Essays in Public Economic Theory

Inaugural-Dissertation
zur Erlangung des Grades

Doctor oeconomiae publicae (Dr. oec. publ.)

im Jahr 2005
an der Volkswirtschaftlichen Fakultät der
Ludwig-Maximilians-Universität München
vorgelegt von

Marco Sahm

Referent: Prof. Dr. Bernd Huber
Korreferent: Prof. Dr. Klaus Schmidt
Promotionsabschlussberatung: 8. Februar 2006

Für Eva

Contents

Preface

Public economic theory analyzes the role of the state within the economy. As in many other textbooks, Musgrave and Musgrave (1984) identify three main functions of the public sector: allocation, stabilization and distribution. The revenue required to meet these demands is mainly collected by taxes. Along those lines one arrives immediately at the categorization according to which public economics is taught also at the University of Munich: tax policy (Steuerpolitik), allocation policy (Allokationspolitik), stabilization policy (Fiskalpolitik), and distribution policy (Sozialstaat).

The essays embedded in the dissertation at hand belong to the fields of taxation and allocation. Accordingly, the book basically consists of two self-contained parts: Part I deals with methods of capital gains taxation, Part II with public goods provision in large economies.

In a world with ever-growing financial markets, the question how to treat capital gains for tax purposes is getting more and more important. And in an economy like Germany, being on its way from a pay-as-you-go pension system to a funded one, it is also a question that affects more and more people. Should the gains from an asset, for example a stock fund, be taxed periodically as they accrue or only at the time when the fund is sold and gains or losses are actually realized? What are the Pros and Cons for one method of taxation or the other? How should the tax scheme look like in order to overcome the disadvantages of one method without running into the problems of the other? How is the market price of an asset affected by the method of taxation? And what impact does it have on welfare and its distribution? Those are some of the questions investigated in Part I which contains two chapters.

Chapter 1 is based on my survey article *Imitating accrual taxation on a realization basis* and discusses the design of the capital gains tax schemes proposed in the literature. Unlike other surveys on that topic, it offers a uniform formal way of describing those proposals. The chosen approach allows to easily compare and evaluate them from a theoretic point of view.

Chapter 2 is an extended version of my paper *Methods of capital gains taxation and the impact on asset prices and welfare*. The study provides a comparison between accrual and realization taxation with respect to the resulting asset prices and welfare. As to my knowledge, it is the first article in this field of literature to explicitly incorporate distributional aspects within the welfare analysis.

Part II of this book presents the so-far output of an ongoing project studying the provision of public goods in large economies. This is joint work together with Felix Bierbrauer from *Max Planck Institute for Research on Collective Goods* in Bonn. We aim at combining the literature on public goods provision in the tradition of Clarke-Groves with the literature on optimal income taxation originating from Mirrlees (1971). Our attempt relies on the basic idea that in general the decisions about financing and providing a public good cannot be separated from each other but should be considered simultaneously.

How many highways should be constructed? How much should be spent on armed forces and police? The fundamental task of public goods provision is to find out the amount that fits best the joint requirements of the economy's members. This is a problem of information aggregation: In order to provide the efficient amount, the social planner has to learn about the individuals' valuation for the public good.

However, the individual answers to the questions raised highly depend on the financing scheme: Who has to contribute how much to the provision of these commodities? Usually, the contributions individuals have to make depend on their abilities to generate income. And so do their answers: Their effective valuation for a public good is the result of the interplay between both their tastes and skills.

Part II also consists of two chapters. Chapter 3 is based on our paper *Public goods provision in a continuum economy with two-dimensional heterogeneity*. In such an economy the presence of aggregate uncertainty causes a problem of information aggregation that has been ignored by the literature so far. The solution we offer requires to modify the Samuelson rule for public goods provision according to the extent of skill heterogeneity among the individuals.

In order to derive this result we develop a new solution concept. Chapter 4 is based on our paper *Robustness to sampling*. It introduces this notion of robustness to a general framework providing a useful tool for a great variety of problems of information aggregation in continuum economies with aggregate uncertainty.

Reading the book

The two parts of the book are self-contained and can be read apart. In principal, the same is true for the two chapters of each part as well. However, in order to avoid redundancies, some cross-references are indicated. An appendix at the end of both Chapter 2 and 3 provides the technical details and longer proofs of the derived results. Items which enter the index are defined on the displayed page and italicized at their first appearance.

Acknowledgments

First and foremost I would like to thank my thesis supervisor Bernd Huber. With his frankness and experience he was very inspiring and helpful in choosing a topic and getting started at the beginning of my doctorate. Later, when I began to develop own ideas, he was always encouraging me and gave me the freedom to pursue the chosen projects.

I am deeply indebted to Klaus Schmidt who agreed to serve as second supervisor on my thesis committee. I benefitted a lot from his insightful comments to my tutorial presentations as well as from his outstanding course on contract theory.

Andreas Haufler, who completes my thesis committee as third examiner, has always shown much interest in my work and has taken the time to discuss it. I am very grateful for his constructive hints and proposals as well as for encouraging me to apply for conferences and publications.

Special thanks go to my colleagues at the *Lehrstuhl für Finanzwissenschaft*: To Ulrich Woitek and Frank Westermann, my representative supervisors, for giving me a new angle on numerous economic topics. To Christoph Eichhorn, who has always been the first to share (economic) ideas with – and who is also the co-author of our article *Monopoly Pricing, Demand Quality and Two-Sided Markets*, which is not part of this dissertation. To Marco Runkel for his detailed comments on one of my papers. To Nadine Riedel for helping me proof-reading. To Gregor Gehauf, Florian Wöhlbier, and Günther Oppermann for introducing me to the chair and the department. To Anita Hofmann for keeping the expenditure of administrative work low. To Christiane Starbatty for finding the requested articles for me. To Florian Ranzi for assisting me in case of technical problems.

I would also like to express my gratitude to my teachers Clemens Fuest, Achim Wambach, and Sven Rady. Starting my doctorate as a mathematician, their doctoral courses deepened my economic thinking a lot.

I thank all participants of the faculty's various seminars I presented my work at for vivid discussions and useful comments. The same gratitude I owe to the numerous participants of the conferences at which I had the chance to talk.

Chapter 2 of this thesis is based on my paper *Methods of capital gains taxation and the impact on asset prices and welfare*. I presented an earlier version of this paper at the Annual Congress of the European Economic Association (EEA) 2005 in Amsterdam, at the Annual Meeting of the Austrian Economic Association (NOeG) 2005 in Innsbruck, and at the Annual Meeting of the Verein für Socialpolitik (VfS) 2004 in Dresden.

Chapter 3 is based on the paper *Public goods provision in a continuum economy with two-dimensional heterogeneity*, which is joint work with Felix Bierbrauer. I presented an earlier version of this paper at the Public Economic Theory Meeting (PET) 2005 in Marseille as well as at the Spring Meeting of Young Economists (SMYE) 2005 in Geneva. For this paper, we are also very grateful for inspiring suggestions from Felix Bierbrauer's supervisor Martin Hellwig.

Of course, special thanks go to my co-author Felix Bierbrauer himself. As a close friend, he has been the one to arouse my interest not only for the specific topic of public goods provision but (public) economics in general. The joint work has been both a lot of fun and a rewarding experience.

Most importantly I am indebted to my family. With their support for and trust in all my decisions and plans, my parents Heidi and Jürgen Sahm helped me a lot. The same notably applies to my wife Eva Sahm. She has always encouraged my work as if it were her own one.

Part I

Methods of capital gains taxation

In most of the world's economies changes in the value of an investor's asset (or entire portfolio) are subject to a tax, the so-called capital gains tax. From a theoretic point of view, there are basically two different methods of collecting this tax: taxation of capital gains upon *accrual* or upon *realization*.

Under an accrual system – sometimes also referred to as 'yield-to-maturity' approach – the tax is payable, in theory, as soon as there is a change in the value of an asset (e.g. by a change in the asset price) or, in practice, periodically. Among others the most severe problems that arise under an accrual tax are those of liquidity and valuation. Some investors might be forced to sell some of their assets, which they would keep hold of otherwise, just in order to pay the tax. For some assets that are not frequently or not publicly traded it can be very costly if not impossible to permanently or periodically assess their value.

For those practical reasons, assets for which such problems arise are mostly taxed upon realization. Under a realization system – sometimes also referred to as 'wait and see' approach – the tax is payable only when the investor sells the asset thereby realizing a gain or a loss. Solving the problems of liquidity and valuation, the realization tax creates a new problem of its own: It equips the investor with a *timing option* that enables him to realize capital losses immediately and defer capital gains in order to save taxes. This so-called *lock-in effect* distorts the investor's optimal liquidation policy and hence possibly his investment decision.

Considering the possible distortions caused by the lock-in effect, there is a natural question arising: Does taxation of capital gains upon realization create a welfare loss? Put differently: Is social welfare smaller under a realization tax than under an accrual tax?

There is a whole branch of the literature on capital gains taxes that implicitly answers in the affirmative and, hence, searches for tax systems avoiding or, at least, reducing the lock-in effect. Chapter 1* surveys the corresponding articles that are found in the economic as well as in the tax law literature. The proposals made are mostly based upon the idea of imitating an accrual tax by *retrospective taxation* on a realization basis in order to circumvent the lock-in effect without running into the problems of liquidity and valuation.

The main contribution of Chapter 1 consists in developing a uniform formal way of describing the different proposals. This uniform approach does not only allow to easily compare the proposed tax schemes and recognize their common grounds or differences but also evaluate and categorize them on an abstract level.

*Chapter 1 is based on my paper *Imitating accrual taxation on a realization basis.*

Chapter 2** returns to the question how social welfare is affected by the method of taxation. The timing option provided by a realization tax has important implications for the investors' optimal consumption and saving behavior and, hence, for equilibrium asset prices. Within a simple general equilibrium model of an exchange economy with heterogeneous agents it is shown that asset prices are higher under a realization based tax system than under an accrual one. However, due to distributional effects, total welfare is not necessarily lower.

Insofar Chapter 2 can be regarded as a more rigorous attempt to clarify the conditions under which the proposals for imitating an accrual tax discussed in Chapter 1 should in fact be implemented in order to attain the economy's normative goal.

**Chapter 2 is based on my paper *Methods of capital gains taxation and the impact on asset prices and welfare*. I presented an earlier version of this paper at the Annual Congress of the European Economic Association (EEA) 2005 in Amsterdam, at the Annual Meeting of the Austrian Economic Association (NOeG) 2005 in Innsbruck, and at the Annual Meeting of the Verein für Socialpolitik (VfS) 2004 in Dresden.

Chapter 1

Imitating accrual taxation on a realization basis

While taxation of capital gains upon accrual poses problems of liquidity and valuation, taxation upon realization evokes the so-called lock-in effect that possibly induces distortions of the investors' liquidation and investment decision. Those problems are discussed in Section 1.1. The tax schemes, proposed in the literature in order to remedy simultaneously the shortcomings of both taxation methods, are mostly based on the idea of imitating accrual taxation on a realization basis. Section 1.2 surveys those proposals and develops a uniform formal way of describing the suggested schemes that allows to easily compare and evaluate them from a theoretic point of view. Section 1.3 concludes with some brief remarks on the practical applicability of the discussed proposals.

1.1 The basic problem

1.1.1 Accrual vs. realization taxation

One rationale for tax systems all over the world is the ability-to-pay principle. Since people's skills are not directly observable, the personal income serves as a proxy and mostly provides the taxable base. If capital gains are considered to be part of this tax basis and if the income tax is collected periodically, then, for the sake of consistency, the taxation of capital gains or losses should be based on changes in the values of assets during the corresponding period. Put differently, this would be in accordance with the Haig-Simons definition of income as the sum of consumption and the change in value of property.

The method of taxing capital gains when they actually accrue is referred to as taxation upon accrual or *yield-to-maturity* approach. In many countries it particularly applies to financial instruments with fixed returns. For example, bonds issued at deep discount are taxed upon accrual in the US as well as in many other OECD countries.[1]

For an asset with fixed returns the price path is, more or less, predetermined. Hence, the tax authorities can easily assess the current value and corresponding tax liability, i.e. administrative costs caused by valuation are low. Moreover, since the holder of such an asset knows the taxes due in advance when he purchases it, there should neither be a problem of liquidity.

However, things look different for assets with contingent returns. The stochastic price path of financial instruments like stock or stock options often is quite volatile. It may happen that investors are forced to sell some of their assets, which they would keep hold of otherwise, just in order to pay the tax due. Moreover, it might be quite costly if not impossible to permanently or periodically assess the current value of volatile assets with contingent returns that are not frequently or not publicly traded. Like Warren (2004) or Shakow (1986), most authors consider the problems of *liquidity* and *valuation* to be the most severe drawbacks of accrual taxation and the main reasons for the fact that in most countries capital gains from assets with contingent returns are taxed upon realization.[2]

Under a realization system – also referred to as *wait-and-see* approach – the tax is payable only when the capital gain or loss is actually realized, i.e. when the asset is sold or otherwise disposed of. On the one hand, this solves the problems of liquidity and valuation. On the other hand, of course, the realization method provides an inconsistency for an income tax otherwise assessed periodically. This inconsistency has several important consequences for the investors' economic behavior.[3]

First, the realization tax provides a timing option that allows the investor to realize capital losses immediately and defer capital gains in order to save taxes. This distortion of the optimal liquidation policy is known as lock-in effect and will be explained in Section 1.1.2 in some more detail.

