[7] Meffert and Gschwend (2010) provide extended discussion of the 2006 election.

BZÖ more than doubled its vote share to 10.7%. It seems likely that a significant fraction of voters preferred the BZÖ already in 2006, but refrained from voting for the BZÖ for fear of wasting their vote, as the BZÖ might have missed the election threshold. In 2008, these voters went for the BZÖ, as there was less uncertainty about the BZÖ entering parliament. Our model can explain such a jump in the vote share without a change in voters' preferences.[8]

The German Green Party was founded in 1980 and failed to pass the 5% threshold in the national election of the same year. They obtained a significant fraction of 1.5% of the votes, though. In the 1983 national election, they successfully passed the threshold. We argue that this success would not have been possible, had they received insufficiently many votes in 1980. Hence, their 1980 voters did not waste their votes but effectively invested them.

On the other hand, our model predicts that if a new party falls short of an endogenous threshold in the first election, its support disappears. There are numerous examples of such developments. Consider, for instance, Proud of the Netherlands, a party founded in 2009. Although early polls showed the party to win up to 10%, it gained a vote share of only 0.6% or 52,937 votes in its first general election in 2010. This is slightly below the Dutch election threshold of 1/150 to enter parliament. Although the party merged with another one to compete in the 2012 general election, its election results decreased significantly to 0.1% or 7,363 votes. Another example is Action for Democratic Progress, a party founded in 1968 in Germany to protest against the German Emergency Acts. The party participated in the federal election of 1969, but gained a vote share of only 0.6%, i.e., less than 200,000 votes. The very same year, the party was dissolved.[9]

3.3 Related Literature

We contribute to the literature on communicative voting, in which voting decisions are used as a signaling device. We introduce a model which specifically captures the uncertainty about the support of a new party. Our setting fits the multi-party election system with proportional representation which is typical for continental Europe. Parties have to pass an election threshold to enter parliament. These thresholds create non-trivial voting incentives, which are at the heart of our

[8]Notice, though, that our analysis has to be changed slightly to a fit the BZÖ case. We present a more clear-cut version in which, along the equilibrium path, the small party actually has no chance of entering parliament in the first period.

[9]See: www.verkiezingsuitslagen.nl/Na1918/Verkiezingsuitslagen.aspx?VerkiezingsTypeId=1, www.dutchnews.nl/news/archives/2012/06/new_dpkp_party_marks_a_turning_1.php, www.bundeswahlleiter.de/de/bundestagswahlen/fruehere_bundestagswahlen/btw1969.html; all retrieved June 20, 2013.

analyses. Related contributions differ not only with respect to the relevant setting but also with respect to the modeling of uncertainty and the addressees of the signal.

Razin (2003) employs a one-period model with common values, in which election results are informative for the legislature about some shock affecting the country. This allows candidates better to adapt policies to the state of the world. In our model, the signal is aimed at the electorate rather than the legislature. We analyze competition between three parties (as opposed to two, as in ibid.). Therefore we can address the coordination problem of voters that arises if each party needs some minimal support to have any effect on policy.

Piketty (2000) allows voters to signal their information about the state of the world by electing specific candidates. They want to influence future voting behavior. The two-period horizon is similar to our setting. Yet, in Piketty's model, voters are uncertain about the optimal policy and the candidate they prefer. In our model, every voter knows with certainty which party she prefers, but is uncertain about the preferences of other citizens. The problem of coordination is not prominent in Piketty's paper, because only two alternatives are offered in each of the elections; two candidates compete in the first election and the winner is challenged by a third candidate in the second election.

In the model proposed by Castanheira (2003), the position of the median voter on a one-dimensional policy space is unknown. Four parties compete under majoritarian rule in two consecutive elections. After the first election, voters update their beliefs about the true position of the median voter. Votes for an extreme party signal that the median voter is at a more extreme position. Given updated beliefs, parties might reallocate their platforms for the next election. Voting for "losers" is a rational strategic choice so as to induce parties to change their policy offers.

In our contribution, voters have incentives to vote for "losers" as well; however, they do not send a signal to parties but rather to other voters. The question is not whether parties adjust their platforms, but whether they have sufficient support to enter parliament.[10] Ex ante there is no clear answer to this question, because there is uncertainty about voters' types.

In Meirowitz and Shotts (2009), there is uncertainty about voters' preferences, as well. Two candidates compete in two consecutive elections. After the first election beliefs are updated and candidates might reallocate their platforms for the next election. They show that for large electorates the signaling effect dominates

[10]We take it as given that three parties compete, but presume one to be new and small. The question whether a new party emerges at all is addressed by Palfrey (1984). In his model, incumbent parties may anticipate entry of new parties and adopt their policies in advance in order to prevent entry.

any pivot effects. The reason is that the probability of being pivotal converges to zero very fast while every vote can signal at least some information.

In all of the above papers, voters' types are independently and identically distributed. We assume aggregate uncertainty in the sense that the joint distribution of voters' types is random. This distribution is drawn according to a commonly known second-order distribution. In contrast to the first approach, this aggregate uncertainty does not depend on the number of voters. In particular, our approach ensures uncertainty even for a large number of voters. For the same reason, Bierbrauer and Hellwig (2011) use this approach in a public economics setting.

Coalition formation plays a central role in proportional elections. In contrast to the aforementioned contributions, we include coalitional bargaining in our model. Therefore, sub-majority parties in parliament can also influence policy. We build upon the proto-coalition bargaining model of Baron and Diermeier (2001).[11] Their two-dimensional policy space is able to capture a situation that is typical for the emergence of a new party. Namely, incumbent parties differ with regard to traditional issues, but have similar views on a new issue such that parliament does not reflect different opinions on that new issue. A new party emerges for the very reason to represent this new issue or some hitherto unrepresented opinion. Examples include green parties which promoted environmental protection or the Pirate Party with its focus on digital rights.

We adapt the Baron and Diermeier (ibid.) setting in various respects. First, we assume that the distribution of voters' preferences is uncertain. Second, we include an intensity parameter in voters' preferences. Third, we introduce a second election period. This gives rise to communicative voting motives that drive our model. Furthermore, we make specific assumptions about the peaks of the parties and the distribution of voters in order to capture the particular situation that makes the election threshold interesting. We also slightly adapt the bargaining procedure.

In terms of empirical literature, we relate to studies that investigate strategic voting.[12] Bargsted and Kedar (2009) examine elections in Israel in 2006. They

[11]Further prominent contributions on coalitional bargaining with a focus on political economy include Baron and Ferejohn (1989), and Austen-Smith and Banks (1988). They do not match our setting as well as Baron and Diermeier (2001), though. Baron and Ferejohn (1989) concentrate on the bargaining between legislators who serve their own district, but do not include a general election stage. Austen-Smith and Banks (1988) propose a more complete model of the political process. Similar to our model, they consider three parties, strategic voters, and endogenous coalition formation. However, they focus on parties' platform choice, whereas our main attention is on voters' decisions. Furthermore, their one-dimensional policy space is not rich enough for our purposes.

[12]Meffert and Gschwend (2011, pp. 2-3) provide a long list of evidence for strategic voting in different countries.

find that the set of perceived coalition possibilities had a profound influence on voter ballots in the sense that it induced non-sincere voting, i.e., voting for parties other than the most preferred one. Baron and Diermeier (2001, p. 934) provide another example for strategic voting behavior in which supporters of a larger party vote for a small party to make them enter parliament. Similarly, Shikano, Herrmann, and Thurner (2009, p. 634) "show that voters' preferences, rather than mapping directly into party choice, are affected by their expectations on small parties' re-entry chances." Meffert and Gschwend (2011) conduct a laboratory experiment embedded in a state-level election in Germany in 2006. Participants in their study were able to form meaningful expectations about election outcomes and possible coalitions. McCuen and Morton (2010) also show strategic voting under proportional representation in a laboratory experiment.

Sobbrio and Navarra (2010) explicitly investigate communicative motives in voting. For a sample of 14 European countries, they find that voting for a "sure loser" is associated with higher education. They see this as evidence that voters consider more than one election period in their voting decision.

Finally, Kricheli, Livne, and Magaloni (2011) propose a model of civil protest against repressive regimes. In a first period, a small group of protesters take to the streets to signal their type. If their protest reaches a critical size, it triggers a mass protest able to overthrow the dictator. Kricheli, Livne, and Magaloni (ibid.) back their theory with an empirical analysis, according to which more repressive regimes tend to be more stable, but protests that do occur pose a higher threat to them. The more repressive a regime, the higher the costs of failed protest. Initial protest is then less likely to occur but constitutes a stronger signal.

3.4 Model

Society consists of n citizens and has to make a two-dimensional policy choice

$$x = (a, b) \in X = [0, 1] \times [0, 1].$$

Three parties, called *Left* (L), *Right* (R), and *Entrant* (E) compete for proportional representation in parliament. In the ballot, each citizen casts a vote for one of the parties. Parties receiving less than $\tau < n/4$ votes do not get any seats in parliament. The seat share of a party that enters parliament equals its vote share corrected for votes unrepresented in parliament.[13] The parties entering parliament bargain to form a government (coalition) which needs to hold at least a

[13] For the sake of tractability, we abstract from non-divisibility issues of seats. Hence, a party's seat share equals the number of its votes divided by the sum of votes of those parties that enter parliament.

simple majority, i.e., half the seats. Government chooses policy. The model consists of two consecutive election periods. Fundamentals stay constant over time, but beliefs may change.