Second, in the presence of several investment opportunities, the investors' *portfolio choice* and investment decision may also be distorted as shown in Section 1.1.3.

Section 1.1.4 demonstrates that, if the method of taxation is not uniform but relies on distinctions, such as the difference between fixed and contingent returns mentioned above, the problem of distortions is heightened by *financial*

[1]See OECD (1994); for a numerical example see also Warren (2004).

[2]See e.g. OECD (1994).

[3]A comprehensive discussion about the pros and cons of the realization requirement can be found in Land (1996).

innovations that undermine the relevant distinction.

Finally, if ordinary income is taxed periodically while capital gains are taxed upon realization, this inconsistency may affect even the production decision, since relative prices for the factors capital and labor possibly change. However, the problems arising from the different character (whether ordinary income or capital gain) are discussed at length in the literature on dual income taxation[4] and will not be addressed any further in this book. Likewise, problems arising from distinctions in source (whether domestic or foreign) or voice (whether debt or equity) of capital gains are ignored.[5]

The study at hand focusses on the timing of taxation under the two controversial methods. In order to simplify the formal analysis, this chapter abstracts from many specific real world conditions and makes the following assumptions with respect to the assets, the capital market, and the tax code:[6]

Assumption 1.1 *The assets under consideration do not produce any intermediate cash flows, and their shares are infinitely divisible.*

Assumption 1.2 *On the capital market, investors are price takers and trade only at equilibrium prices. Transaction costs are zero. There are no restrictions on short sales.*

Assumption 1.3 *All capital gains and losses are taxed at the constant rate $0 < \tau < 1$. No distinction is made between the short term and long term status of capital gains and losses.*

1.1.2 The optimal liquidation policy

In order to illustrate the lock-in effect arising under a realization tax, Constantinides (1983) investigates the optimal liquidation policy of an investor. Thus the discussion focusses on the timing of income under a realization tax.

He considers a single good exchange economy with a risky asset (stock), which is taxed upon realization at the rate τ, and a riskless, tax exempt bond with yield $r > 0$.[7] Suppose the actual price for one share of stock at time t to be P_t and let $T > t$ be some arbitrary future time. Consider the problem

[4] See e.g. Sørensen (1994), Alstadsæter (2003), Fjaerli and Lund (2001).

[5] See Warren (2004) for a short discussion of those distinctions creating problems for income taxation. A very similar categorization of the basic weaknesses of current tax systems can be found in Alworth (1998).

[6] The assumptions are similar to those made in Constantinides (1983).

[7] Equivalently, one could assume that the bond is taxed upon accrual and yields the after tax rate r.

of an investor, who holds one share of stock with basis P_0 purchased at some previous time 0 and aims at maximizing the after-tax value of his investment at time T deciding on his liquidation policy. In particular, compare the following two strategies: Either (H) hold the asset till time T, which then yields

$$W_H := P_T - \tau(P_T - P_0) = (1-\tau)P_T + \tau P_0,$$

or (R) sell and repurchase the asset at time t. To repurchase the asset after taxation upon realization at time t in case of a realized gain, i.e. for $P_t > P_0$, the investor must borrow the tax payment $\tau(P_t - P_0)$ at the riskless rate r by going short in the bond; in case of a loss, i.e. for $P_t < P_0$ he can lend this amount buying shares of the bond. Compounding interest continuously, his after tax proceeds at time T are given by

$$\begin{aligned} W_R &:= P_T - \tau(P_T - P_t) - \tau(P_t - P_0)e^{r(T-t)} \\ &= (1-\tau)P_T + \tau P_0 - \tau(P_t - P_0)(e^{r(T-t)} - 1). \end{aligned}$$

A comparison of those values shows that

$$W_H \geq W_R \quad \Longleftrightarrow \quad P_t \geq P_0,$$

i.e. in case of accrued gains the 'hold'-strategy is preferable whereas in case of losses the 'sell and repurchase'-strategy should be chosen.[8] Put differently, taxation upon realization provides a timing option that should be exerted only in the case of capital losses, whereas gains are locked-in. Therefore, this result is labelled the lock-in effect. Since $T > t$ has been chosen arbitrarily, it can be summarized as follows:

Theorem 1.1 (Constantinides (1983)) *If Assumptions 1.1–1.3 hold, then the optimal liquidation policy under a realization tax is to realize losses immediately and defer gains as long as possible.*

1.1.3 The impact on portfolio choice

The previous section has demonstrated that taxation upon realization causes a distortion of the investor's liquidation decision. In this section the question under consideration is whether the lock-in effect may distort the investor's portfolio choice as well. Following Auerbach (1991), it is shown that in the

[8] Note that this conclusion is independent of the investor's expectations on the future asset price P_T.

presence of alternative investment opportunities there are realistic circumstances under which the answer is yes.

Look at the following two-period version of the above framework with two riskless assets: A, yielding the rate of return ρ, and B, yielding the rate of return i. Consider an investor, who holds one share of asset A with actual price P_1 and basis P_0. Suppose that a capital gain has accrued ($P_1 > P_0$). The investor aims at maximizing the after-tax value of his investment in period 2 deciding on the structure of his portfolio. In particular, compare the following two strategies: Either (H) hold asset A till period 2, which then yields

$$\begin{aligned} W_H &= P_1(1+\rho) - \tau\left[P_1(1+\rho) - P_0\right] \\ &= P_1(1+\rho) - \tau\left[(P_1 - P_0) + \rho P_1\right], \end{aligned}$$

or (R) sell the asset A and reinvest in asset B yielding

$$\begin{aligned} W_R &= \left[P_1 - \tau(P_1 - P_0)\right]\left[1 + i(1-\tau)\right] \\ &= P_1(1+i) - \tau\left[(P_1 - P_0)\left[1 + i(1-\tau)\right] + iP_1\right]. \end{aligned}$$

A comparison of those values for identical rates of return $\rho = i$ shows, again, the preferability of the 'hold' strategy:

$$W_H - W_R = \tau(P_1 - P_0)i(1-\tau) > 0.$$

Since $\frac{\partial W_H}{\partial \rho} = (1-\tau)P > 0$, the after tax yield W_H is increasing in the rate of return ρ, and hence there is a cut-off value $\rho_{\min} < i$ such that for all ρ with $\rho_{\min} < \rho < i$ the inequality $W_H > W_R$ still holds. Economically spoken, the investor is willing to keep hold of asset A even for a range of pre-tax returns that fall short of the pre-tax return of the alternative asset B. In such cases, the realization tax in fact distorts the investor's portfolio choice: Tax considerations lock-in the accrued gains and avoid an efficient reallocation of the portfolio.

1.1.4 The role of financial innovations

As emphasized by many authors, e.g. Alworth (1998), Boadway and Keen (2003) or Warren (2004), financial innovations often aggravate the distorting character of a realization tax for the following reason: If the method of taxation is not uniform but relies on distinctions, such as the difference between fixed and contingent returns mentioned above, the investors might be able to circumvent the relevant distinction by the use of new financial instruments.

The notion of financial innovations is not always used consistently in the literature. According to Alworth (1998) it includes the invention of new trading strategies as well as the design of new assets and synthetic portfolio positions for the purposes of risk management and (tax) arbitrage.

The 'sell and repurchase' strategy that has been proven to be favorable in the case of losses (see Section 1.1.2) may be seen as an example of such an innovative trading strategy. It is labelled a *wash-sale*, because the transaction is conducted just in order to qualify for a tax credit washing away the losses. However, the tax codes of most countries like the US do not allow a tax credit for transactions identified as wash sales.[9]

The most prominent example for a synthetic portfolio position that undermines the distinction between fixed and contingent returns is the famous *put-call-parity* for European options.[10] An European option is a contract which provides the buyer with the right to buy (call) or sell (put) a specific quantity of an underlying asset (e.g. stock) at a specific price K on a specified future date T from or to the seller respectively. Let S, P and C be the actual prices for a share of stock that does not pay dividends, a put option, and a call option with the same strike price K and expiration date T, respectively. Moreover, let i be the pre-tax rate of return of a zero coupon bond with maturity at date T. The assumption that no arbitrage is possible on the financial market implies the following identity, known as put-call-parity:

$$K(1+i)^{-T} = S + P - C. \tag{1.1}$$

Hence, investing an amount of $K(1+i)^{-T}$ in the zero coupon bond is just equivalent to the synthetic position of buying a share of stock as well as a put option and selling short a call option. Though each of the assets within this synthetic position has contingent returns, their combination generates a fixed one that equals the return of the zero coupon bond. Put differently, the put-call-parity makes the distinction between fixed and contingent returns obsolete.

However, while the return of the zero coupon bond is usually taxed upon accrual, the gains or losses arising from each of the assets within the synthetic portfolio are usually taxed upon realization (see Section 1.1.1). This means that the investor himself can decide, which method of taxation should be applied to a position with fixed return i by investing either in the zero coupon bond or the synthetic portfolio. In particular, he can pick realization taxation providing him with the timing option even for instruments with fixed returns.

[9] See e.g. Warren (2004).

[10] See e.g. Alworth (1998), Boadway and Keen (2003), Warren (2004).

1.2 Formulaic taxation

1.2.1 The different solution concepts

Warren (2004) categorizes the actual and proposed solutions to the problems discussed in the previous section as follows: *transactional analysis*, *anti-avoidance provisions*, taxation of changes in the market values, and *formulaic taxation*.

Transactional analysis

Transactional analysis means "to analyze the components of a new transaction in order to achieve consistent treatment with other, more familiar, assets" (Warren, 2004, p. 904). The approach includes *disaggregation* (also labelled *bifurcation*) as well as *integration* of (new) financial instruments, as described by Shuldiner (1992) or Warren (1993).

For example, the put-call-parity (1.1) suggests to integrate the transactions of selling (short) a share of stock as well as the corresponding put option and buying the corresponding call option in a joint position that should be taxed like the disposal of the equivalent zero coupon bond.[11] As pointed out by Warren (1993), it might be possible in theory but is surely difficult in practice to properly identify such offsetting positions, since it requires case by case considerations inducing high administrative costs.

Anti-avoidance provisions

Anti-avoidance provisions are regulations "that authorize disallowance of tax benefits in certain transactions motivated by tax avoidance" (Warren, 2004, p. 920). The basic idea behind those rules is the so-called *substance over form principle*: Transactions should be taxed according to their economic substance rather than their judicial or contractual form.

For example, as already mentioned in the preceding section, most tax codes do not allow a tax credit for wash-sales: The transactions embodied in the corresponding 'sell and repurchase' strategy do not really alter the economic position of the investor but are conducted for tax purposes only. However, in practice it is not always as easy as for a wash-sale to identify whether a transaction is mainly motivated by tax avoidance. As for the transactional analysis, it often requires case by case considerations inducing high administrative costs (see Warren (1993)).

[11] See Warren (2004) for a rich variety of other examples of transactional analysis.

Besides the practical problems of difficult identification and expensive bureaucracy, the solution concepts of transactional analysis and anti-avoidance provisions are affected with a major theoretical drawback: They might be useful to maintain the tax-relevant distinctions, such as the difference between fixed and contingent returns, but they do not solve the basic problem of capital gains taxation upon realization, namely the inconsistency in timing and the resulting distortions of the investors' liquidation decision and portfolio choice. By the very nature of this problem it cannot be solved just altering the institutional framework at an ad-hoc basis but solely by a comprehensive adjustment in the taxation method itself. The following two types of solution represent such concepts (see Alworth et al. (2002)).

Taxation of changes in the market values

The most naive possibility would be, of course, to shift completely to a pure accrual tax, i.e. (periodically) taxing the changes in the market value of an asset independent of a realization event. However, the problems of liquidity and valuation discussed in Section 1.1.1 have made many economists sceptical whether this approach is viable (see Alworth et al. (2002)).

Nevertheless Shakow (1986) argues in favor of a wider use of accrual taxation. For many assets and transactions, he proposes solutions to the drawbacks of liquidity and valuation. In particular, for publicly traded assets with observable market prices, valuation poses no problem: their value can be permanently assessed by 'marking them to market'. That is the reason why taxation upon accrual is sometimes also referred to as *mark to market* approach. In fact, in many countries like the US (see Warren (2004)) or Italy (see Alworth et al. (2002)), accrual taxation is applied to a growing variety of assets which meet those requirements.

However, the mark to market approach does not solve the liquidity problem and cannot be used for assets without publicly observable prices. Therefore, "the realization requirement is widely considered to be essential to make the tax system administrable" (Land, 1996, p. 48).

Formulaic taxation

Sticking to the requirement of raising taxes on a realization basis one has to deal with the lock-in effect and the induced distortions of the investor's liquidation and investment decision (see Sections 1.1.2 and 1.1.3). The proposed solutions to that problem suggest to carry on collecting taxes upon realization but attempt to assess the tax due in such a way as to avoid or, at least, reduce the distortions mentioned above. Because the tax liability is calculated from the realized gain by a certain tax formula, Warren (2004) refers to this approach as formulaic taxation. The basic idea behind formulaic taxation consists in imitating accrual taxation on a realization basis by imputing a certain rate of return on the deferred gains (see Boadway and Keen (2003) as well as Warren (2004)).

The remainder of this section surveys the various tax formulas suggested in the corresponding literature. Inspired by the work of Boadway and Keen (2003), it is the first article in that field to offer a uniform formal way of describing the different proposed tax schemes.

The several proposals differ with respect to the specific rates of return used to accumulate gain on the one hand and imputed on deferred gains on the other hand as well as with respect to the informational requirements concerning the price path and the holding period of the taxable asset. Those parameters are decisive for both the practical applicability[12] and the extend to which the lock-in effect and the resulting distortions can be overcome.

The big advantage and the main contribution of the work at hand lies in the fact that by using a uniform formal description, those parametric differences become apparent directly from the tax formulas. This allows to compare and evaluate them from a purely theoretic point of view.

Each of the following subsections has a closer look to one proposal of formulaic taxation. For each proposed tax scheme the corresponding tax formulas both in continuous and discrete time are derived and illustrated by a numerical example. Moreover, the economic intention of each approach as well as its success in combatting the valuation problem on the one hand and the lock-in effect on the other hand are discussed.

In the formal analysis, the tax liability of an investor who, at time T, sells an asset purchased at some previous time 0 is computed. To this end, Assumptions 1.1–1.3 are supposed to hold; in particular, the relevant assets are assumed to generate no intermediate cash-flows. The asset price at time $0 \leq s \leq T$ is denoted P_s and the riskless rate of return i, for simplicity, supposed to be constant over time. The formulas in continuous time are derived by just calculating the relevant integral; the discrete time formulas are computed by evaluating the corresponding sum making consistently use of the following relation for *geometric series*: for any constant c,

$$\sum_{s=1}^{T} c^s = c\frac{1-c^T}{1-c}. \tag{1.2}$$

[12] Of course, the practical applicability is lower the higher the informational requirements.

In order to illustrate how the tax formulas work, a numerical example is provided throughout employing the following parameters:

$$T = 2, \quad P_0 = 100, \quad P_1 = 180, \quad P_2 = 200, \quad i = 0,1, \quad \tau = 0,25. \tag{1.3}$$

Before discussing eight different proposals for formulaic taxation made in the literature, two benchmark cases are analyzed: the status quo, i.e. taxation upon realization, and a hypothetical scheme that serves as a reference case, because it is completely equivalent to an accrual tax.

1.2.2 The status quo: realization taxation

Usually, the tax liability under a realization tax is just the difference between the selling price and the purchase price times the tax rate: $\tau(P_T - P_0)$. While the drawbacks of realization taxation have been discussed at length, note that its informational requirements are very low: only the purchase price and the selling price must be observed.

However, additionally knowing the length of the holding period T, one can determine the average internal rate of return. This can be used to state the following identities that will prove useful for later references.

Continuous time formula

Computing the average internal rate of return in continuous time g yields

$$g = \frac{1}{T} \ln \left(\frac{P_T}{P_0} \right). \tag{1.4}$$

The implicit value of the asset at time s is then given by $P_s^{\text{impl}} = P_T e^{-g(T-s)}$. Consequently,

$$\text{Tax}^c_{\text{real}} = \int_0^T \tau g P_T e^{-g(T-s)} ds = \tau(P_T - P_0), \tag{1.5}$$

expressing that if the realized gains are distributed uniformly all over the holding period, the realization method will be equal to permanently taxing the implicit gains $gP_T e^{-g(T-s)}$ at the rate τ (without imputing any interest on the deferred tax payments).

Discrete time formula

The average internal rate of return in discrete time γ is given by

$$\gamma = \sqrt[T]{\frac{P_T}{P_0}} - 1, \tag{1.6}$$

the implicit value of the asset at time s by $P_s^{\text{impl}} = P_T(1+\gamma)^{-(T-s)}$, and hence

$$\text{Tax}^d_{\text{real}} = \sum_{s=1}^{T} \tau P_T(1+\gamma)^{-T}\left[(1+\gamma)^s - (1+\gamma)^{s-1}\right] = \tau(P_T - P_0) \qquad (1.7)$$

with an interpretation analogous to that given above.

Example

For the values specified in (1.3) the realization tax due is given by

$$\text{Tax}_{\text{real}} = 0,25 \times (200 - 100) = 25.$$

1.2.3 The reference case: full equivalence

If liquidity was the only problem under an accrual tax, in theory it would be possible to adjust taxes paid upon realization to those that would have accrued if the asset had been marked to market. The necessary adjustment is determined by ex post calculating the accrued gain and implicit tax due for each tax period between purchase and realization of the asset, and imputing the net-of-tax period-by-period internal rate of return on the deferred tax payments.[13] Since such a scheme would perfectly mimic an accrual tax on a realization basis in the sense that "after paying the adjusted tax at realization the investor's terminal wealth from investing into a particular asset would be equivalent to the terminal wealth under accrual taxation" (Alworth et al., 2002, p. 4), it is referred to as *full equivalence* method.

Continuous time formula

In continuous time, the value of the asset at time s times the internal rate of return at that date leads to the gain $g_s P_s$ at time s. This gain is taxed at rate τ and the deferred tax payment is compounded forward to the time of realization T at the internal rate of return, leading to the following formula:

$$\text{Tax}^c_{\text{acc}} = \int_0^T e^{\int_s^T g_t(1-\tau)dt} \tau g_s P_s ds, \qquad (1.8)$$

[13] See Alworth et al. (2002). Note that holding the asset for one more period can be interpreted in terms of taxation as 'reinvesting' the deferred tax payments into the asset. Hence the implicit tax due appreciates at the internal rate of return. However, since it should not be taxed again in the subsequent periods, the net-of-tax return applies.

where $g_s := \frac{dP_s/ds}{P_s}$ is the internal rate of return at time s. Usually there is no hope that the price path of the asset is in fact differentiable.[14] Therefore, the integral should not be taken literally but rather understood as an illustrative notation similar to that used in stochastic integration.[15]

Discrete time formula

In discrete time, charging the tax due as described above leads to the following formula:

$$\text{Tax}_{\text{acc}}^{d} = \sum_{s=1}^{T} \tau(P_s - P_{s-1}) \prod_{t=s+1}^{T} [1 + \gamma_t(1-\tau)], \tag{1.9}$$

where $\gamma_s := \frac{P_s}{P_{s-1}} - 1$ is the internal rate of return in period s.

Example

For the values specified in (1.3) the tax due under the full equivalence method is given by

$$\begin{aligned} \text{Tax}_{\text{acc}} &= 0,25 \times (180 - 100) \left[1 + \frac{200}{180}(1 - 0,25)\right] + 0,25 \times (200 - 180) \\ &\approx 26,67. \end{aligned}$$

Discussion

Eliminating the lock-in effect, however, the full equivalence method would require to observe the entire price path between purchase and realization of the asset, i.e. not solve the problem of valuation. Hence, in practice it would be applicable only where the mark to market approach applies anyway.

1.2.4 Retrospective taxation at the actual rate (Vickrey, 1939)

Full, i.e. ex-post, equivalence between taxation upon accrual and realization is sufficient but not necessary to eliminate the lock-in effect. Vickrey (1939) proposes a retrospective tax scheme that is rather similar to the full equivalence method. But instead of imputing the period-by-period internal rate

[14] For example, if, as usual, the price path of common stock is modelled to follow an Itô-process (see e.g. Welcker et al. (1992)), it would be nowhere differentiable with probability 1 (see e.g. von Weizsäcker and Winkler (1990)).

[15] See e.g. von Weizsäcker and Winkler (1990).

of return on the deferred tax payments, it uses the riskless rate of return. This makes investors indifferent between accrual and realization based taxes ex-ante, i.e. at the moment they decide whether to sell the asset or keep on holding it, because it equates the certainty-equivalent after-tax returns (see Alworth et al. (2002) and Auerbach (1991)). Imputing interest on the deferred tax payments at the riskless rate could be interpreted as the government lending these virtual payments to the investor for him to reinvest them in the asset.

Continuous time formula

In continuous time, the tax due is given by

$$\text{Tax}_V^c = \int_0^T e^{i(1-\tau)(T-s)} \tau g_s P_s ds, \tag{1.10}$$

where, again, $g_s := \frac{dP_s/ds}{P_s}$. The remark made on the continuous time formula under the full equivalence method applies here as well.

Discrete time formula

In discrete time, the tax due amounts to

$$\text{Tax}_V^d = \sum_{s=1}^{T} \tau(P_s - P_{s-1})[1 + i(1-\tau)]^{T-s}. \tag{1.11}$$

Example

For the values specified in (1.3) the tax due under retrospective taxation a la Vickrey (1939) is given by

$$\begin{aligned} \text{Tax}_V &= 0,25 \times (180 - 100)\,[1 + 0,1 \times (1 - 0,25)] + 0,25 \times (200 - 180) \\ &= 26,5. \end{aligned}$$

Discussion

Eliminating the lock-in effect, however, with the same informational requirement to observe the entire price path of the asset, the applicability of this approach does not exceed the one under the full equivalence method.

1.2.5 Retrospective taxation at the constant rate (Meade, 1978)

In order to reduce the informational requirements, Meade (1978) proposes to retrospectively allocate the realized gain as if the asset would have appreciated uniformly at the implicit average internal rate of return over the holding period and impute the net-of-tax riskless rate of return on the corresponding virtual tax payments. Hence, only the purchase and sell prices P_0 and P_T as well as the length of the holding period T have to be observable.

Continuous time formula

For taxation in continuous time this leads to a tax payment of

$$\begin{aligned} \text{Tax}^c_M &= \int_0^T e^{i(1-\tau)(T-s)} \tau g P_T e^{-g(T-s)} ds \\ &= \begin{cases} T\tau g P_T & \text{if} \quad g = i(1-\tau) \\ \frac{\tau g}{g-i(1-\tau)}(1-e^{[i(1-\tau)-g]T})P_T & \text{else} \end{cases}, \end{aligned} \tag{1.12}$$

with g defined as in (1.4).

Discrete time formula

In discrete time the tax formula is given by

$$\begin{aligned} \text{Tax}^d_M &= \sum_{s=1}^{T} \tau P_T[(1+\gamma)^{-(T-s)} - (1+\gamma)^{-(T-(s-1))}][1+i(1-\tau)]^{T-s} \\ &= \begin{cases} T\tau \frac{\gamma}{1+\gamma} P_T & \text{if} \quad \gamma = i(1-\tau) \\ \frac{\tau\gamma}{\gamma-i(1-\tau)}(1-[\frac{1+i(1-\tau)}{1+\gamma}]^T)P_T & \text{else} \end{cases}, \end{aligned} \tag{1.13}$$

where (1.2) is used and γ is defined as in (1.6).

Example

For the values specified in (1.3) the tax due under retrospective taxation a la Meade (1978) is given by

$$\begin{aligned} \text{Tax}_M &= 0,25 \times \left(\frac{200}{\sqrt{2}} - \frac{200}{2}\right)[1+0,1\times(1-0,25)] + 0,25\times\left(200 - \frac{200}{\sqrt{2}}\right) \\ &\approx 25,78. \end{aligned}$$

Discussion

From a comparison of the formulas (1.12) and (1.5) as well as (1.13) and (1.7) respectively, it becomes apparent that the status quo taxation upon realization can be interpreted as an application of retrospective taxation with an imputed interest of $i = 0$.

Unfortunately, the proposed method can only alleviate but not eliminate the lock-in effect. Like in the example at hand, for an asset with above-normal rates of return initially, the investor still has an incentive to keep on holding it because this allows him to spread the uniform accrual pattern retrospectively imputed for his gain over several periods. “Likewise, an asset that had declined in value would offer its owner the incentive to sell” (Auerbach, 1991, p. 168).

1.2.6 Continuous yield to maturity (Land, 1996)

Land (1996) proposes a tax formula that requires to know only the purchase and realization price of the asset but is still able to eliminate the lock-in effect. His scheme might be regarded as a version of retrospective taxation similar to that of Meade (1978). But instead of imputing the riskless rate of return on the deferred gains, it compounds interest on the internal rate of return. This may be interpreted as if the government would become an equity partner entitled to a share of the investor’s asset equal to the deferred tax payments.

Continuous time formula

The proposed formula is given by

$$\begin{aligned} \mathrm{Tax}_L^c &= \int_0^T e^{g(1-\tau)(T-s)} \tau g P_T e^{-g(T-s)} ds \\ &= P_T - P_0^\tau P_T^{1-\tau}, \qquad (1.14) \end{aligned}$$

where, again, g is defined as in (1.4). Interestingly, for the method to be able to eliminate selective realization and, hence, the lock-in effect, *continuous* compounding is necessary. The name of the proposal is due to that observation. The reason why there is a deferral advantage remaining in discrete time lies in the fact that interest is compounded on the deferred tax payment for an accrued gain only for all subsequent periods but not for the period in which this gain accrues. With decreasing length of the taxation period this error decreases and vanishes in the limit as the length tends towards zero.