Our analysis focuses on voters' ballot decisions rather than on the question whether to vote at all.[14] Voters in our model have a strictly positive benefit from voting, so our results are robust if voting costs are sufficiently small. Alternatively, consider voters who have already made their way to the voting booth. Now they just need to decide which party to vote for. This is the decision we analyze.

3.4.1 Parties

Each party $\mathcal{J} \in \{L, R, E\}$ is characterized by its peak $(a_{\mathcal{J}}, b_{\mathcal{J}})$. Each period, parties gain utility $u^{\mathcal{J}}(x)$ from policy x, and might receive benefits $s_{\mathcal{J}}$ from holding office, like positions in the cabinet or seats on the boards of public companies.[15] Party $\mathcal{J}$'s overall utility in any given period is

$$U^{\mathcal{J}}(x, s_{\mathcal{J}}) = u^{\mathcal{J}}(x) + s_{\mathcal{J}} = -(a - a_{\mathcal{J}})^2 - (b - b_{\mathcal{J}})^2 + s_{\mathcal{J}}\,.$$

The total spoils of holding office are $S > 0$. The parties split these spoils among them such that $\sum_{\mathcal{J}} s_{\mathcal{J}} = S$. In addition, $s_{\mathcal{J}} \geq 0$ if $\mathcal{J}$ is not in government: parties outside the governing coalition cannot be forced by coalition members to pay the ruling parties; thus a negative value of $s_{\mathcal{J}}$ is allowed only for coalitions members. Note that the specification allows for unlimited side payments among members of the government.

We interpret the first policy dimension as a traditional issue ('left vs. right') over which the incumbent parties L and R struggle. L prefers $a = 0$, hence $a_L = 0$, while $a_R = 1$. The second policy dimension, b, is new in the sense that it has not been much of an issue in the past. This is reflected in identical peaks, namely $b_L = b_R = 0$. The entrant party E disagrees on the new issue and likes to initiate a change in b by offering a political alternative, thus $b_E = 1$. Let it emerge from the left spectrum such that $a_E = 0$.[16] To summarize, we have

$$(a_L, b_L) = (0, 0), \qquad (a_R, b_R) = (1, 0), \qquad (a_E, b_E) = (0, 1).$$

[14] The paradox of not voting is the focus of a related signaling-voting model recently proposed by Aytimur, Boukouras, and Schwager (2013).

[15] For a detailed discussion of $s_{\mathcal{J}}$ see Baron and Diermeier (2001, p. 935).

[16] Our results would be the same for a symmetric variation of the model in which the new party emerged from the right spectrum. The new party could also position itself somewhere between left and right. This would require a richer model in terms of voter preferences, though, and would complicate the model without enhancing the insights with respect to the investment equilibrium.

Parties cannot make binding commitments. After each election, a bargaining process determines the policy in the current period. Only parties represented in parliament participate in the bargaining. Section 3.5 provides details about bargaining and policy choice.

3.4.2 Voters

Voters have preferences on policies. They differ, in particular, with respect to the salience of the two policy dimensions. A voter of type $\theta^i \equiv (a^i, b^i, \alpha^i) \in \Theta$ has the per-period utility function

$$u(x; \theta^i) = -\alpha^i(a - a^i)^2 - (1 - \alpha^i)(b - b^i)^2,$$

where (a^i, b^i) corresponds to the peak with respect to the two policy dimensions and $\alpha^i \in (0, 1)$ describes the relative salience of the two dimensions from the perspective of the voter. There is a common discount factor $\delta > 0$. The second period can alternatively be interpreted as some long-run steady state spanning multiple periods. In that case, δ serves as a weighting factor and might be above 1.

To keep the model as simple as possible, we consider only four different types of voters: *left* (θ^L), *right* (θ^R), *passionate* (θ^P) and *moderate* (θ^M), with

$$\theta^L = (0, 0, 1/2), \qquad \theta^R = (1, 0, 1/2), \qquad \theta^P = (0, 1, \alpha^P), \qquad \theta^M = (0, 1, \alpha^M).$$

Left and right types mirror the preferences of the respective parties. Traditional and new issues are of equal importance to them. The supporters of the entrant party share its peak, but differ in their assessment of the importance of the new policy dimension b. Some of them are more passionate for the new issue, while others have a stronger attachment to the traditional issue and are hence more moderate in their support for the entrant. Formally, we have $\alpha^P < \alpha^M$. Figure 3.2 depicts the preferences of the four types.

The distribution of voter types is unknown. In contrast to previous literature, in which voters' types are determined by *iid* draws, the whole distribution of voters' types is drawn at once from a joint distribution. This creates aggregate uncertainty, even for large electorates. Let N_j be the random number of voters of type θ^j. We denote a distribution of voters' types by $N = (N_R, N_L, N_M, N_P)$ with $N_R + N_L + N_M + N_P = n$. N is randomly drawn according to a common prior P. Through the following assumption concerning the support of P, we restrict attention to the specific scenarios addressed in the introduction.

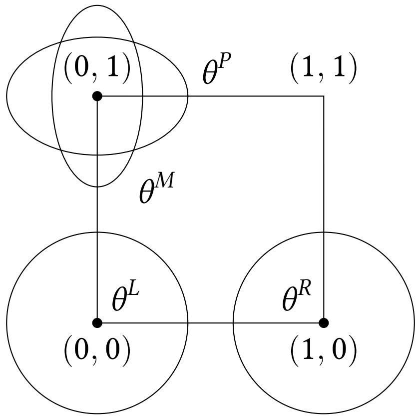

Figure 3.2: Black dots depict the peaks of the four types. Circles and ellipses indicate typical indifference curves. Notice that passionates and moderates share the same peak, but differ with respect to the relative importance of the two issues.

Assumption 3.1. *The prior P is such that there are constants $n_R \in (\tau, n/2)$ and $\underline{n_M} \geq 3$, and*

$$\Pr(N_R = n_R) = 1, \qquad \Pr(N_L = y) \begin{cases} > 0 & \text{if } \tau < y < n/2 - 1 \\ = 0 & \text{otherwise} \end{cases}$$

$$\Pr(N_M = y) \begin{cases} > 0 & \text{if } \underline{n_M} \leq y < \tau \\ = 0 & \text{otherwise} \end{cases}, \qquad \Pr(N_P = y) \begin{cases} > 0 & \text{if } y < \tau - 2 \\ = 0 & \text{otherwise} \end{cases}$$

Assumption 3.1 is illustrated in Figure 3.3. It guarantees that the number of voters on the right wing, $N_R = n_r$, and on the left wing, $N_L + N_M + N_P = n - n_R$, are known, e.g., due to previous (non-modeled) elections. The distribution within the left wing is uncertain, though. None of the two groups of entrant supporters

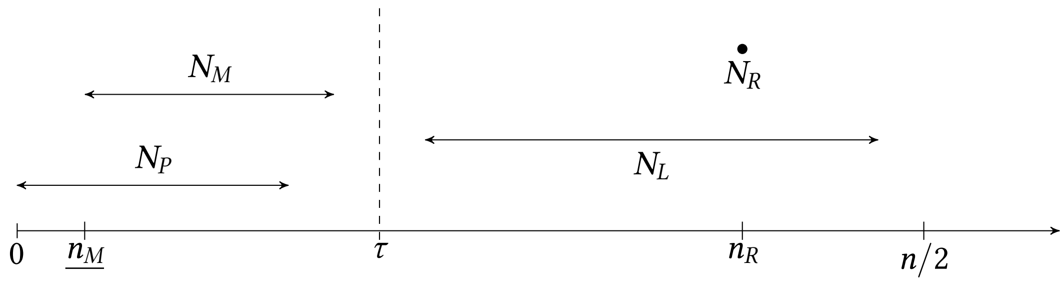

Figure 3.3: The figure illustrates Assumption 3.1. The number of right types, N_R, is know to equal n_R, strictly between the election threshold τ and an outright majority. The numbers of passionates, N_P, moderates, N_M, and left types, N_L, are random variables. Their support is known to be in the depicted intervals.

is able to meet the election threshold on its own. Jointly they have a chance to do so. The specification allows for a correlation of N_P and N_M, but we do not require it. The number of moderate types has a commonly known lower limit of $\underline{n_M}$. Finally, both L and R have more than τ supporters but less than the absolute majority of voters. Notice that Assumption 3.1 restricts the support, but not the probabilities.

3.4.3 The Election Game

There are two consecutive elections. Each consists of a ballot, in which all voters cast their vote, and a subsequent bargaining stage, in which parties determine policy. The exact timing is a follows:

- $t = 0$ N is drawn according to prior P, voters learn their own type θ^i.
- $t = 1$ Voting takes place.
  - Parties above the election threshold enter parliament.
  - Coalition bargaining determines the governing coalition.
  - The governing coalition implements policy, payoffs are realized.
- $t = 1.5$ Voters observe the result of the first election and update their beliefs.
- $t = 2$ Repetition of $t = 1$.

Notice that nature only turns once, at the beginning of the game; the preference distribution remains constant during the whole game but is not public information. Therefore, voters will want to update their beliefs using the first election result in order to make better-informed decisions in period two.