Discrete time formula

However, for the sake of completeness, the discrete time version of the proposed method is provided as well:

$$\begin{aligned}\text{Tax}_L^d &= \sum_{s=1}^{T} \tau P_T[(1+\gamma)^{-(T-s)} - (1+\gamma)^{-(T-(s-1))}][1+\gamma(1-\tau)]^{T-s} \\ &= P_T - P_0[1+\gamma(1-\tau)]^T, \qquad (1.15)\end{aligned}$$

where, again, (1.2) is used and γ is defined as in (1.6).

Example

For the values specified in (1.3) the tax due according to the continuous time formula is given by

$$\text{Tax}_L^c = 200 - 100^{0,25} \times 200^{1-0,25} \approx 31,82.$$

Using the discrete time formula one derives

$$\begin{aligned}\text{Tax}_L^d &= 0,25\times\left(\frac{200}{\sqrt{2}} - \frac{200}{2}\right)\left[1+(\sqrt{2}-1)(1-0,25)\right] + 0,25\times\left(200-\frac{200}{\sqrt{2}}\right) \\ &\approx 28,22.\end{aligned}$$

Discussion

Despite its low informational requirements and the ability to circumvent the lock-in effect, the suggested method is not free from shortcomings. Besides the administrative burden that it puts for assets with intermediate cash-flows (see Land (1996) or Warren (2004)) its implementation would certainly create severe problems of political acceptance. As one can see from a glance at formula (1.14), the proposed scheme is not only not linear but not even monotone in the realized gain either. Figure 1.1 illustrates, for example, that an investment of P_0 would produce no refund if it became worthless over time, whereas the effective marginal tax rate for growing gains tends towards 100 %.

1.2.7 Expected value taxation (Shuldiner, 1992)

The basic idea behind *expected value taxation* put forward by Shuldiner (1992) is to tax anticipated gains on an accrual basis but unanticipated gains upon realization. This requires to determine an expected rate of return for

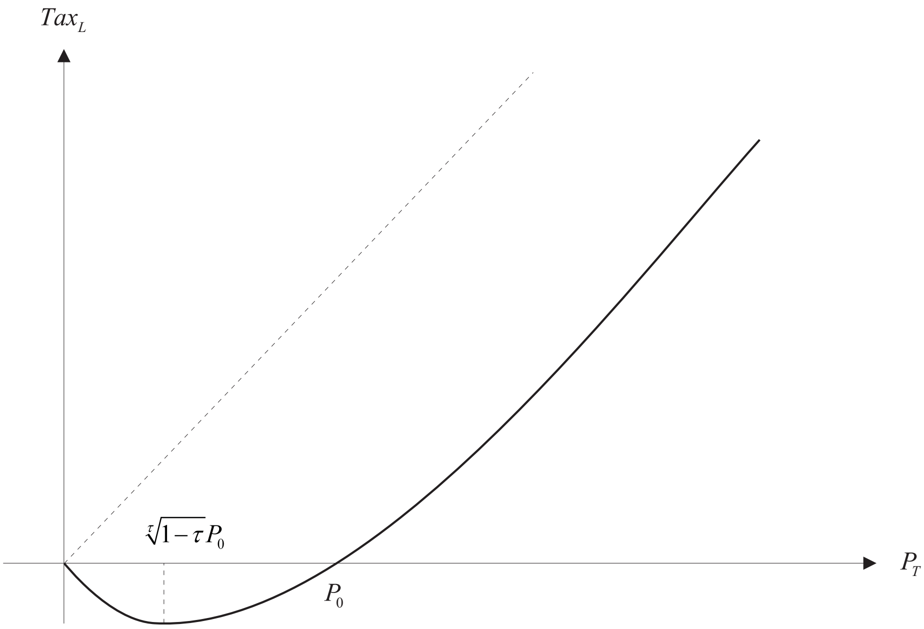

Figure 1.1: Tax under the continuous yield to maturity approach

each asset with contingencies and then periodically tax the virtual (anticipated) gain that accrues according to this predefined rate till the date the actual gain is realized. At realization, when contingencies are resolved, the unanticipated gain is taxed as the difference between actual sell price and expected value of the asset according to the predefined expected rate of return.

Continuous time formula

Let r be the expected (average) rate of return. At time s the expected value of the asset is then given by $P_0 e^{rs}$. According to the original proposal of Shuldiner (1992), the resulting virtual gain rP_0e^{rs} should then be taxed upon accrual at the rate τ. To compute the final value for this stream of tax payments at the time of realization T, compound interest at the net-of-tax riskless rate of return. Adding the tax for the unanticipated gain yields

$$\text{Tax}_S^c = \int_0^T e^{i(1-\tau)(T-s)} \tau r P_0 e^{rs} ds + \tau(P_T - P_0 e^{rT})$$

$$= \begin{cases} T\tau r P_0 e^{rT} + \tau(P_T - P_0 e^{rT}) & \text{if } r = i(1-\tau) \\ \frac{\tau r}{r - i(1-\tau)}(e^{rT} - e^{i(1-\tau)T})P_0 + \tau(P_T - P_0 e^{rT}) & \text{else.} \end{cases} \quad (1.16)$$

Discrete time formula

Analogously one derives the discrete time formula

$$\begin{aligned} \text{Tax}_S^d &= \sum_{s=1}^{T} \tau P_0[(1+\rho)^s - (1+\rho)^{s-1}][1+i(1-\tau)]^{T-s} \\ &+ \quad \tau[P_T - P_0(1+\rho)^T] \\ &= \begin{cases} T\tau\frac{\rho}{1+\rho}P_0(1+\rho)^T + \tau[P_T - P_0(1+\rho)^T] & \text{if } \rho = i(1-\tau) \\ \frac{\tau\rho}{\rho - i(1-\tau)}([1+\rho]^T - [1+i(1-\tau)]^T)P_0 & \\ + \quad \tau[P_T - P_0(1+\rho)^T] & \text{else,} \end{cases} \end{aligned} \quad (1.17)$$

where, again, (1.2) is used, and now the expected (average) rate of return is denoted ρ.

Example

For the values specified in (1.3) and additionally assuming $\rho = i$, the tax due according to expected value taxation is given by

$$\begin{aligned} \text{Tax}_S &= 0,25 \times 100 \times \left[(1,1-1)\left[1+0,1\times(1-0,25)\right] + (1,1^2 - 1,1)\right] \\ &\quad +0,25 \times (200 - 100 \times 1,1^2) \\ &\approx 25,19. \end{aligned}$$

Discussion

One big challenge with this approach is to determine an appropriate expected rate of return. Given such a rate, the informational requirements are low – only the purchase and sell price have to be known – but choosing this rate in a suitable way may require to learn prices (e.g. forward prices) that are not publicly observable (see Shuldiner (1992) or Warren (2004)).

Comparing the formulas (1.16) and (1.5) as well as (1.17) and (1.7) respectively, the status quo realization tax can be regarded as the particular case of expected value taxation in which the expected rate of return is supposed to equal zero ($r = \rho = 0$), i.e. the entire gain is considered unanticipated. On the other hand, if the expected rate of return equals the actual internal rate of return ($r = g$ or $\rho = \gamma$), the tax liability under expected value taxation equals the one under retrospective taxation a la Meade (1978), as a comparison of (1.16) and (1.12) or (1.17) and (1.13), respectively, shows. The special case in which the expected rate of return equals the riskless rate of return[16] is discussed in Section 1.2.10.

However, even if an appropriate expected rate of return is available, expected value taxation can only alleviate but not remedy the lock-in effect since the incentive to defer unanticipated gains still remains (see Shuldiner (1992) or Warren (2004)).

1.2.8 Imputing interest on basis (Cunningham and Schenk, 1992)

The proposal made by Cunningham and Schenk (1992) is based upon the idea that ex-ante, an investor who purchases an asset assumes it to appreciate at least at the riskless rate of return. Otherwise he could invest at the riskless rate. Hence, from an ex-ante point of view, the tax liability should be (at least) as high as for an asset appreciating at the riskless rate. Consequently, the authors suggest to simply impute the riskless interest rate on the basis of an asset and tax the gains that would accrue if the asset appreciated at the riskless rate of return. If the tax payments are deferred till realization, interest will be compounded at the net-of-tax riskless rate.

Continuous time formula

This procedure yields the following continuous time formula:

$$\begin{aligned} \text{Tax}^c_{CS} &= \int_0^T e^{i(1-\tau)(T-s)} \tau i P_0 e^{is} ds \\ &= P_0 \left[e^{iT} - e^{i(1-\tau)T} \right]. \qquad (1.18) \end{aligned}$$

Discrete time formula

Applying (1.2), the discrete time formula is given by

$$\text{Tax}^d_{CS} = \sum_{s=1}^{T} \tau P_0 [(1+i)^s - (1+i)^{s-1}][1 + i(1-\tau)]^{T-s}$$

[16] Auerbach (1991) argues that under the assumption of optimal portfolio choice by the investors, in equilibrium the certainty-equivalent value of the expected rate of return must equal the riskless rate of return.

$$= \quad P_0 \left[(1+i)^T - (1+i(1-\tau))^T\right]. \qquad (1.19)$$

Example

For the values specified in (1.3), the tax due according to the method suggested is given by

$$\begin{aligned} \text{Tax}_{CS} &= 0,25 \times 100 \times \left[(1,1-1)\left[1+0,1\times(1-0,25)\right] + (1,1^2-1,1)\right] \\ &\approx 5,44. \end{aligned}$$

Compared to the liabilities under the other schemes the tax payment is quite low due to the fact that only the virtual gain that would have accrued at the riskless rate is taxed but not the much higher actually realized one.

Discussion

Although under the proposed scheme an investor's tax liability does not depend on his accrued gain at all, the method is not able to eliminate the lock-in effect. The investor has an incentive to selectively realize losses and defer gains in order to qualify for a tax basis as low as possible. Despite the very low informational requirements – only the purchase price and the length of the holding period must be observable – installing such a scheme would presumably cause severe problems of political acceptability. Its ex-ante character, i.e. the fact that the tax liability does in no way reflect the ex-post realized gain, is usually not compatible with the traditional notions of fairness and solidarity that try to find compensation ex-post, after contingencies have been resolved. A similar critique applies to the method proposed in the next section (see Auerbach (1991)).

1.2.9 Retrospective taxation at the riskless rate (Auerbach, 1991)

In his seminal article *Retrospective Capital Gains Taxation* Auerbach (1991) proposes a tax scheme very similar to the previous method. But instead of imputing interest on the basis, he suggests to retrospectively tax the asset as if it would have appreciated to its realization price at the riskless rate of return.

Continuous time formula

In continuous time, the virtual asset price at time s is derived by discounting interest on the realization price yielding $P_T e^{-i(T-s)}$. The virtual gain $iP_T e^{-i(T-s)}$ at time s is then taxed at rate τ. Imputing the net-of-tax riskless rate of return on the deferred tax payments leads to the following tax formula:

$$\begin{aligned}\text{Tax}_A^c &= \int_0^T e^{i(1-\tau)(T-s)} \tau i P_T e^{-i(T-s)} ds \\ &= P_T \left[1 - e^{-i\tau T}\right]. \end{aligned} \tag{1.20}$$

Discrete time formula

The discrete time formula is derived by an analogous procedure and applying (1.2):

$$\begin{aligned}\text{Tax}_A^d &= \sum_{s=1}^{T} \tau P_T \left[(1+i)^{-(T-s)} - (1+i)^{-(T-(s-1))}\right] [1 + i(1-\tau)]^{T-s} \\ &= P_T \left[1 - \left(\frac{1+i(1-\tau)}{1+i}\right)^T\right]. \end{aligned} \tag{1.21}$$

Example

For the values specified in (1.3), the tax due according to the method suggested is given by

$$\begin{aligned}\text{Tax}_A &= 0,25 \times 200 \times \left[\left(\frac{1}{1,1} - \frac{1}{1,1^2}\right)[1 + 0,1 \times (1 - 0,25)] + \left(1 - \frac{1}{1,1}\right)\right] \\ &\approx 8,99. \end{aligned}$$

Here as well, the tax payment is relatively low because only the virtual gain, that would have been necessary for the asset to accumulate at the riskless rate to the selling price, is taxed but not the much higher one actually realized.

Discussion

Using the assumption of optimal portfolio choice, Auerbach (1991) shows that facing this tax scheme, in equilibrium an investor is ex-ante indifferent between either holding the asset for one more period or selling the asset and reinvesting the proceeds at the riskless rate of return. Hence, despite its low informational requirements – only the sell price P_T and the length of the holding period T must be observable – the method is able to eliminate the lock in effect and the induced distortions.[17] However, ex-post the tax liability does not depend on the gain actually realized. Therefore, with respect to concerns of fairness, the scheme faces a similar critique as the proposal of Cunningham and Schenk (1992).

1.2.10 Arbitrary gain reference date (Bradford, 1995)

This critique might be regarded as the motivation for the work of Bradford (1995). Introducing the notion of a *gain reference date (GRD)*, he proposes a tax scheme, which is able to eliminate the lock-in effect with comparable low informational requirements but still allows to charge tax for the actual realized gain. The method may be looked upon as a combination and enhancement of the proposals made by Auerbach (1991), Shuldiner (1992), as well as Cunningham and Schenk (1992) and works as follows:

Continuous time formula

Before the investor makes any transaction, a so-called arbitrary gain reference date D is fixed. If an asset, purchased at basis P_0, is held for the time T and then sold at price P_T, in order to determine the resulting gain at date D, the basis would be compounded forward and the sell price would be discounted back to that date at the riskless rate. This 'unanticipated' gain as of time D, i.e. the difference $P_T e^{-i(T-D)} - P_0 e^{iD}$ is taxed at rate θ, which is also set in advance and may differ from the tax rate τ for 'anticipated' gains. The total tax liability at realization is then composed of three parts: (1) the tax on the 'unanticipated' gain compounded forward from the GRD D to the realization date T at the net-of-tax riskless rate of return, and the tax on 'anticipated' gains – compounded forward at the same rate – resulting from (2) imputing the riskless rate of return on the basis P_0 from the time of purchase 0 till the GRD D (see equation (1.18)) and (3) discounting the riskless rate of return from the sell price P_T from the time of realization T back to the GRD D (see equation (1.20)).

$$\begin{aligned}
\text{Tax}_B^c &= \theta\left[P_T e^{-i(T-D)} - P_0 e^{iD}\right] e^{i(1-\tau)(T-D)} \\
&+ \ P_0\left[e^{iD} - e^{i(1-\tau)D}\right] e^{i(1-\tau)(T-D)} \\
&+ \ P_T\left[1 - e^{-i\tau(T-D)}\right]
\end{aligned}$$

[17] In fact, this scheme is the only one to eliminate the lock in effect getting along with these informational requirements (see Auerbach (1991) and Auerbach and Bradford (2002)).

$$
\begin{aligned}
&= P_T\left[1-(1-\theta)e^{-i\tau(T-D)}\right] \\
&- P_0 e^{i(1-\tau)T}\left[1-(1-\theta)e^{i\tau D}\right]
\end{aligned} \tag{1.22}
$$

Discrete time formula

In discrete time, an analogous procedure yields the following tax formula:

$$
\begin{aligned}
\mathrm{Tax}_B^d &= \theta\left[P_T(1+i)^{-(T-D)} - P_0(1+i)^D\right][1+i(1-\tau)]^{T-D} \\
&+ P_0\left[(1+i)^D - (1+i(1-\tau))^D\right][1+i(1-\tau)]^{T-D} \\
&+ P_T\left[1-\left(\frac{1+i(1-\tau)}{1+i}\right)^{T-D}\right] \\
&= P_T\left[1-(1-\theta)\left(\frac{1+i(1-\tau)}{1+i}\right)^{T-D}\right] \\
&- P_0[1+i(1-\tau)]^T\left[1-(1-\theta)\left(\frac{1+i(1-\tau)}{1+i}\right)^{-D}\right].
\end{aligned} \tag{1.23}
$$

Example

For the values specified in (1.3) and additionally assuming $D=1$ to be the GRD and the tax rate for anticipated and unanticipated gains to coincide, i.e. $\theta=\tau$, the tax due according to the proposed scheme is given by

$$
\begin{aligned}
\mathrm{Tax}_B &= 0,25\times\left(\frac{200}{1,1}-100\times 1,1\right)[1+0,1\times(1-0,25)] \\
&\quad +0,25\times 100\times(1,1-1)[1+0,1\times(1-0,25)] \\
&\quad +0,25\times 200\times\left(1-\frac{1}{1,1}\right) \\
&\approx 26,53.
\end{aligned}
$$

Discussion

Consider some special cases with respect to the parameters D and θ. First note by a comparison of equations (1.20) and (1.22) or (1.21) and (1.23) respectively, that for $D=\theta=0$, i.e. if the gain reference date is set to equal the purchase date and unanticipated gains remain untaxed, the proposed tax scheme will reduce to the one of Auerbach (1991). Second, if $\theta=0$ and $D=T$, the tax formulas (1.22) and (1.23) will equal those under the approach of Cunningham and Schenk (1992) given in (1.18) and (1.19), respectively. Finally, if $\theta = \tau$ and $D = T$, the tax formulas (1.22) and (1.23) will equal those under the approach of Shuldiner (1992) as given in (1.16) and (1.17) respectively, assuming that the expected rate of return r or ρ respectively, is equal to the riskless rate i. Note however, that the GRD is set in advance and cannot be influenced by the investors realization decision, i.e. $D = T$ may happen only by chance. Exactly this is the reason why the approaches of Shuldiner (1992) and Cunningham and Schenk (1992) fail to eliminate the lock-in effect, whereas the method of Bradford (1995) succeeds.

As mentioned, the proposed method requires only little information – the prices P_0, P_T and the length of the holding period T must be known –, eliminates the lock-in effect, and alleviates the problem of ex-post fairness by taxing the actual gain. Despite those advantages the political acceptance for this approach would presumably be quite low because its rather complex tax formula would counteract the effort made in order to reach more simplicity and transparency in the tax code.

1.2.11 Generalized cash-flow taxation (Auerbach and Bradford, 2002)

The approach suggested by Auerbach and Bradford (2002) is somewhat odd in this review of income tax schemes since it actually represents a cash-flow tax. However, as the authors show, it is equivalent to an income tax collected upon accrual. The basic idea of *generalized cash-flow taxation* can best be understood reconsidering the tax formulas (1.22) or (1.23), respectively, under the approach of Bradford (1995).

Continuous time formula

The difference in equation (1.22), for example, can be regarded as consisting of a (cash-flow) tax at rate $\left[1-(1-\theta)e^{-i\tau(T-D)}\right]$ on the sale at time T and a deduction of the present value of a (cash-flow) tax at rate $\left[1-(1-\theta)e^{i\tau D}\right]$ on the purchase at time 0. It seems natural to disentangle those taxes by collecting two separate tax payments via a cash-flow tax. In continuous time, a positive (negative) cash-flow P_s at time s would then cause a tax liability (credit) according to the following formula:

$$\text{Tax}^c_{AB} = P_s\left[1-(1-\theta)e^{-i\tau(s-D)}\right], \tag{1.24}$$

where D and θ are parameters fixed in advance that could be given a similar interpretation to that in the scheme of Bradford (1995).[18]

Discrete time formula

In discrete time, the analogous formula is given by

$$\mathrm{Tax}^d_{AB} = P_s \left[1 - (1-\theta) \left(\frac{1+i(1-\tau)}{1+i} \right)^{s-D} \right]. \tag{1.25}$$

Example

For the values specified in (1.3), with $D = 1$ and $\theta = \tau$ the generalized cash flow tax generates a credit of

$$\mathrm{Tax}^{AB}_0 = 100 \times \left[1 - (1 - 0,25) \times \left(\frac{1,1}{1 + 0,1 \times (1 - 0,25)} \right) \right] \approx 23,26$$

at purchase and a liability of

$$\mathrm{Tax}^{AB}_2 = 200 \times \left[1 - (1 - 0,25) \times \left(\frac{1 + 0,1 \times (1 - 0,25)}{1,1} \right) \right] \approx 53,41$$

at realization. The final value of the tax payments adds up to

$$\mathrm{Tax}_{AB} = \mathrm{Tax}^{AB}_2 - \mathrm{Tax}^{AB}_0 \times [1 + 0,1 \times (1 - 0,25)]^2 \approx 26,53,$$

the same value as under the approach of Bradford (1995).

Discussion

Note that the cash-flow tax rate under the proposed scheme is increasing over time, and that at just the right rate to simulate an accrual tax at rate τ (see Auerbach and Bradford (2002)). Therefore, the approach is able to eliminate the lock-in effect.[19]

The informational requirement is very low – only the amount of cash-flow must be observable – but installing the proposed scheme would mean to switch from an income tax system to a cash-flow tax. Such a switch is frequently discussed but in most countries far from being realized.

[18]See Auerbach and Bradford (2002) for a discussion how the values for these parameters should be chosen appropriately.

[19]However, a rising cash-flow tax rate discourages current deductible investment (see Auerbach and Bradford (2002)).

1.2.12 Schematic comparison and evaluation

The most important characteristics of the proposed tax schemes analyzed in this chapter are (1) their informational requirements, (2) the rate of return used to assess the accumulated (real or virtual) gain, and (3) the rate of return imputed on deferred tax payments. Those characteristics are decisive for the practical applicability of the method as well as for its ability to eliminate the lock-in effect. Table 1.1 is summarizing those attributes and the corresponding success in combatting the induced distortions.

Proposal	Informational requirements	Rate of return used to accumulate gains	Rate of return imputed on deferred taxes	Eliminates the lock-in effect
Status quo	P_0, P_T	g	0	no
Full equivalence	$(P_s)_{s=0}^T$	$(g_s)_{s=0}^T$	$[(1-\tau)g_s]_{s=0}^T$	yes
Vickrey (1939)	$(P_s)_{s=0}^T$	$(g_s)_{s=0}^T$	$(1-\tau)i$	yes
Meade (1978)	P_0, P_T, T	g	$(1-\tau)i$	partly
Land (1996)	P_0, P_T, T^a	g	$(1-\tau)g$	yes[b]
Shuldiner (1992)	P_0, P_T, T	r	$(1-\tau)i$	partly
Cunningham and Schenk (1992)	P_0, T	i	$(1-\tau)i$	no
Auerbach (1991)	P_T, T	i	$(1-\tau)i$	yes
Bradford (1995)	P_0, P_T, T	i	$(1-\tau)i$	yes
Auerbach and Bradford (2002)	P_s	—	$(1-\tau)i$	yes

[a] T is needed only for taxation in discrete time.
[b] The lock-in effect is eliminated entirely only for taxation in continuous time.

Table 1.1: Characteristics of the proposed tax schemes

1.3 Concluding remarks

This chapter has surveyed the proposals made in the literature in order to simultaneously solve the problems of accrual and realization taxation based on the idea of imitating an accrual tax on a realization basis. In order to describe the different tax formulas, a uniform formal approach has been introduced, which allows to easily compare and evaluate the proposed schemes from a purely theoretic point of view. As the proposal made by Bradford (1995) shows, in theory it is possible to overcome the problems of an accrual system, namely liquidity and valuation, as well as those of a realization one, namely the lock-in effect, at the same time and still tax the actually realized gain. However, this comes at the cost of a rather complex and opaque tax formula that is not very likely to be implemented in practice.

In order to keep the analysis simple a list of assumptions has been made (see Assumptions 1.1–1.3) abstracting from numerous and often varying real world conditions. In particular, the problems arising from the following three realities have been neglected: (1) assets that produce intermediate cash-flows, e.g. dividends; (2) tax rates or interest rates that vary over time or from individual to individual; (3) credits that authorities might allow for inflation. How the respective schemes have to be modified in order to cope with these issues is usually discussed in the corresponding original articles. However, such a discussion is beyond the scope of the survey at hand. Here the focus is rather on basic concepts than on workable tax codes.

So far there has been only little attempt to really implement formulaic capital gains taxation. Alworth et al. (2002) describe the Italian experience of having experimented with both the mark to market approach proposed by Shakow (1986) (see Section 1.2.1) as well as introducing retrospective taxation of capital gains along the lines suggested alternatively by Vickrey (1939), Meade (1978), Auerbach (1991), Bradford (1995), and Auerbach and Bradford (2002). The Italian 1998 tax reform "introduced an accruals based regime in a number of situations. Where this was not feasible, various types of retrospective capital gains taxation were introduced with the purpose of 'equalizing' realizations based taxes with those resulting from a system based on accruals" (Alworth et al., 2002, p. 3). This 'equalizer' could be regarded as a version of the tax formulas proposed by Vickrey (1939) and Meade (1978) and was introduced with delay in January 2001. Though it worked quite well creating only few distortions,[20] the 'equalizer' was first suspended and

[20] Alworth et al. (2002, p. 3) assert in an empirical study that the distortions from the 'equalizer' "do not appear to be very significant."

then abolished only seven months after its introduction due to problems of political acceptance. Alworth et al. (2002, p. 29) judge that as a "warning signal" and comment on it as follows:

> "There exists a gap between the methods and concepts used by economists in analyzing issues like the 'lock-in' effect and the methods and concepts that can be used in the political and legal arena for upholding ex-post adjustments of realized income. Even the more sophisticated adjustments such as those put forward by Auerbach are not immune from political pressures and 'ex-post' equity considerations. ... [I]t is difficult to argue in the political arena that net-returns would have been the same as under the accrual system if portfolio allocations had been changed as would be dictated by rational behavior."

Put differently, if politics aims at a successful introduction of efficient formulaic taxation, the lessons to be learned from the Italian experience suggest to improve the social acceptance for 'ex-ante' concepts of fairness as well as for complex tax formulas.

An entirely different question is whether it is desirable after all to replace the status quo method of realization taxation by an accrual tax or an imitating scheme. One might argue that the preferential tax treatment provided by the advantage of deferral has social value. In a *first-best* world, as Auerbach (1991, p. 173) shows, a "tax benefit for capital assets need not be provided via a distortionary deferral advantage." However, in a *second-best* world in which the method of taxation is the only *policy instrument* available, things are not that clear. Chapter 2 will throw some light on the impact the method of capital gains taxation has on (the distribution of) welfare.