We detail on the coalitional bargaining stage in Section 3.5. For now, suppose that policy in period t is given by some random variable X^t. At each election, voters (correctly) anticipate the bargaining outcome. We consider the following *election game*: Let v_J^t be the number of votes for party J in period t, and let $V^t := \{(v_R^t, v_L^t, v_E^t) \in \mathbb{N}_0^3 | \sum_J v_J^t = n\}$ be the set of all possible results in period t. Recall that Θ is the type set. Behavior of voter i is given by a pure strategy $\sigma_i = (\sigma_i^1, \sigma_i^2)$, which specifies voting decisions $\sigma_i^1 \colon \Theta \to \{R, L, E\}$ for period one, and $\sigma_i^2 \colon (\Theta, V^1) \to \{R, L, E\}$ for period two. Each strategy profile $\sigma := \{\sigma_1, \ldots, \sigma_i, \ldots, \sigma_n\}$ induces distributions of votes, $v^1(\sigma)$ and $v^2(\sigma)$, which are random due to the uncertainty about N.[17] Parties' vote shares determine their

[17] Notice that our strategy definition is restrictive. In general, the second-period voting decision could depend not only on voters' type and the previous election outcome but also on the parties' action and first-period policy. Our main results, however, remain valid if we allow for these general strategies.

seat shares, accounting for those parties failing the election threshold. Policy is endogenously derived from parties' seat shares, as detailed in Section 3.5. Thus there is a mapping $X^t(\sigma)$ from strategy profiles into the policy space. The mapping is random due to the randomness in types and the nature of the bargaining process.

Anticipating policy choice $X^t(\sigma)$, voter i chooses strategy σ_i in order to maximize expected utility

$$U(\sigma_i, \sigma_{-i}; \theta^i) = \mathbb{E}[u(X^1(\sigma_i, \sigma_{-i}); \theta^i)|\theta^i] + \delta\mathbb{E}[u(X^2(\sigma_i, \sigma_{-i}); \theta^i)|\theta^i],$$

given the strategy profile σ_{-i} of all other voters. We look for a Perfect Bayesian Equilibrium of the election game.[18]

3.5 Coalition Bargaining

After each election, the parties represented in parliament determine policy. If a party has a seat share of more than 50%, this party alone decides about policy and receives all spoils of office. If no party has such a majority, a coalition is necessary to form government. Then one party is randomly selected to be the *proposer*. The selection probability is equal to the seat share. The proposer makes a take-it-or-leave-it offer to a subset of parties that (including the proposer) holds at least 50% of seats in the parliament. The offer specifies a policy and a distribution of the spoils of office s_J. All members of the proposed coalition have to accept the proposal; if at least one of them rejects, a default policy is implemented and no party gains any spoils of office.[19]

The default policy is linked to the (non-modeled) past. As we model an old struggle over the traditional issue only, a natural candidate for the default policy is the one who emerges as a compromise between left and right. We thus assume

[18] The solution concept allows only for uncertainty about the distribution of policy preferences in the population. There is no strategic uncertainty, because voters know which strategy is played by other voters. It is embedded in the equilibrium concept that every voter correctly anticipates others' strategies. There is no coordination failure in the sense that some voters believe equilibrium A is played, while others play equilibrium B. By using Perfect Bayesian Equilibrium, we highlight the effects that arise from the unknown type distribution N.

[19] This modeling of a coalitional bargaining process within a general election model was proposed by Baron and Diermeier (2001). We make the simplification that coalitional bargaining only takes place if necessary, i.e., if no party has more than 50% of seats. In terms of empirical evidence, the Baron-Diermeier model is supported by Diermeier and Merlo (2004), who explore data of coalition formation in 11 European countries spanning the years 1945–1997. They find that a selection probability corresponding to seat shares fits the data quite well and outperforms selection according to the seat share ranking.

that the default policy x_d is equal to $(1/2, 0)$. The results do not depend on this specific value, but are robust to some perturbations, as shown in Lemma 3.2 below.

Due to the quasi-linearity of parties' preferences in spoils of office, the policy implemented by some coalition depends only on the preferences of its members, not on the default policy nor the relative seat shares within the coalition.[20] The default policy does determine how the spoils of office are distributed between coalition members. Those are not relevant for voters, though.

What is important to voters is that for each possible coalition there is a unique induced policy. Therefore, the ballot affects the probability distribution over the set of possible coalitions. In fact, no more than four policies are possible as outcomes of the bargaining game when we restrict attention to elections in which the Entrant misses an outright majority.

Lemma 3.1. *Let the seat share of party E be below 50%. Then in any equilibrium and in any period the coalition resulting from coalitional bargaining is $\{L\}$, $\{L, E\}$, $\{R\}$ or $\{R, E\}$ and the respective policy is x_L, x_{LE}, x_R or x_{RE} with*

$$x_L = (0, 0), \qquad x_{LE} = \big(0, 1/2\big)\,, \qquad x_R = (1, 0), \qquad x_{RE} = \big(1/2, 1/2\big)\,.$$

Furthermore, if no party has an outright majority of more than 50%, it is optimal for R to propose coalition $\{R, E\}$ and for L and E to propose $\{L, E\}$.

With respect to conditional probabilities of equilibrium policies, Lemma 3.1 implies the following. If E fails to enter, x_L occurs with probability

$$\Pr[v_L^t > v_R^t] + \Pr[v_L^t = v_R^t]/2$$

while x_R occurs with complementary probability. Neither x_{LE} nor x_{RE} occurs. If all three parties enter parliament and each receives less than 50% of votes, then x_{LE} occurs with probability $[v_L^t(\sigma) + v_E^t(\sigma)]/n$ and x_{RE} occurs with complementary probability $v_R^t(\sigma)/n$.

Figure 3.4 illustrates the location of possible equilibrium policies. These do not depend on the seat shares. A coalition of, say, Right and Entrant will always implement x_{RE}, independently of their relative size. The seat share does determine the likelihood of the respective policies, though. The default policy is relevant for the proposer's decision which coalition to opt for. The Right Party, for instance,

[20] This is true also for more complex bargaining structures where parties can make counter-proposals. In this case, reservation utilities of parties depend on their endogenous continuation values rather than default policies. Party preferences over spoils and policies ensure that the policy implemented by coalition C does not depend on the particular bargaining process. For further discussion, see Baron and Diermeier (2001).

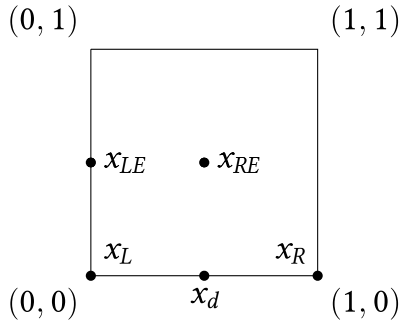

Figure 3.4: The coalitional bargaining process has a limited number of policy outcomes. There will either be a single-party government of Left or Right, implementing x_L or x_R, or a two-party coalition including the Entrant, implementing x_{LE} or x_{RE}.

proposes to the Entrant because the default would be very bad for the Entrant and thus the Entrant is willing to accept a high side payment to Right in order to prevent the status quo. Right can thus secure a high overall payoff. Its optimal proposal could change if the default policy was located elsewhere. But, according to the following result, Lemma 3.1 is in fact robust to some perturbation in x_d.[21]

Lemma 3.2. *Lemma 3.1 is valid for all default policies $x_d = (a_d, b_d)$ with*

$$x_d \in \{(a_d, b_d) \in [0,1]^2 \mid b_d \leq 1/4 \leq a_d \text{ and } a_d^2 + b_d^2 \leq 1/3\}.$$

Having analyzed party behavior after an election, we now turn to the voting decision and to the question why a voter should support a small party that is expected to fail at the election threshold.

3.6 The Investing Equilibrium

Put yourself in the situation of a voter in the booth. Left and right types expect their preferred parties to enter parliament and can vote sincerely. Supporters of the small party have a harder time. They like E to enter parliament which requires voting for E. On the other hand, if the Entrant fails the election threshold, a vote for the Entrant is wasteful, and voters would be better off voting for the left party. If the probability of E failing the threshold is sufficiently high, voting for them is thus never optimal in a one-shot game. The existence of a second period changes the game, though. Suppose it is certain that E will make it into parliament in

[21] On the other hand, if the default policy is outside the set given in Lemma 3.2, the parties propose different coalitions: If $b_d > 1/4$, then R proposes a coalition with L, as the default policy is very bad for L in this case. If $a_d < 1/4$, then E makes R an offer, as the default policy puts R at disadvantage. Finally, if $a_d^2 + b_d^2 > 1/3$, R proposes a grand coalition of all three parties.

$t = 2$ if passionate and moderates coordinate on voting for them. Then voting E is optimal in $t = 2$ and all supporters of the entrant are quite happy with the policy outcome. Call this the *entrance scenario.*

Now suppose voting in period one affects the probability of the entrance scenario. This additional incentive could induce supporters of the entrant to vote for the entrant even though this is wasteful in the short run. We show that this is indeed the case: passionate types vote for the entrant in the first period, actually knowing for sure that E will miss the threshold. By using their vote as a signal, they change the beliefs about the type distribution and foster the entrance scenario in the second period. Moderate types choose their second-best alternative in the first period and only jump the entrant bandwagon if it promises success in period two. To simplify the exposition of the investing equilibrium, we impose the following assumption on the prior P.