Chapter 2

The impact on asset prices and welfare

Taxation of capital gains upon realization instead of accrual provides incentives to hold winners as long as possible and sell losers immediately. This so-called lock-in effect possibly distorts the liquidation and investment decision and hence is usually regarded as harmful. This chapter analyzes the impact the method of taxation has on asset prices and welfare within a simple general equilibrium model of an exchange economy with heterogeneous agents. It is shown that asset prices are higher under a realization based tax system than under an accrual one. However, due to distributional effects, total welfare is not necessarily lower.

2.1 The basic problem revisited

As explained in Chapter 1.1, solving the problems of liquidity and valuation arising under an accrual tax, the taxation of capital gains upon realization possibly distorts the investors' liquidation decision and portfolio choice. The following example reconsiders this problem and shows that the lock-in effect may arise in an even simpler framework.

The owner of an asset with basis P_0, actual price P_1 and final payout P_2 decides on either selling and repurchasing the asset in period 1 or holding the asset till period 2 in order to maximize his period-2-payout after taxation at the constant rate $0 < \tau < 1$.

Under an accrual tax the investor obviously is indifferent between the two strategies. Both of them leave him with the same after-tax payout in period 2 equal to

$$W^{\text{acc}} := \frac{P_1 - \tau(P_1 - P_0)}{P_1} \left[P_2 - \tau(P_2 - P_1)\right].$$

Note that under the *hold strategy* he still has to liquidate part of the asset in period 1 in order to fulfill his tax liability. Under a realization system the after-tax payout in period 2 following the 'hold'-strategy is

$$W_H^{\text{real}} := P_2 - \tau(P_2 - P_0),$$

whereas the *sell-and-repurchase strategy* yields

$$W_R^{\text{real}} := \frac{P_1 - \tau(P_1 - P_0)}{P_1} \left[P_2 - \tau(P_2 - P_1)\right],$$

the same payout as under an accrual tax. A comparison of the two strategies shows that

$$W_H^{\text{real}} \geq W_R^{\text{real}} \quad \Leftrightarrow \quad (P_0 \geq P_1 \geq P_2) \vee (P_0 \leq P_1 \leq P_2). \tag{2.1}$$

If one assumes that, as in most of the relevant cases, the investor expected the asset to appreciate when he purchased it and still does in period 1 ($P_0, P_1 \leq P_2$), the optimal liquidation policy according to (2.1) suggests to choose W_H^{real} if $P_0 \leq P_1$ and W_R^{real} otherwise, i.e. defer gains (as long as possible) and realize losses (immediately).

As shown in Section 1.1.2, the same result is derived by Constantinides (1983) from a similar situation (see Theorem 1.1), but where firstly the asset is risky, secondly there is an alternative investment opportunity, which thirdly is taxed upon accrual, i.e. different taxation methods coexist.[1] The above analysis shows that the result hinges on neither of those additional assumptions.[2]

However, following Auerbach (1991) and additionally assuming the existence of an alternative investment opportunity, Section 1.1.3 proves that the investor finds it optimal to keep hold of an asset with accrued capital gains instead of selling it and buying the alternative one even for some (expected) pre-tax rates of return smaller than the alternative pre-tax rate. This indicates that besides the distortion of the optimal liquidation policy a realization-based tax possibly leads to inefficient portfolio selection and a distortion of the investment decision.

Usually both distorting effects arising from taxing capital gains upon realization are summarized and labelled the lock-in effect. Nevertheless for analytical purposes it is worth while distinguishing between one and the other:

[1] The assumption of a riskless, tax exempt bond by Constantinides (1983) is equivalent to the assumption of a riskless bond taxed upon accrual, which yields the same after tax rate.

[2] However, the above analysis requires the assumption that the investor expects the asset to appreciate ($P_0, P_1 \leq P_2$), which is not necessary for the result in Constantinides (1983) to hold (see Section 1.1.2).

The effect on the optimal liquidation policy is always present and referred to as *primary lock-in effect* for the remainder of this chapter. The effect on the investment decision arises only in the presence of alternative investment opportunities and is referred to as *secondary lock-in effect*. Note again that the assumptions of uncertainty and the coexistence of different taxation methods are not necessary for those effects to occur. Hence the analysis stated below surrenders these assumptions in order to isolate the lock-in effect from possibly additional effects due to risk[3] and the concomitance of different taxation methods. Moreover, to begin with, only the impact of the primary lock-in effect is investigated.

Considering the distortions caused by the lock-in effect, there is a natural question arising: Does taxation of capital gains upon realization do harm creating a welfare loss? Put differently: Is social welfare smaller under a realization tax than under an accrual tax? The answer usually given in the economic literature is yes[4], but the reasoning is rather based on heuristic considerations than proper analysis in a formal model (e.g. Kovenock and Rothschild (1987)). The present paper tries to fill the gap and examines the question more closely.

Of course a welfare analysis within the framework used in the above example, where asset prices are exogenously given, is not very fruitful as it neglects the impact a specific method of capital gains taxation has on asset prices. To take this *price effect* into account but still keep the analysis tractable, a simple general equilibrium model of an exchange economy with heterogeneous agents is investigated. It is shown that in the presence of accrued capital gains asset prices are higher under a realization tax than under an accrual system. Since the realization system creates incentives to defer accrued gains to later periods, actual total demand for the asset and thus its price increase. However, the impact the method of taxation has on welfare is ambiguous: Due to distributional effects, total welfare is not necessarily smaller under a realization system than under an accrual one. While a realization system discriminates agents without accrued capital gains, it is in favor of individuals holding assets with such gains.

The remainder of this chapter is structured as follows: Section 2.2 offers a short review of the related literature. Section 2.3 specifies the model by introducing some basic and technical assumptions and establishes the consumer's problem of utility maximization under different regimes of capital gains taxation. The analysis shows that comparative statics results within

[3]For the relation between *risk taking* and capital gains taxation see e.g. Sandmo (1985).

[4]This positive answer is the implicit motivation for the proposed tax schemes surveyed in Chapter 1.

the existing literature often are due to special assumptions, mostly with respect to the consumers' utilities. The details of this analysis are provided in the Appendix 2.6. In Section 2.4 the impact of the method of taxation on asset prices and welfare is analyzed by comparing an accrual based system with a realization tax. The results are illustrated in Appendix 2.7 by the example of quasi-linear logarithmic preferences within a slightly extended version of the model. Section 2.5 concludes discussing some possible extensions of the model.

2.2 Review of the literature

While the papers discussed in Chapter 1 ask how a realization based tax system should look like to circumvent the distortions raised by the lock-in effect, there is another branch of the literature on capital gains taxation that puts a different question: How will portfolio selection and asset prices be affected if capital gains are taxed upon realization?

As in Auerbach (1991) the impact on portfolio selection can be analyzed within a partial equilibrium framework, where (expected) asset prices or, equivalently, (expected) pre-tax rates of return are exogenously given. Balcer and Judd (1987) show that the method of capital gains taxation as well as investor's individual horizons for saving will affect optimal portfolio composition. Similarly, in a simulation model, Dammon et al. (2001) show that the optimal dynamic consumption and portfolio decision is a function of the investor's age, initial portfolio holdings, and tax basis. Kovenock and Rothschild (1987) compute the effective tax rates under a realization system and compare the net returns of different portfolio strategies. However, if one wants to take price effects into account, a general equilibrium model has to be engaged.

As pointed out, for example by Lang and Shackelford (2000), whenever investigating price effects of capital gains taxation one has to be aware of an impact that arises independently from the method of taxation, be it accrual or realization based: A higher tax rate lowers the after-tax return of an asset which in turn results in a lower demand for the assets and hence, given a fixed supply, a lower asset price. This so-called *capitalization effect* is opposed to the lock-in effect that occurs only under a realization based system: A higher tax rate induces bigger incentives to postpone the realization of accrued capital gains resulting in higher demand for those assets and hence, given a fixed supply, higher asset prices. The intuition behind the capitalization effect can also be stated as follows: The owner of an asset with accrued capital gains will sell it only if he is compensated for the tax advantage he foregoes by selling it. Thus prices must rise. Empirical studies mostly show the dominance of the capitalization effect (e.g. Lang and Shackelford (2000), Rendleman and Shackelford (2003)), but also evidence for the lock-in effect can be found (e.g. Landsman and Shackelford (1995)).

There is a broad literature modelling the capitalization effect in different settings under the assumption of an accrual tax.[5] In contrast, so far only few articles exist that explicitly account for the fact that capital gains are usually taxed upon realization, and hence are able to incorporate the lock-in effect.

Constantinides (1983) develops a capital asset pricing model under the assumptions of a realization tax and perfect capital markets. However, using short-selling strategies, in his model investors are able to separate their liquidation decision from their consumption and saving decision and thus to defer tax payments until so-called ‘events of forced liquidation’ (e.g. death). Consequently, the lock-in effect is capitalized in the asset prices only to the extend such events occur.

Stiglitz (1983) shows that under realistic assumptions, by applying sophisticated trading strategies investors on perfect capital markets can avoid not only the payment of realization based capital gains taxes but all income taxes. This provides an indication and Poterba (1987) supports empirical evidence that the assumptions of perfect capital markets, especially the one of unlimited short-selling, are not sustainable if one wants to describe a reality in which investors pay a considerable amount of capital gains taxes.

Klein (1999) engages a general equilibrium model to study the impact of capital gains taxation on asset prices and portfolio selection under the assumptions of imperfect capital markets where short-selling is not allowed. In a multi period setting, finitely many individuals maximize their utility from consumption by periodically deciding on how much to consume and save given their *initial endowments*. The investment opportunities are exogenously given and consist in a riskless asset taxed upon accrual and finitely many risky assets taxed upon realization. His findings can be summarized as follows: The pre-tax returns for assets with accrued capital gains are smaller, i.e. their prices are higher, than for assets without accrued gains. The lock-in effect is capitalized in asset prices and might overcompensate for the capitalization effect. Put differently, asset prices may increase by higher taxes. Moreover, the optimal portfolio selection depends not only on one’s own amount of accrued capital gains and saving horizon but also on the amounts of accrued capital gains and the saving horizons of all other investors.

The complexity of Klein’s framework makes it very hard if not impossible

[5] e.g. Auerbach (1979), Gordon and Bradford (1980), Collins and Kemsley (2000)

to use it for a welfare analysis. His model is rich in the sense that it does not only account for the pure effects of capital gains taxation,[6] but also for possibly additional effects arising from uncertainty and the coexistence of different taxation methods. The idea of the analysis presented below is to simplify the model in order to separate the different effects from each other. The aim is to remain in the position to analyze the impact the method of capital gains taxation has on asset prices but in addition to get into a position that allows to undertake a welfare analysis. The model described in the following sections accomplishes that by renouncing risk and the coexistence of different tax systems. To begin with, only the impact of the primary lock-in effect will be analyzed, i.e. there is a single saving opportunity. Furthermore, as compared to the framework of Klein (1999), the model uses stronger assumptions concerning the initial endowments of the agents, whereas it gets by with much weaker assumptions on consumers' preferences.

2.3 A simple general equilibrium model

In this section, the model outlined above is developed more formally and used to derive some comparative statistics results for different regimes of capital gains taxation.

2.3.1 Basic assumptions

Consider a two period ($t \in \{1, 2\}$) exchange economy with two agents ($i \in \{1, 2\}$), who are price takers and trade only at equilibrium prices. Given their initial endowments, they maximize their utility $U^i(c_1^i, c_2^i)$ from private consumption in period 1, c_1^i, and 2, c_2^i, deciding on how much to save in period 1. To guarantee the existence of a Walrasian equilibrium,[7] assume the preferences to be continuous, strictly convex, strongly monotone, and, moreover, such that consumption in period t is a normal commodity, as it is common for large aggregates.[8] For analytical convenience, assume that preferences can be expressed by twice continuously differentiable utility functions

$$U^i : (\mathbb{R}_0^+)^2 \rightarrow \mathbb{R} \quad \text{such that} \quad \frac{\partial^2 U^i}{\partial {c_t^i}^2} < 0 < \frac{\partial U^i}{\partial c_t^i}.$$

For the time being and the reasons mentioned above let there be a single saving opportunity: one arbitrarily divisible share of an asset with basis

[6] i.e. the capitalization effect and the (primary and secondary) lock-in effect

[7] See e.g. Mas-Colell et al. (1995), Section 17.C.

[8] See e.g. Mas-Colell et al. (1995), Section 2.E.

P_0 and safe payout P_2 in period 2. The asset can be interpreted as a real investment in the following way: At some prior time[9] the amount of P_0 consumption goods has been invested in a project that has a safe payout of P_2 consumption goods in period 2 but cannot be liquidated before (in period 1). The analysis aims at finding the asset price P_1, i.e. the price a share of this investment is traded at in period 1.

2.3.2 Specific assumptions

The most general framework would allow for arbitrary amounts of consumption goods in both periods and an arbitrary division of the share in period 1 as the consumer's initial endowments. However, such a setting would create the problem of identifying the seller and the buyer of the asset respectively. As in Klein (1999) one had to separately look at the cases where consumer 1 is either the buyer or the seller or no trade takes place. To avoid this problem the following assumption is made:

Assumption 2.1 *At the beginning of period* 1*, consumer* 1 *holds one unit of the asset purchased in period* 0 *with basis* $P_0 \leq P_2$ *but no other wealth, whereas consumer* 2 *has initial wealth* W *in consumption goods but no shares. Besides the payout of the asset, none of them has any additional income in period* 2*.*

Therefore, at any equilibrium consumer 1 sells shares to consume in period 1 and consumer 2 buys them to consume in period 2.[10] Admittedly this situation is not the only possible scenario but an economically interesting and relevant one: If one aims at analyzing distributional effects of the taxation method, one has to introduce some source of heterogeneity between the agents. And the easiest way to do so is by different initial endowments. In this model, there is heterogeneity between the individuals with respect to their initial share ownership in a way that fits pretty well the situation at the housing market in the UK.[11] In Great Britain a relatively huge amount of total wealth is held in the form of housing equity (see e.g. Banks et al. (2002)). Accordingly, like in Assumption 2.1, on the one hand there is the group of house owners endowed with a considerable amount of that asset

[9]For the analysis of the accrual tax regime it is convenient to assume that this time lies in period 0 in order to guarantee that no capital gains taxes have been paid for the asset until period 1.

[10]In particular, by Assumption 2.1, short-selling does never take place. Hence, the assumption of no short-selling is not necessary but, of course, additionally could be made like in Klein (1999) or Dammon et al. (2001).

[11]I would like to thank Frank Cowell from LSE for pointing out this example to me.

and, on the other hand, a group of people without housing wealth. Capital gains on that market have a striking influence on wealth distribution and consumption in the UK (see e.g. Henley (1998), Disney et al. (2003)), which explains the vivid discussion about how to tax them.

Under Assumption 2.2, the agents' different initial endowments imply heterogeneity with respect to their accrued capital gains as well.

Assumption 2.2 *In equilibrium the asset price* P_1 *in period* 1 *satisfies*

$$P_0 \leq P_1 \leq P_2.$$

Again, this assumption is primarily made to abbreviate the analysis and avoid case differentiation. But now it is not an assumption on the input but the outcome of the model and thus seems to be quite strong at a first glance. Three different arguments can be made in order to defend it: First of all, it can be shown that for any given utility functions with the above properties there are values for the parameters P_0, P_2 and W such that an equilibrium with $P_0 \leq P_1 \leq P_2$ exists, i.e. the analysis does not refer to the empty set. Secondly, there is no risk in the model and thus a constellation at which the asset is appreciating over time seems to be most plausible to look at. Finally, if one wants to study the influence of accrued capital gains, the relevant situations are those in which at least $P_0 \leq P_1$.

2.3.3 The problem under different tax regimes

In period 1 the asset is traded. Let S^i be the fraction of the asset consumer i possesses after trade has taken place. The market clearing condition requires

$$S^1 + S^2 = 1. \tag{2.2}$$

Capital gains are taxed at the constant rate $0 \leq \tau < 1$ in each period after trade or rather payout and before consumption have taken place. Now the consumer's problem can be stated and analyzed under different tax regimes. To highlight some basic equilibrium properties first look at

The problem without taxation.

By the choice of $S^i \in [0, 1]$ consumer i maximizes $U^i(c_1^i, c_2^i)$ such that the budget constraints

$$\begin{aligned} c_1^1 &= (1 - S^1)P_1 \\ c_2^1 &= S^1 P_2 \end{aligned} \tag{2.3}$$

and

$$\begin{aligned} c_1^2 &= W - S^2 P_1 \\ c_2^2 &= S^2 P_2 \end{aligned} \tag{2.4}$$

respectively hold. From the first order conditions

$$\frac{\partial U^i / \partial c_1^i}{\partial U^i / \partial c_2^i} = \frac{P_2}{P_1} \tag{2.5}$$

price dependent demand functions $S^i(P_1)$ for the asset can be derived, and the market clearing condition (2.2) delivers an equilibrium price P_1^*. Applying the implicit function theorem on the first order conditions yields the following comparative statics results:[12]

$$\frac{dS^1}{dP_1} \gtrless 0, \quad \frac{dS^2}{dP_1} < 0. \tag{2.6}$$

An increasing asset price P_1 makes consumption in period 1 relatively cheaper, i.e. has a negative substitution effect on consumption in period 2 and hence decreases savings S^i for both consumers. While the income effect is also negative for agent 2, it is positive for agent 1. Therefore the asset demand of consumer 2 is decreasing in the asset price, whereas the effect is not clear cut for consumer 1.

Under the assumptions made, there is – as mentioned above – always an equilibrium price but it is not necessarily unique. However, as illustrated by figure 2.1, in the case of multiple equilibria for almost every combination of values P_0, P_2 and W a locally isolated equilibrium price P_1^* exists such that aggregated asset demand is decreasing in P_1, i.e.[13]

$$\frac{d(S^1 + S^2)}{dP_1} < 0\,. \tag{2.7}$$

The following analysis refers always to this type of equilibrium.

The problem with taxation upon accrual

Under an accrual tax the consumer's budget constraints are given by

$$\begin{aligned} c_1^1 &= (1 - S^1)P_1 - \tau(P_1 - P_0) \\ c_2^1 &= S^1[P_2 - \tau(P_2 - P_1)] \end{aligned} \tag{2.8}$$

[12]The formal analysis is presented in Appendix 2.6.1.

[13]See e.g. Mas-Colell et al. (1995), Section 17.D.

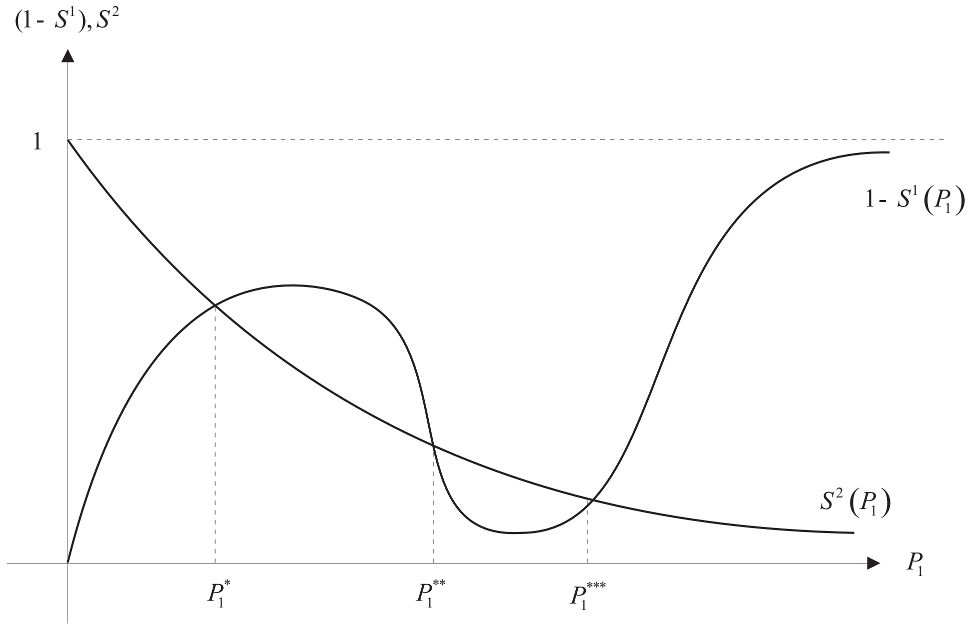

Figure 2.1: Multiple equilibria

and

$$\begin{aligned} c_1^2 &= W - S^2 P_1 \\ c_2^2 &= S^2[P_2 - \tau(P_2 - P_1)] \end{aligned} \quad (2.9)$$

respectively. Note that the taxable base of consumer 1 in period 1 is not only the realized part but his entire capital gain. The first order conditions for a maximum become

$$\frac{\partial U^i / \partial c_1^i}{\partial U^i / \partial c_2^i} = \frac{P_2 - \tau(P_2 - P_1)}{P_1} . \quad (2.10)$$

Now the consumers' asset demand depends not only on the price P_1 but also on the tax rate τ. Additionally, consumer 1's consumption and hence asset demand responds to changes in accrued capital gains, i.e. changes in the basis P_0. Consequently, in general one has $S^1 = S^1(P_1, \tau, P_0)$ and $S^2 = S^2(P_1, \tau)$ respectively and hence the equilibrium price $P_1^* = P_1^*(\tau, P_0)$ is also a function of tax rate and accrued capital gains.

With the same reasoning as before, the comparative statics results of expression (2.6) hold in the presence of an accrual tax.[14] Investigating changes

[14]The formal analysis is presented in Appendix 2.6.2.

in the tax rate one gets a somewhat surprising result:

$$\frac{dS^1}{d\tau} \gtreqless 0, \quad \frac{dS^2}{d\tau} \gtreqless 0 \quad \text{and thus} \quad \frac{dP_1^*}{d\tau} \gtreqless 0. \tag{2.11}$$

Because income and substitution effect work in opposite directions with respect to consumption in period 1, in the absence of additional assumptions, a change in the tax rate has no clear-cut effect on the consumers' saving decision and hence the price effect is ambiguous as well. This is often overlooked by the literature describing the depressing nature of the capitalization effect[15]: Even under an accrual system an increasing tax rate does not necessarily result in decreasing asset prices.

Moreover, in the existing literature, the amount of capital gains is often thought to be neutral under an accrual tax in the sense that it does not affect the saving and investment decision.[16] However, this is not true if taxes are collected only periodically like in the model presented here. To see this first note that consumer 2's problem and thus asset demand are not altered by a change in the accrued capital gain of consumer 1, i.e. by a change in P_0. In contrast, a lower accrued gain, i.e. higher P_0, has a pure income effect increasing consumer 1's consumption in both periods. Since his period 2 consumption does not depend directly on P_0 it can only be augmented by saving more, i.e.[17]

$$\frac{dS^1}{dP_0} \geq 0. \tag{2.12}$$

Applying the implicit function theorem on the market clearing condition (2.2), and employing inequalities (2.7) and (2.12) yield the following result:

$$\frac{dP_1^*}{dP_0} = -\frac{\frac{\partial S^1}{\partial P_0}}{\frac{\partial (S^1+S^2)}{\partial P_1}} \geq 0. \tag{2.13}$$

Put differently, as a reaction to higher accrued gains total asset demand and hence the equilibrium price may decrease. This means, in particular, that in the presence of capital gains taxes asset prices do not only depend on future (expected) payoffs but possibly also on past prices, even if the tax is levied upon accrual.

The problem with taxation upon realization

While consumer 2's budget constraints (2.9) and first order condition (2.10) do not alter under a realization system, consumer 1 now pays taxes in period

[15] See e.g. Lang and Shackelford (2000).
[16] See e.g. Auerbach (1991).
[17] The formal analysis is presented in Appendix 2.6.2.

1 only for the realized part of his capital gains and, hence, faces the following constraints

$$\begin{aligned} c_1^1 &= (1 - S^1)[P_1 - \tau(P_1 - P_0)] \\ c_2^1 &= S^1[P_2 - \tau(P_2 - P_0)] \end{aligned} \quad (2.14)$$

resulting in the first order condition

$$\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1} = \frac{P_2 - \tau(P_2 - P_0)}{P_1 - \tau(P_1 - P_0)} . \quad (2.15)$$

As a consequence, if there is heterogeneity among the agents with respect to their accrued capital gains, i.e. if $P_0 < P_1$, consumers do no longer face the same relative prices for consumption in period 1 and 2 respectively.[18] Put differently: Whereas under an accrual tax, the marginal rates of substitution for consumer 1 and 2 coincide, they differ under a realization tax. Hence, the resulting equilibrium allocation under taxation upon realization cannot be Pareto efficient. This result may be seen as a formal justification for the branch of literature surveyed in Chapter 1, that tries to find a way of circumventing the lock-in effect by simulating an accrual system on a realization basis. However, as Proposition 2.2 will show, the result does not imply that the equilibrium allocation resulting from an accrual tax Pareto dominates the equilibrium allocation under a realization tax.

Comparative statics show that relations (2.6) and (2.11) still hold under a realization tax.[19] This means in particular that, as in the model of Klein (1999), an increasing tax rate may possibly lead to higher asset prices.

However, the prediction of Klein (1999) that the pre-tax returns for assets with accrued capital gains are smaller, i.e. their prices are higher, than for assets without such accrued gains cannot be verified in this setting. Yet, in contrast to the case of an accrual system[20], under a realization tax higher accrued capital gains can possibly increase asset prices as the following comparative statics show:[21]

$$\frac{dS^1}{dP_0} \gtrless 0 \quad (2.16)$$

and hence

$$\frac{dP^1}{dP_0} = -\frac{\frac{\partial S^1}{\partial P_0}}{\frac{\partial (S^1+S^2)}{\partial P_1}} \gtrless 0 .$$

[18]For $P_0 = P_1$ the FOC's (2.10) and (2.15) coincide. For $P_0 < P_1$, by inequality (2.1) consumption in period 2 is relatively cheaper for consumer 1 than 2.