Assumption 3.2. *Voters of type θ^R are rather pivotal in the decision between R and L than in the decision whether E hits the equilibrium threshold. In particular,*

$$2\Pr\left[N_L + N_M \in \{n_R, n_R - 1\} | \theta^R\right] > \delta \Pr\left[N_P = \tau - \underline{n_M} | \theta^R\right] \left(3\frac{n_R}{n} - 1\right).$$

Assumption 3.2 simplifies exposition. But there is an outcome-equivalent equilibrium which works without Assumption 3.2 (see Proposition 3.2). Also, for $n_R/n \leq 1/3$, the assumption is trivially satisfied.

Proposition 3.1. *For every prior P consistent with Assumptions 3.1 and 3.2, the following strategies constitute a Perfect Bayesian equilibrium if α^P is sufficiently small and α^M is sufficiently large, as specified by conditions* (3.1) *and* (3.2) *in the proof.*

- *Left and right types vote for their respective party in both periods.*
- *Passionate types vote for E in the first period, while moderate types vote for L. In the second period, they both vote for E if the first-period votes of E, v_E^1, meet a threshold e, defined as*

  $$e = \tau + 1 - \underline{n_M}$$

  Otherwise all of them vote for L.

On the equilibrium path, the entrant party gets a positive vote share that is below the election threshold in the first election. It is common knowledge that it will not enter parliament. Nevertheless, passionate voters have an incentive to vote for the entrant in order to increase the probability of entry in the next election. In the second period, the new party either receives no votes at all or gets sufficient votes to enter parliament. While the entrant attracts only passionate supporters in the first period, once it has established itself its electorate becomes more moderate.

Corollary 3.1. *In the investing equilibrium, the ex-ante probability that the new party E enters parliament is* $\Pr(N_P \geq \tau + 1 - \underline{n_M})$*, the probability of sufficient passionate supporters.*

After the first election, voters believe that the number of passionate voters equals v_E^1, the votes for the entrant party. Thus, the entrant has sufficient support to meet the threshold τ if $v_E^1 \geq \tau - \underline{n_M}$. Yet, the equilibrium threshold e departs from that value as $+1$ is added to the threshold. This guarantees entry (conditional on meeting e) even if a non-passionate voter pretends to be a passionate type by voting E in $t = 1$. As a consequence, the in-equilibrium incentives of passionate and moderate voters have an identical structure. For both types, voting L instead of E yields a short-term benefit of

$$\Pr[pivotal|\theta^i] \cdot \left(u(x_L;\theta^i) - u(x_R;\theta^i)\right)$$

as voting L increases the chances of L rather than R building the first-period government. In the long run, however, voting E rather than L creates an additional value[22] of

$$\delta\Pr[pivotal|\theta^i] \cdot \left(\frac{n_R u(x_{RE};\theta^i) + (n - n_R) u(x_{LE};\theta^i)}{n} - u(x_L;\theta^i)\right),$$

as voting for E in period 1 increases the entrant party's probability to enter parliament in period 2. While the probabilities and the gains from changing the election outcome differ for the two types of entrant supporters, the basic trade-off between short- and long-run effects is the same. This analogy requires adding $+1$ to the endogenous threshold. Otherwise, there is a chance that a moderate voter voting for E triggers coordination for E in a state in which entrant supporters fall short of the threshold τ. This results in R having an outright majority and implementing x_R, the worst outcome for entrant supporters. The $+1$ excludes this case for individual deviations.

We call a voter 'passionate' if the long-run incentives outweigh the short-run benefits, and 'moderate' otherwise. While both effects are small, as they depend on the probability of being pivotal, they are both on the same scale.[23]

[22] To be precise, voting for E has a long-run benefit for entrant supporter i only if $\alpha^i < 3/(n_R/n + 3)$. Otherwise, this voter cares so much about the traditional dimension that she prefers the certain alternative x_L over the lottery with x_{LE} and x_{RE}.

[23] We thank Daniel Krämer for bringing that point to our attention.

Robustness
It is possible to describe the investing equilibrium without Assumption 3.2.

Proposition 3.2. *For every prior P consistent with Assumption 3.1, the following strategies constitute a Perfect Bayesian equilibrium if* α^P *is sufficiently small and* α^M *is sufficiently large, as specified by conditions* (3.1) *and* (3.2) *in the proof of Proposition 3.1.*

- *Left and right types vote for their respective party in both periods.*
- *Passionate types vote for E in the first period, while moderate types vote for L. In the second period, they both vote for E if the first period votes of E,* v_E^1*, meet a threshold e, defined as*

$$e = \begin{cases} \tau + 1 - \underline{n_M} + \max\{0, n_R - v_R^1\}, & \text{if } n_R/n > 1/3 \\ \tau + 1 - \underline{n_M}, & \text{if } n_R/n \leq 1/3. \end{cases}$$

  Otherwise all of them vote for L.

For $n_R/n > 1/3$, right types favor entry of the entrant party over an absolute majority of the left party. Thus, they might want to deviate from the equilibrium path by voting for *E*. To guarantee that such a deviation is not profitable, the endogenous threshold increases if less than n_R votes for *R* are observed in $t = 1$. On the other hand, if $n_R/n \leq 1/3$, then right types prefer an absolute left majority over the entrance of *E*. If they could raise the threshold by voting *L* (instead of *R*), they might want to do so to decrease chances of entrance of *E* in $t = 2$. To eliminate this incentive, the threshold is independent of v_R^1 in the case of $n_R/n \leq 1/3$.

So far, we concentrated on one particular equilibrium. The investing equilibrium serves our goal of explaining voting patterns as described in our motivation. In the following, we discuss another natural equilibrium candidate.

3.7 Sincere Voting

Voting is *sincere* if a voter chooses to vote for the party that matches the voter's preferences most closely. With two alternatives, sincere voting typically is a weakly dominant strategy, which makes it a prominent candidate for predicting behavior. Indeed, many political models take sincere voting for granted. With more than two alternatives, however, there is typically no dominant strategy, not even in the weak sense. This is true for our model as well. The question whether sincere voting is strategically optimal is far from trivial. We present a negative result in the following proposition.

Proposition 3.3. *For any prior consistent with Assumptions 3.1, sincere voting is not an equilibrium of the game.*

The intuition is that voters learn the distribution of types after the first election. Therefore, they can increase their utility by making their voting decision dependent on this information. In particular, a moderate voter can deviate to L instead of E in the second period, if she knows that E cannot enter parliament as the number of moderate and passionate types in the population is too low.

This points to a second equilibrium candidate: In the first period, everyone votes sincerely. But if voters learn that E cannot enter parliament, moderate and passionate types vote for L instead of E in the second period. Otherwise, there is sincere voting in the second period also. We call this strategy *conditionally sincere voting*. Notice that this strategy – at least in the second period – is similar to the strategies used in the investing equilibrium. Yet, we show that conditionally sincere voting is not always optimal, even under a condition in which the investing equilibrium exists. To see why, consider a moderate voter switching to L in the first period. Define P^L as the change in probabilities of L forming the government caused by such behavior.[24]

$$\begin{aligned} P^L := &\frac{1}{2}\Pr\left[N_P + N_M < \tau \wedge N_L \in \{n_R - 1, n_R\}\right] + \Pr\left[N_P + N_M = \tau \wedge N_L \geq n_R\right] + \\ &+ \frac{1}{2}\Pr\left[N_P + N_M = \tau \wedge N_L = n_R - 1\right]. \end{aligned}$$

There are three terms. First, the new party does not pass the threshold anyhow and does not enter parliament. Then the voter could be pivotal in the decision between L and R. The second and third terms capture the case of the new party missing the threshold exactly for the one voter.

Assumption 3.3. *A moderate type is rather pivotal in the decision between R and L than in the decision whether E enters parliament. In particular,*

$$4P^L > \left(4 - \frac{n_R}{n}(1+\delta)\right)\Pr\left[N_P + N_M = \tau \middle| \, \theta^M\right].$$

Given this assumption, a voter of type θ^M has a profitable deviation if she is sufficiently moderate.

[24] All probabilities are conditional on $\theta^i = \theta^M$. The dependence is suppressed for notational convenience.

Proposition 3.4. *For any prior P consistent with Assumptions 3.1 and 3.3, conditionally sincere voting is not an equilibrium of the game if* α^M *is sufficiently large, as specified by equation* (3.3) *in the proof.*

In the corresponding one-period model, sincere voting is no equilibrium either.

Corollary 3.2. *Consider a one-shot version of the model with only one election period. Sincere voting is not an equilibrium of the game for any prior consistent with Assumptions 3.1 and 3.3 if* α^M *is sufficiently large, as specified by equation* (3.3) *with* $\delta = 0$.

Finally, we conclude that the result of Proposition 3.4 is consistent with previous results.

Lemma 3.3. *For every prior consistent with Assumptions 3.1 and 3.3, there are* α^P *and* α^M *such that condition* (3.3) *in the proof of Proposition 3.4 and conditions* (3.1) *and* (3.2) *in the proof of Proposition 3.1 are satisfied.*

Therefore, for some parameter values the investing equilibrium exists, while neither sincere voting nor conditionally sincere voting are equilibria. Then the investing equilibrium can explain the observed voting pattern while sincere voting cannot.

3.8 Conclusion

In this chapter, we set up a model in which voters coordinate their voting intentions by using their votes as a signal. This allows others to learn about the distribution of preferences in the population. If passionate supporters of a new party are convinced about its importance, they vote for this new party, although they lose their influence on the composition of parliament in the current election. However, they might influence the results of future elections by pushing more moderate supporters to vote for the new party in future elections.

By extending the time horizon of voters' perceptions, we offer a strategic explanation for a voting behavior that looks wasteful at first sight. The explanation is also valid in situations when sincere voting is not strategically optimal.

Our results suggest some paths for future research. While we concentrate on election thresholds in a proportional system, the basic intuition carries over to majoritarian systems. There are differences, though, as the effective election threshold in a majoritarian system is an endogenous object equal to the minimum number of votes to secure a majority. Including the possibility to abstain from voting might prove interesting as well. In such an extension, the emerging new party might get initial support from former non-voters. The idea would be that non-voters did not vote because they are more or less indifferent between the incumbent parties. Voting for the new party is thus less costly for them even if the party fails the threshold. In this respect, such voters mirror the preferences of the passionates types in our model. Another field in which the basic result of our model might be applied are the dynamics of civil protests. Starting a civil protest can involve high personal cost, but yields the prospect of others joining in once a movement has gained a critical mass. Meirowitz and Tucker (2013) and Kricheli, Livne, and Magaloni (2011) analyze these kinds of scenarios, albeit with quite a different modeling approach.

Appendix 3.A Proofs

Proof of Lemma 3.1. If either L or R has strictly more than 50% of seats, this party implements its peak. If neither has an outright majority, then a coalition of any two parties holds at least 50% of seats. Suppose party J is the proposer and proposes coalition C. If it makes an offer that is not accepted, x_d is implemented and no spoils are distributed. Such an outcome is dominated by proposing x_d alongside an equal share of S to all parties. Thus, we are looking for the best offer that is accepted. Obviously, party J always proposes $s_I = 0$ to any party I outside the coalition C. Then the optimal proposal of party J for coalition C solves

$$\max_{a,b,s_J,s_{-J}} u^J(a,b) + s_J$$

$$\text{subject to } u^K(a,b) + s_K \geq u^K(x_d) \ \forall K \in C \ \text{ and } \sum_{K \in C} s_K \leq S.$$

Note that the solution of that problem maximizes $\sum_{K \in C} u^K$. We calculate the optimal proposal for the three possible proposers.

1. L as proposer:

In a coalition with R, policy $(1/2, 0)$ will be implemented. As this is equal to x_d, L can grab all spoils of office and gets utility $u^L(x_d) + S = -1/4 + S$.
In a coalition with E, policy $x_{LE} = (0, 1/2)$ will be implemented. E likes x_{LE} much better than x_d and is actually willing to "pay" L for offering a coalition. In fact, E is willing to accept $s_E = u^E(x_d) - u^E(x_{LE}) = -1$. Hence L gets utility $u^L(x_{LE}) + S + 1 = 3/4 + S$ which is bigger than its utility from a coalition with R.
If L proposes a *grand coalition* containing all three parties, the value-maximizing policy is $x_{LRE} = (1/3, 1/3)$. Furthermore, the necessary payments to coalition members are $s_R = u^R(x_d) - u^R(x_{LRE}) = 11/36$ and $s_E = u^E(x_d) - u^E(x_{LRE}) = -25/36$. Thus, L receives utility $u^L(x_{LRE}) + S - 11/36 + 25/36 = S + 1/6$ which is smaller than its utility from a coalition with E only. If L has exactly 50% of seats, it could propose a single-party government and receive $0 + S$ if L is chosen as proposer. L prefers, however, to propose a coalition with E and policy $x_{LE} = (0, 1/2)$.

2. R proposes a coalition with E and a policy offer $x_{RE} = (1/2, 1/2)$, as

$$U^R = \begin{cases} u^R(x_{RE}) + u^E(x_{RE}) - u^E(x_d) + S = 1/4 + S & \text{in a coalition with } E \\ u^R(x_{RL}) + u^L(x_{RL}) - u^L(x_d) + S = -1/4 + S & \text{in a coalition with } L \\ u^R(x_{RLE}) + u^L(x_{RLE}) - u^L(x_d) & \\ \qquad + u^E(x_{RLE}) - u^E(x_d) + S = 1/6 + S & \text{in a grand coalition} \\ u^R(x_R) = 0 + S & \text{in a single-party gov't.} \end{cases}$$

3. E proposes a coalition with L and a policy offer $x_{LE} = (0, 1/2)$, as

$$U^E = \begin{cases} u^E(x_{LE}) + u^L(x_{LE}) - u^L(x_d) + S = -1/4 + S & \text{in a coalition with } L \\ u^E(x_{RE}) + u^R(x_{RE}) - u^R(x_d) + S = -3/4 + S & \text{in a coalition with } R \\ u^E(x_{RLE}) + u^L(x_{RLE}) - u^L(x_d) & \\ \qquad +u^R(x_{RLE}) - u^R(x_d) + S = -5/6 + S & \text{in a grand coalition.} \end{cases}$$

□

Proof of Lemma 3.2. We only have to consider the case when no party has an absolute majority. Given a coalition C, the implemented policy just depends on the members of the coalition C and remains unchanged from Lemma 3.1. Differing the default policy only changes the distribution of the spoils of office. Denote the default policy x_d by (a, b). Again, we go through the cases of the three possible proposers.

1. L proposes LE if $b \leq a$ and $a^2 + b^2 \leq -1/6 + 2a$, as

$$U^L = \begin{cases} -1/2 + S + a^2 + (1-b)^2 & \text{in a coalition with } E \\ -1/2 + S + (1-a)^2 + b^2 & \text{in a coalition with } R \\ -4/3 + S + a^2 + (1-a)^2 + b^2 + (1-b)^2 & \text{in a grand coalition.} \end{cases}$$

2. R proposes RE if $b \leq 1/4$ and $a^2 + b^2 \leq 1/3$, as

$$U^R = \begin{cases} -1 + S + a^2 + (1-b)^2 & \text{in a coalition with } E \\ -1/2 + S + a^2 + b^2 & \text{in a coalition with } L \\ -4/3 + S + 2a^2 + b^2 + (1-b)^2 & \text{in a grand coalition.} \end{cases}$$

3. E proposes LE if $a \geq 1/4$ and $a^2 + b^2 \leq -1/6 + 2a$, as

$$U^E = \begin{cases} -1 + S + (1-a)^2 + b^2 & \text{in a coalition with } R \\ -1/2 + S + a^2 + b^2 & \text{in a coalition with } L \\ -4/3 + S + a^2 + (1-a)^2 + 2b^2 & \text{in a grand coalition.} \end{cases}$$

In summary, all the conditions are satisfied if $b \leq 1/4 \leq a$ and $a^2 + b^2 \leq 1/3$. □

Proof of Proposition 3.1. Using backward induction, we begin with the second period. First, let party E's vote share in period one be small, i. e., $v_E^1 < e$. According to the equilibrium strategies, $v^2 = (v_R^2, v_L^2, v_E^2) = (n_r, n - n_r, 0)$. Then, given Assumption 3.1, L has an outright majority, forms a single-party government, and implements $x_L = (0,0)$. An individual deviation will make a difference only if L looses its outright majority. If it does, it will result in either a tie between R and L or an outright majority for R. Both alternatives are worse for left, moderate, and passionate types. So these voters cannot profit from a deviation. A right type would like L to loose its majority, but all right types vote R already, so a deviation would actually reduces R's prospects. Thus right types have no profitable deviation either.

Second, consider $v_E^1 \geq e$. On the equilibrium path, $v^2 = (n_r, n - n_r - N_M - N_P, N_M + N_P)$. Coalition bargaining yields equilibrium outcomes as described in Lemma 3.1. Hence, with probability $v_R^2/n = n_R/n =: r$ party R is chosen to form a coalition and the policy is x_{RE}, whereas with probability $(1-r)$ the proposer is either L or E and in both cases the policy is x_{LE}. Note that, by Assumption 3.1 and the definition of e, no party gains an absolute majority. Neither will L or E gain such a majority if they receive one additional vote. Again, deviation is not profitable for any type: (a) By deviating (to E or L) a right type voter would reduce the chances of a $\{R,E\}$-coalition in favor of $\{L,E\}$, thus reducing his payoff. (b) A left type voter switching to R reduces the chances of a $\{L,E\}$-coalition, and both alternatives, $\{R,E\}$ or even single party government $\{R\}$, are less-preferred alternatives. Switching to E has no effect at all. (c) Finally, moderate and passionate voters can not improve for reasons analogous to (b); switching to R reduces the expected payoff, switching to L does not change it.

In summary, there is no profitable deviation in period two. Before proceeding to period one, we state the period-two payoffs along the different equilibrium paths: When the Entrant fails to enter parliament, the period-two payoffs are

$$\mathbb{E}(u(X^2;\theta)|v_E^1 < e;\theta) = u(0,0;\theta) = \begin{cases} -\frac{1}{2} & \text{for } \theta = \theta^R \\ 0 & \text{for } \theta = \theta^L \\ \alpha^P - 1 & \text{for } \theta = \theta^P \\ \alpha^M - 1 & \text{for } \theta = \theta^M . \end{cases}$$

When the Entrant joins parliament, the payoffs are

$$\mathbb{E}(u(X^2;\theta)|v_E^1 \geq e;\theta) = \begin{cases} -r\frac{1}{4} - (1-r)\frac{5}{8} = \frac{1}{8}(3r-5) & \text{for } \theta = \theta^R \\ -r\frac{1}{4} - (1-r)\frac{1}{8} = -\frac{1}{8}(1+r) & \text{for } \theta = \theta^L \\ -\frac{1}{4}(1-\alpha^P(1-r)) & \text{for } \theta = \theta^P \\ -\frac{1}{4}(1-\alpha^M(1-r)) & \text{for } \theta = \theta^M . \end{cases}$$

Period 1: Given equilibrium strategies and Assumption 3.1, only L and R enter parliament and either x_L or x_R is implemented depending on the vote shares. E does not enter parliament, even if an individual voter switches her vote to E.

Voters may have incentives to deviate in order to change the probabilities of policies x_L and x_R (*short-term incentive*) or in order to change voting behavior in period 2 (*long-term incentive*). In the following, we rule out any one-stage deviation.

Incentives of a left type: Switching to R is unprofitable; it has a negative short-term effect and no long-term effect. Switching to E has a negative short-term effect as well. Furthermore, it increases the chance of meeting threshold e. Left types, however, prefer the very opposite with respect to period 2.

Incentives of a right type: A vote for L has a negative short-term effect (compared to voting for R) and there is no long-run effect. Voting for E instead of R can be pivotal in two respects: In the short run, it might lead to a left government rather than a right one. In the long run, it could induce a coalition government rather then a left one. The total change in utility is

$$\begin{aligned}
&dU = \Pr\left[N_M + N_L \in \{n_R - 1, n_R\} | \theta = \theta^R\right] \frac{u(x_L;\theta^R) - u(x_R;\theta^R)}{2} \\
&+ \delta \Pr\left[N_P + 1 = e | \theta = \theta^R\right] \left(\mathbb{E}(u(X^2;\theta^R)|v_E^1 \geq e;\theta^R) - \mathbb{E}(u(X^2;\theta^R)|v_E^1 < e;\theta^R)\right) \\
&= -\frac{1}{4}\Pr\left[N_M + N_L \in \{n_R - 1, n_R\} | \theta = \theta^R\right] + \delta \Pr\left[N_P + 1 = e | \theta = \theta^R\right] \frac{3r-1}{8}\,.
\end{aligned}$$

The deviation is not profitable if $dU \leq 0$, i.e., if

$$2\Pr\left[N_M + N_L \in \{n_R - 1, n_R\} | \theta = \theta^R\right] \geq \delta \Pr\left[N_P = \tau - \underline{n_M} | \theta = \theta^R\right](3r-1)\,.$$

Assumption 3.2 ensures this condition.

Incentives of a moderate type: Switching the vote from L to R is clearly not profitable. Switching to E lowers short-term expected utility, but may help E to enter parliament later, if it triggers coordinated voting for E in period 2. This is unprofitable if the short-term losses

$$\begin{aligned}
&\left(\Pr\left[N_L + N_M = n_R | \theta^M\right] + \Pr\left[N_L + N_M = n_R + 1 | \theta^M\right]\right) \frac{u(x_R;\theta^M) - u(x_L;\theta^M)}{2} \\
&\qquad = -\left(\Pr\left[N_L + N_M = n_R | \theta^M\right] + \Pr\left[N_L + N_M = n_R + 1 | \theta^M\right]\right) \frac{\alpha^M}{2}
\end{aligned}$$

are higher than the long-term gains from switching

$$\delta \Pr\left[N_P = \tau - \underline{n_M} | \theta^M\right] \left(\mathbb{E}(u(X^2; \theta^M) | v_E^1 \geq \tau + 1 - \underline{n_M}; \theta^M) - u(x_L; \theta^M)\right)$$
$$= \delta \Pr\left[N_P = \tau - \underline{n_M} | \theta^M\right] \frac{1}{4} \left(3 - \alpha^M(r+3)\right).$$

This requires

$$\alpha^M \geq \frac{3\delta \Pr\left[N_P = \tau - \underline{n_M} | \theta^M\right]}{2\Pr\left[F_1\right] + \delta \Pr\left[N_P = \tau - \underline{n_M} | \theta^M\right](3+r)} \tag{3.1}$$

with the event F_1 defined by $N_L + N_M \in \{n_R, n_R + 1\}$.

Incentives of a passionate type: Passionate types have a short term incentive to switch their vote to L to increase the probability of L getting the majority. On the other hand, switching to L reduces the chances that $v_E^1 \geq \tau + 1 - \underline{n_M}$, thus decreasing the probability of E entering parliament in period 2. The short-term benefits from switching are:

$$\left(\Pr\left[N_L + N_M = n_R | \theta^P\right] + \Pr\left[N_L + N_M + 1 = n_R | \theta^P\right]\right) \frac{u(x_L; \theta^P) - u(x_R; \theta^P)}{2}$$
$$= \Pr\left[F_2 | \theta^P\right] \frac{\alpha^P}{2}$$

with the event F_2 defined by $N_L + N_M \in \{n_R - 1, n_R\}$. The long-term losses from switching are:

$$-\delta \Pr\left[N_P = \tau + 1 - \underline{n_M} | \theta^P\right] \left(\mathbb{E}(u(X^2; \theta^P) | v_E^1 \geq \tau + 1 - \underline{n_M}; \theta^P) - u(x_L; \theta^P)\right)$$
$$= -\delta \Pr\left[N_P = \tau + 1 - \underline{n_M} | \theta^P\right] \frac{1}{4} \left(3 - \alpha^P(r+3)\right)$$

We find that the deviation is unprofitable if $\Pr[F_2 | \theta^P] = 0$ or

$$\alpha^P \leq \frac{3\delta \Pr\left[N_P = \tau + 1 - \underline{n_M} | \theta^P\right]}{2\Pr\left[F_2 | \theta^P\right] + \delta \Pr\left[N_P = \tau + 1 - \underline{n_M} | \theta^P\right](r+3)}. \tag{3.2}$$

Switching the vote from E to R is unprofitable; it has a negative short-term effect, and the long-term effect is negative if $\alpha^P \leq 3/(3+r)$, which is implied by (3.2).

Note that conditions (3.1) and (3.2) are compatible with $0 \leq \alpha^P < \alpha^M < 1$. □

Proof of Corollary 3.1. Corollary 3.1 follows immediately from the statement of Proposition 3.1. □

Proof of Proposition 3.2. The proof proceeds analogously to the proof of Proposition 3.1. Only the incentives of a right type change.

A vote for L has a negative short-term effect (compared to voting for R). It suffices to show that the long-run effect is non-positive:

$r = n_R/n > 1/3$: Voting for L increases the threshold e and reduces the probability of E entering parliament. Payoff $\mathbb{E}(u(X^2;\theta^R)|v_E^1 < e;\theta^R) = -1/2$ becomes more likely, whereas payoff $\mathbb{E}(u(X^2;\theta^R)|v_E^1 \geq e;\theta^R) = (3r-5)/8$ becomes less likely. Yet the R type voter does not profit from the deviation, as $-1/2 < (3r-5)/8 \Leftrightarrow r > 1/3$.

$r \leq 1/3$: In this case, switching has no long-run effect, as the threshold e is not affected.

Voting for E instead of R has a negative short-term effect. It suffices to show that the long-run effect is non-positive.

$r > 1/3$: Voting for E not only increases the number of votes E receives, but also the threshold e. Consequently, the probability of E meeting the threshold does not change. Hence switching has no long-term effect.

$r \leq 1/3$: In this case, switching does increase the probability that E meets the threshold. Payoff $\mathbb{E}(u(X^2;\theta^R)|v_E^1 < e;\theta^R) = -1/2$ becomes less likely, whereas payoff $\mathbb{E}(u(X^2;\theta^R)|v_E^1 \geq e;\theta^R) = (3r-5)/8$ becomes more likely. Yet the R type voter does not profit from the deviation as $-1/2 \geq (3r-5)/8 \Leftrightarrow r \leq 1/3$.

Hence, there is no profitable deviation for a right type. □

Proof of Proposition 3.3. Suppose everyone is voting sincerely in both periods. Now, consider the following deviation. A voter of type θ^M votes sincerely in the first period. If the entrant receives less than τ votes in the first period, she votes for L instead of E in the second period. This deviation could make L win the election instead of R increasing the voter's utility by α^M.

After the first election, voters learn the sum of passionate θ^P and moderate types θ^M. Assumption 3.1 guarantees that with positive probability the sum of passionate θ^P and moderate types θ^M is smaller than the election threshold τ. In addition, with positive probability the number of left types θ^L equals n_R or $n_R - 1$. Finally, both events together occur with positive probability. Hence, the deviation is profitable in expectation. Consequently, voting sincerely in both periods is not an equilibrium. □

Proof of Proposition 3.4. First, notice that the set of priors consistent with Assumptions 3.1 and 3.3 is non-empty.

Now, consider a voter of type θ^M in period one. Everyone else votes sincerely. By deviating from sincere voting, the voter may be pivotal in two respects. First, switching to L may help L to win a majority if E does not enter parliament. Second, switching may make E fail the threshold τ in period one, changing period-one government and implying a left government in period two instead of a three-party parliament. To simplify notation, let $u_K = u(x_K; \theta^M)$ for $K = L, R$ and let $u_E = ru(x_{RE}; \theta^M) + (1-r)u(x_{LE}; \theta^M)$ (with $x_{(\cdot)}$ as given in Lemma 3.1). Then, by deviating to L instead of E in the first period, the moderate's utility changes by

$$
\begin{aligned}
dU = {} & \Pr[N_M + N_P = \tau \wedge N_L + 1 < n_R | \theta^M] \left(u_R - u_E + \delta(u_L - u_E)\right) \\
& + \Pr[N_M + N_P = \tau \wedge N_L + 1 = n_R | \theta^M] \left((u_R + u_L)/2 - u_E + \delta(u_L - u_E)\right) \\
& + \Pr[N_M + N_P = \tau \wedge N_L + 1 > n_R | \theta^M] \left(u_L - u_E + \delta(u_L - u_E)\right) \\
& + \Pr[N_M + N_P < \tau \wedge N_L + 1 = n_R | \theta^M] \left((u_R + u_L)/2 - u_R\right) \\
& + \Pr[N_M + N_P < \tau \wedge N_L = n_R | \theta^M] \big(u_L - (u_R + u_L)/2\big) \\
= {} & \Pr[N_M + N_P = \tau \wedge N_L + 1 < n_R | \theta^M] \left(u_R - u_E + \delta(u_L - u_E)\right) \\
& + \Pr[N_M + N_P = \tau \wedge N_L + 1 = n_R | \theta^M] \left((u_R + u_L)/2 - u_E + \delta(u_L - u_E)\right) \\
& + \Pr[N_M + N_P = \tau \wedge N_L + 1 > n_R | \theta^M] \left(u_L - u_E + \delta(u_L - u_E)\right) \\
& + \Pr[N_M + N_P < \tau \wedge N_L \in \{n_r - 1, n_R\} | \theta^M] \big((u_L - u_R)/2\big)
\end{aligned}
$$

Further rearrangement yields

$$
\begin{aligned}
dU = {} & \Pr[N_M + N_P = \tau | \theta^M](\delta u_L - (1+\delta)u_E) \\
& + u_R \Bigg(\Pr[N_M + N_P = \tau \wedge N_L + 1 < n_R | \theta^M] + \frac{1}{2}\Pr[N_M + N_P = \tau \wedge N_L + 1 = n_R | \theta^M] \\
& \qquad - \frac{1}{2}\Pr[N_M + N_P < \tau \wedge N_L \in \{n_r - 1, n_R\} | \theta^M] \Bigg) \\
& + u_L \Bigg(\frac{1}{2}\Pr[N_M + N_P = \tau \wedge N_L + 1 = n_R | \theta^M] + \Pr[N_M + N_P = \tau \wedge N_L + 1 > n_R | \theta^M] \\
& \qquad + \frac{1}{2}\Pr[N_M + N_P < \tau \wedge N_L \in \{n_r - 1, n_R\} | \theta^M] \Bigg) \\
= {} & \Pr[N_M + N_P = \tau | \theta^M](\delta u_L - (1+\delta)u_E) + u_R\big(\Pr[N_M + N_P = \tau | \theta^M] - P^L\big) + u_L P^L \\
& = P^L(u_L - u_R) + \Pr[N_M + N_P = \tau | \theta^M](u_R - u_E + \delta(u_L - u_E)) \,.
\end{aligned}
$$

We plug in $u_L = \alpha^M - 1$, $u_R = -1$ and $u_E = -(1 - \alpha^M(1-r))/4$ and rearrange into

$$
dU = \alpha^M(T\delta + P^L) - \alpha^M(1-r)T(1+\delta)/4 - 3T(1+\delta)/4
$$

$$= \alpha^M \left(T\delta(3/4 + r/4) + P^L - (1 - r)T/4 \right) - 3T(1 + \delta)/4$$

with $T = \Pr[N_M + N_P = \tau | \theta^M]$. Then dU is greater 0 if

$$\alpha^M > \frac{3T(1 + \delta)}{4P^L - (1 - r)T + \delta(3 + r)T} \tag{3.3}$$

Assumption 3.3 ensures that the fraction (3.3) is smaller than one. To sum up, the deviation to L is profitable for the moderate type, if inequality (3.3) is satisfied. □

Proof of Corollary 3.2. The statement of Corollary 3.2 is a direct consequence of Proposition 3.4 with $\delta = 0$. □

Proof of Lemma 3.3. For every prior which is consistent with Assumption 3.3, condition (3.3) is feasible. In particular, $\alpha^M = 1$ ensures condition (3.3). In addition, for every prior consistent with Assumption 3.1, there are α^P and α^M with $\alpha^P < \alpha^M$ that satisfy conditions (3.1) and (3.2). In particular, $\alpha^P = 0$ and $\alpha^M = 1$ guarantee both conditions. □

Bibliography

Aigner, R. (forthcoming). “Environmental Taxation and Redistribution Concerns”. *FinanzArchiv / Public Finance Analysis.*

Aigner, R. and F. Bierbrauer (2013). “Taxing Wall Street: The Case of Boring Banking”. *mimeo.*

Aigner, R. and M. Lang (2013). “Investing Your Vote – On the Emergence of Small Parties”. *mimeo.*

Atkinson, A. B. and J. E. Stiglitz (1976). “The design of tax structure: Direct versus indirect taxation”. *Journal of Public Economics* 6, 55–75.

Auerbach, A. and R. Gordon (2002). “Taxation of Financial Services under a VAT”. *American Economic Review*, 411–416.

Austen-Smith, D. and J. Banks (1988). “Elections, coalitions, and legislative outcomes”. *American Political Science Review* 82, 405–422.

Aytimur, R. E., A. Boukouras, and R. Schwager (2013). “Voting as a Signaling Device”. *Economic Theory* forthcoming.

Bargsted, M. and O. Kedar (2009). “Coalition-targeted duvergerian voting: how expectations affect voter choice under proportional representation”. *American Journal of Political Science* 53, 307–323.

Baron, D. P. and D. Diermeier (2001). “Elections, Governments, and Parliaments in Proportional Representation Systems”. *Quarterly Journal of Economics* 116, 933–967.

Baron, D. P. and J. Ferejohn (1989). “Bargaining in legislatures”. *American Political Science Review* 83, 1181–1206.

Bierbrauer, F. (2013). “Tax incidence for fragile financial markets”. *mimeo.*

Bierbrauer, F. and P. Boyer (2010). “The Pareto-Frontier in a simple Mirrleesian model of income taxation”. *MPI Collective Goods Preprint* 2010/16.

Bierbrauer, F. and M. Hellwig (2011). “Mechanism Design and Voting for Public-Good Provision”. *MPI Collective Goods Preprint* 2011/31.

Boadway, R. and M. Keen (2003). “Theoretical Perspectives on the Taxation of Capital Income and Financial Services”. In: *Taxation of Financial Intermediation.* Ed. by P. Honohan. New York: World Bank and Oxford University Press, 31–80.

Borselli, F. (2009). “A Sensible Reform of the EU VAT Regime for Financial Services”. *International VAT Monitor* 5, 375–383.
Bovenberg, A. L. (1999). “Green tax reforms and the double dividend: an updated reader’s guide”. *International Tax and Public Finance* 6, 421–443.
Bovenberg, A. L. and R. A. de Mooij (1994). “Environmental levies and distortionary taxation”. *American Economic Review* 84, 1085–1089.
——— (1997). “Environmental Levies and Distortionary Taxation: Reply”. *American Economic Review* 87, 252–253.
Bovenberg, A. L. and L. Goulder (2002). “Environmental taxation and regulation”. In: *Handbook of Public Economics*. Ed. by A. J. Auerbach and M. Feldstein. Vol. 3. Amsterdam: Elsevier, 1471–1545.
Brett, C. and J. Weymark (2008). “Public Good Provision And The Comparative Statics Of Optimal Nonlinear Income Taxation”. *International Economic Review* 49, 255–290.
Caminal, R. (1997). “Financial intermediation and the optimal tax system”. *Journal of Public Economics* 63, 351–382.
Castanheira, M. (2003). “Why vote for losers?” *Journal of the European Economic Association* 1, 1207–1238.
Chia, N.-C. and J. Whalley (1999). “The tax treatment of financial intermediation”. *Journal of Money, Credit, and Banking*, 704–719.
Cremer, H., F. Gahvari, and N. Ladoux (1998). “Externalities and optimal taxation”. *Journal of Public Economics* 70, 343–364.
Da Silva, V. A. (2006). “Duverger’s Laws: Between social and institutional determinism”. *European Journal of Political Research* 45, 31–41.
Dahlby, B. (2008). *The Marginal Cost of Public Funds: Theory and Applications*. MIT Press.
de la Feria, R. and B. Lockwood (2012). “Opting for Opting-In? An Evaluation of the European Commission’s Proposals for Reforming VAT on Financial Services”. *Fiscal Studies* 31, 171–202.
Diamond, P. A. and J. A. Mirrlees (1971). “Optimal taxation and public production I: Production efficiency”. *American Economic Review* 61, 8–27.
Diermeier, D. and A. Merlo (2004). “An empirical investigation of coalitional bargaining procedures”. *Journal of Public Economics* 88, 783–797.
Dixit, A. (1985). “Tax Policy in Open Economies”. In: *Handbook of Public Economics*. Ed. by A. Auerbach and M. Feldstein. Vol. 1. North-Holland, Amsterdam; New York: Elsevier, 313–374.
Downs, A. (1957). *An economic theory of democracy*. New York: Harper Collins.
Duverger, M. (1951). *Les Partis politiques*. Paris: Colin.
——— (1954). *Political parties*. New York: John Wiley.

Edgar, T. (2001). "Exempt treatment of financial intermediation services under a value-added tax: an assessment of alternatives". *Canadian tax journal* 49, 1133–1219.

Ekins, P. and S. Dresner (2004). *Green taxes and charges: reducing their impact on low-income households.* Joseph Rowntree Foundation.

Fullerton, D. (1997). "Environmental levies and distortionary taxation: Comment". *American Economic Review* 87, 245–251.

Fullerton, D. and G. Metcalf (1998). "Environmental taxes and the double-dividend hypothesis: Did you really expect something for nothing?" *Chicago-Kent Law Review.*

Gauthier, S. and G. Laroque (2009). "Separability and public finance". *Journal of Public Economics* 93, 1168–1174.

Genser, B. and P. Winker (1997). "Measuring the Fiscal Revenue Loss of VAT Exemption in Commercial Banking". *FinanzArchiv / Public Finance Analysis* 54, 563–585.

Goulder, L. (1995). "Environmental taxation and the double dividend: a reader's guide". *International Tax and Public Finance* 2, 157–183.

Grubert, H. and J. Mackie (2000). "Must financial services be taxed under a consumption tax?" *National Tax Journal* 53, 23–40.

Guesnerie, R. (1998). *A contribution to the pure theory of taxation.* Cambridge University Press.

Hamlin, A. and C. Jennings (2011). "Expressive political behaviour: Foundations, scope and implications". *British Journal of Political Science* 41, 645–670.

Hammond, P. (1979). "Straightforward individual incentive compatibility in large economies". *Review of Economic Studies* 46, 263–282.

Hellwig, M. (2010). "A generalization of the Atkinson-Stiglitz (1976) theorem on the undesirability of nonuniform excise taxation". *Economics Letters* 108, 156–158.

HM Treasury (2009). *Tax ready reckoner and tax reliefs.* London.

Holmstrom, B. and J. Tirole (1997). "Financial intermediation, loanable funds, and the real sector". *Quarterly Journal of economics* 112, 663–691.

Honohan, P. (2003). "Avoiding the Pitfalls in Taxing Financial intermediation". In: *Taxation of Financial Intermediation.* Ed. by P. Honohan. New York: World Bank and Oxford University Press, 1–30.

Huizinga, H. (2002). "A European VAT on financial services?" *Economic Policy* 17, 497–534.

International Montary Fund (2010). "A Fair and Substantial Contribution by the Financial Sector". In: *Financial Sector Taxation – The IMF's Report to the G-20* *and Background Material*. Ed. by S. Claessens, M. Keen, and C. Pazarbasioglu. Washington DC: International Monetary Fund, 2–73.

Jack, W. (2000). "The Treatment of Financial Services under a Broad-Based Consumption Tax". *National Tax Journal* 53, 841–852.

Jacobs, B. and R. A. de Mooij (2011). "Pigou Meets Mirrlees: On the Irrelevance of Tax Distortions for the Second-Best Pigouvian Tax". *CESifo Working Paper Series* 3342.

Jacobs, B. (2012). "A Contribution to the Theory of the Marginal Cost of Public Funds". *mimeo.*

Kaplow, L. (2012). "Optimal control of externalities in the presence of income taxation". *International Economic Review* 53, 487–509.

Keen, M. (2011). "Rethinking the taxation of the financial sector". *CESifo Economic Studies* 57, 1–24.

Keen, M., R. Krelove, and J. Norregaard (2010). "The Financial Activities Tax". In: *Financial Sector Taxation – The IMF's Report to the G-20 and Background Material*. Ed. by S. Claessens, M. Keen, and C. Pazarbasioglu. Washington DC: International Monetary Fund, 118–143.

Kopczuk, W. (2003). "A note on optimal taxation in the presence of externalities". *Economics Letters* 80, 81–86.

Kosonen, K. (2012). "Regressivity in environmental taxation: myth or reality?" In: *Handbook of Research on Environmental Taxation*. Ed. by J. Milne and M. S. Andersen. Cheltenham: Edward Elgar, 161–174.

Kricheli, R., Y. Livne, and B. Magaloni (2011). "Taking to the Streets Theory and Evidence on Protests under Authoritarianism". *mimeo.*

Lockwood, B. (2014). "How Should Financial Intermediation Services be Taxed?" In: *Taxation and Regulation of the Financial Sector*. forthcoming. MIT Press.

McCuen, B. and R. B. Morton (2010). "Tactical coalition voting and information in the laboratory". *Electoral Studies* 29, 316–328.

Meade, J. (1952). "External economies and diseconomies in a competitive situation". *Economic Journal* 62, 54–67.

Meffert, M. F. and T. Gschwend (2010). "Strategic coalition voting: Evidence from Austria". *Electoral Studies* 29, 339–349.

——— (2011). "Polls, Coalition Signals, and Strategic Voting: An Experimental Investigation of Perceptions and Effects". *European Journal of Political Research*, 1–32.

Meirowitz, A. and K. W. Shotts (2009). "Pivots versus signals in elections". *Journal of Economic Theory* 144, 744–771.

Meirowitz, A. and J. A. Tucker (2013). "People Power or a One-Shot Deal? A Dynamic Model of Protest". *American Journal of Political Science* 57, 478–490.

Metcalf, G. (1999). "A distributional analysis of green tax reforms". *National Tax Journal* 52, 655–681.

——— (2003). "Environmental levies and distortionary taxation: Pigou, taxation and pollution". *Journal of Public Economics* 87, 313–322.

Mirrlees, J. A. (1971). "An exploration in the theory of optimum income taxation". *Review of Economic Studies*, 175–208.

——— (1972). "On producer taxation". *Review of Economic Studies* 39, 105–111.

Oates, W. (1995). "Green taxes: Can we protect the environment and improve the tax system at the same time?" *Southern Economic Journal* 61, 915–922.

Orosel, G. O. and R. Schöb (1996). "Internalizing Externalities in Second-Best Tax Systems". *Public Finance* 51, 242–57.

Palfrey, T. R. (1984). "Spatial Equilibrium with Entry". *Review of Economic Studies* 51, 139–156.

Pigou, A. (1932). *Economics of welfare.* Macmillan. London.

Piketty, T. (2000). "Voting as Communicating". *Review of Economic Studies* 67, 169–191.

Pirttilä, J. and M. Tuomala (1997). "Income tax, commodity tax and environmental policy". *International Tax and Public Finance* 4, 379–393.

Poddar, S. and M. English (1997). "Taxation of financial services under a value-added tax: Applying the cash-flow approach". *National Tax Journal* 50, 89–112.

PricewaterhouseCoopers (2006). *Study to Increase the Understanding of the Economic Effects of the VAT Exemption for Financial and Insurance Services – Final Report to the European Commission.* 2 November 2006.

Rausch, S., G. Metcalf, and J. Reilly (2011). "Distributional impacts of carbon pricing: A general equilibrium approach with micro-data for households". *Energy Economics* 33, 20–33.

Razin, R. (2003). "Signaling and Election Motivations in a Voting Model with Common Values and Responsive Candidates". *Econometrica* 71, 1083–1119.

Sandmo, A. (1975). "Optimal taxation in the presence of externalities". *Swedish Journal of Economics* 77, 86–98.

Schöb, R. (1997). "Environmental taxes and pre-existing distortions: The normalization trap". *International Tax and Public Finance* 4, 167–176.

——— (2005). "The Double Dividend Hypothesis of Environmental Taxes: A Survey". In: *The International Yearbook Of Environmental And Resource Economics 2005/2006.* Ed. by H. Folmer and T. Tietenberg. Cheltenham: Edgar Elgar, 223–279.

Shikano, S., M. Herrmann, and P. W. Thurner (2009). "Strategic Voting under Proportional Representation: Threshold Insurance in German Elections". *West European Politics* 32, 634–656.

Sobbrio, F. and P. Navarra (2010). "Electoral participation and communicative voting in Europe". *European Journal of Political Economy* 26, 185–207.

Solsten, E., ed. (1994). *Austria: A Country Study*. Washington: GPO for the Library of Congress.

Sydsaeter, K. et al. (2008). *Further mathematics for economic analysis*. Financial Times Prentice Hall.

The Economist (2008). "Let them heat coke – How green taxes hurt the poor". *The Economist* June 12th, 2008.

West, S. (2005). "Equity implications of vehicle emissions taxes". *Journal of Transport Economics and Policy* 39, 1–24.

Weymark, J. (1987). "Comparative static properties of optimal nonlinear income taxes". *Econometrica* 55, 1165–1185.

White, S., M. Wyman, and S. Oates (1997). "Parties and voters in the 1995 Russian Duma election". *Europe-Asia Studies* 49, 767–798.

Zee, H. H. (2005). "A New Approach to Taxing Financial Intermediation Services Under a Value–Added Tax". *National Tax Journal* 58, 77–92.