[19]The formal analysis is presented in Appendix 2.6.3.

[20]Remember inequality (2.13).

[21]The formal analysis is presented in Appendix 2.6.3.

The intuition behind this result is as follows: Now, a smaller accrued gain, i.e. higher P_0, does not only result in a positive income effect on consumption in both periods but also in a substitution effect such that consumption in period 2 becomes more expensive. Thus the overall effect on consumption in period 2 and hence on saving is ambiguous.

Comparative statics results

The results derived so far can be summarized in the following way: In general the optimal saving decision of an investor depends on his accrued capital gain, even under an accrual tax. Under an accrual tax P_1 is increasing in P_0, i.e. the higher the accrued capital gain the lower the asset price. Under a realization tax, in general the impact of a change in P_0 on P_1 is ambiguous, i.e. the effect possibly can but does not necessarily have to be reverted. Moreover, the influence of a change in the tax rate on asset prices is not clear cut either, not even under an accrual system. In contrast to that somehow unsatisfactory ambiguity with respect to the tax rate as an instrument of public policy, the next section will show, that comparing the impact of taxation methods on asset prices leads to clear cut results.

2.4 Price and welfare effects

In this section the impact the method of capital gains taxation has on asset prices and welfare is investigated. To this end, the equilibrium allocation under an accrual based taxation system is compared to the equilibrium outcome under a realization based one.

2.4.1 The impact on asset prices

To carry out a welfare analysis, first one has to learn how equilibrium prices are affected by a certain method of taxation.

Proposition 2.1 (Price effect) *Given a fixed tax rate $0 < \tau < 1$, in the presence of accrued capital gains ($P_0 < P_1$) the equilibrium asset price P_1^* is higher under a realization tax than under an accrual system.*

The result is intuitive: Compared to accrual taxation, the realization system creates an incentive for agent 1 to defer part of his accrued gains till period 2 in order to save taxes, while consumer 2 is not affected by the method of taxation. Thus total demand for the asset and hence its price are higher; more formally:

Proof. Consumer 2's problem and hence asset demand $S^2(P_1)$ are the same under both taxation methods.[22] As for any given $P_0 < P_1 \leq P_2$ inequality (2.1) ensures

$$\frac{P_1 - \tau(P_1 - P_0)}{P_1}\left[P_2 - \tau(P_2 - P_1)\right] \leq P_2 - \tau(P_2 - P_0),$$

consumer 1's problem under a realization tax differs from the one under an accrual tax in two ways: As one can see from the intertemporal budget constraints resulting from (2.8) and (2.14) respectively,

$$c_2^{1\,\mathrm{acc}} = \frac{P_2 - \tau(P_2 - P_1)}{P_1}\left[P_1 - \tau(P_1 - P_0)\right] - \frac{P_2 - \tau(P_2 - P_1)}{P_1}\,c_1^{1\,\mathrm{acc}} \quad \text{and}$$

$$c_2^{1\,\mathrm{real}} = P_2 - \tau(P_2 - P_0) - \frac{P_2 - \tau(P_2 - P_0)}{P_1 - \tau(P_1 - P_0)}\,c_1^{1\,\mathrm{real}},$$

first his budget set is larger. Secondly consumption in period 2 is relatively cheaper. Since consumption in period 2 is a normal commodity, both income and substitution effect are positive with respect to period 2 consumption of consumer 1, and hence $c_2^{1\,\mathrm{real}} > c_2^{1\,\mathrm{acc}}$. The situation is illustrated in figure 2.2. However, this is only possible by higher savings, because for any given S^1 and $P_0 < P_1 \leq P_2$ by the constraints (2.8) and (2.14)

$$c_2^{1\,\mathrm{acc}} = S^1[P_2 - \tau(P_2 - P_1)] > S^1[P_2 - \tau(P_2 - P_0)] = c_2^{1\,\mathrm{real}}$$

holds, i.e. ceteris paribus consumer 1's consumption in period 2 is higher under an accrual than a realization system. Therefore consumer 1's demand $S^1(P_1)$ and thus total demand for the asset are higher under a realization tax. By inequality (2.7), that results in a higher equilibrium price P_1^*.

□

2.4.2 The impact on welfare

In order to investigate *welfare effects* in a framework with taxes, generally one has to take into account how *tax revenues* are spent in the public sector. As stated by the following Lemma, in the present setting this problem fortunately can be neglected: Assuming that the government does not discount,[23]

[22] This observation, of course, is due to the special structure of the model, where consumer 2 holds the asset for exactly one period and, hence, the methods of taxation are equivalent.

[23] This could be modelled explicitly by assuming that, for example, the government had to finance a public good in period 2 but no access to the capital market.

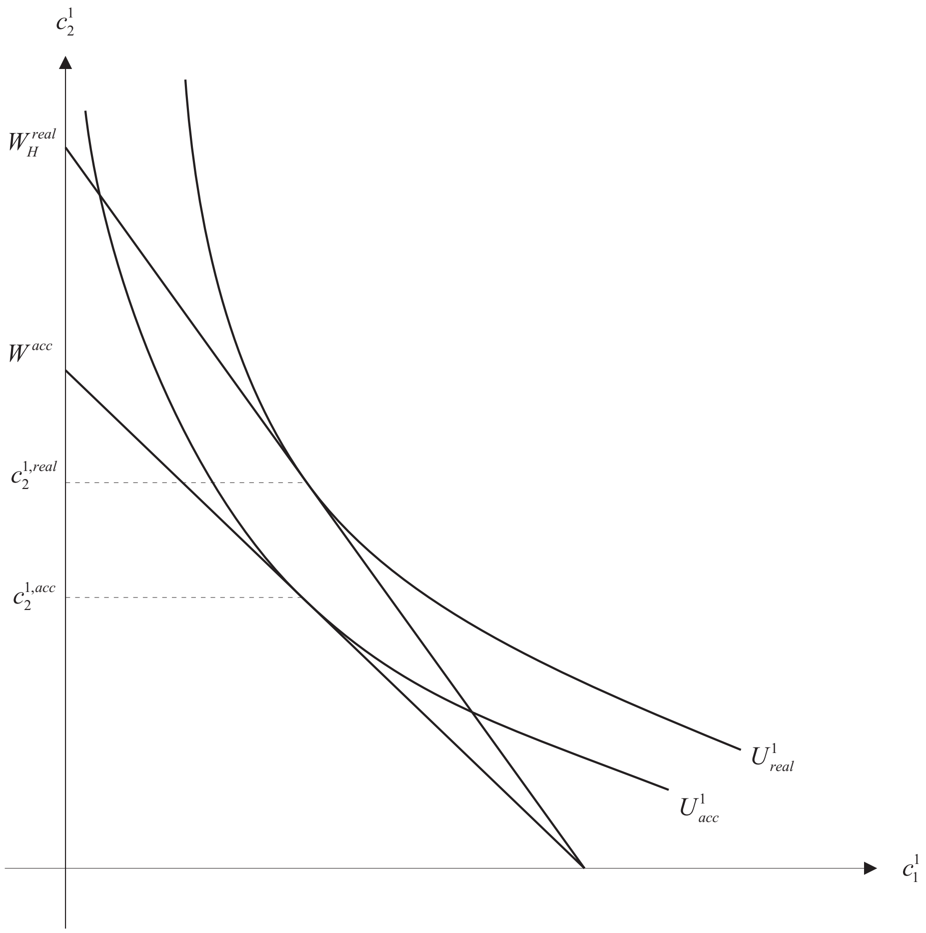

Figure 2.2: Consumer 1's problems

total tax revenue and hence expenditures are not affected by the method of taxation.

Lemma 2.1 *Given a fixed tax rate* $0 < \tau < 1$ *and no intertemporal discounting by the public authorities* ($\rho = 0$)*, the present value of total tax revenue is the same under an accrual based tax system and a realization based one respectively.*

Proof. Compare the present value of total tax revenue under an accrual system T_{acc} and a realization one T_{real} respectively:

$$T_{\text{acc}} = \tau(P_1 - P_0) + \frac{1}{1+\rho}\left[\tau S^1(P_2 - P_1) + \tau(1 - S^1)(P_2 - P_1)\right]$$

$$
\begin{aligned}
&\overset{\rho=0}{=} \quad \tau(P_2 - P_0)\,, \\
T_{\text{real}} &= \quad \tau(1-S^1)(P_1-P_0) + \frac{1}{1+\rho}\left[\tau S^1(P_2-P_0) + \tau(1-S^1)(P_2-P_1)\right] \\
&\overset{\rho=0}{=} \quad \tau(P_2 - P_0)\,.
\end{aligned}
$$

□

Note in particular that under neither system the revenue depends on the asset price P_1 in period 1. By means of Proposition 2.1 and Lemma 2.1 it is possible to prove the following statement concerning total welfare.

Proposition 2.2 (Welfare effect) *Given a fixed tax rate $0 < \tau < 1$ and no intertemporal discounting by the authorities ($\rho = 0$),*[24] *the equilibrium allocation under neither of both taxation methods Pareto dominates the other. In the presence of accrued capital gains ($P_0 < P_1$) consumer 1's utility is higher and consumer 2's utility is lower under a realization tax than under an accrual system.*

Proof. By Lemma 2.1, total tax revenue and thus expenditure is equal under each of both methods. Put differently, there is no effect on utilities caused by different public spending. As seen in the proof of Proposition 2.1 the method of taxation has no direct effect on the problem of consumer 2 whereas the budget set and hence utility of consumer 1 is larger under a realization tax than an accrual one. Additionally, by Proposition 2.1 the equilibrium price P_1 is higher under a realization tax, which has, compared to the situation under an accrual tax, two opposed effects: On the one hand, this further relaxes the budget constraint (2.14) of consumer 1 and thus increases his budget set and, hence, utility. On the other hand, this tightens the budget constraint (2.9) of consumer 2 and thus decreases his budget set and utility.

□

As has been pointed out in Section 2.3, the equilibrium allocation under a realization tax cannot be Pareto efficient, while the one under an accrual tax may be. However, Proposition 2.2 shows that the Pareto welfare criteria are not able to give a political device on the preferability of one taxation method or the other.

[24] Using the continuity of the relevant functions, one easily verifies that Proposition 2.2 will still hold for sufficiently small discount rates $\rho > 0$ if the tax rate is adopted to the corresponding method of taxation in a way such that total tax revenue remains unchanged.

In the light of the large literature on how to avoid the lock-in effect of capital gains taxation upon realization surveyed in Chapter 1, the above result is somewhat surprising. However, there are some authors, like Kovenock and Rothschild (1987), who doubt there is a strong negative welfare effect arising from a realization tax. The analysis presented here may be seen as a further justification for this point of view.

2.4.3 Distributional aspects

In view of Proposition 2.2, without further assumptions efficiency does not provide a fully convincing justification for an accrual tax on normative grounds. If the policy space is restricted to the choice of the taxation method, the Pareto welfare criterion is not able to answer the question whether a realization or accrual tax should be established. To decide on the method of taxation, a weaker concept of social welfare has to be employed, which involves the aggregation of individual utilities. Note that any such utility aggregation, for example by a social welfare function, implicitly incorporates an interpersonal comparison, i.e. a certain ideal of how utility should be *distributed* within the economy. The example provided in the Appendix 2.7 shows that the optimal method of taxation depends on the social welfare function employed. Hence, in the political process, the distributional norm is decisive for the method of taxation.

For example, if the norm is fairness, the above result may give a hint on how this decision might look like: As demonstrated in Proposition 2.2, compared to an accrual based tax system, a realization based one discriminates consumer 2 while it favors consumer 1, i.e. it is in favor of agents with relatively large accrued capital gains. It should be easy to find empirical evidence for the claim that such gains occur more often among 'wealthy' individuals than 'poor' as well as more often among 'elderly' people than 'young'. Given these presumptions, a realization based tax system in which the lock-in effect is present has to be refused if the political aim is to 'close the gap' and 'reduce the burden of future generations' respectively.

2.4.4 Concerns of optimal taxation

From the viewpoint of *optimal taxation* the analysis presented so far investigates the question, whether taxation upon accrual or realization is preferable for a given tax rate τ. In the light of lemma 2.1, this is equivalent to the assumption of an exogenous revenue requirement. As discussed above, the answer depends upon the welfare criterion used by the planner.

Alternatively, one may ask the following question: What is the optimal tax rate τ_M given a certain method of taxation M? And more specific: Do the optimal tax rates under accrual and realization taxation coincide or differ (systematically)?

The problem can be studied within a slight extension of the above framework, in which the planner uses the tax revenues to provide a public good at the end of period 2. Although the revenue for a given tax rate τ is identical under both taxation methods, as shown in Lemma 2.1, one easily verifies the following result, that is illustrated in the example provided in Appendix 2.7: In general, the optimal tax rates under accrual and realization taxation differ, but not systematically (e.g. always $\tau^*_{\text{acc}} \leq \tau^*_{\text{real}}$). This result is due to the fact that the method of taxation alters the distribution of utilities in the economy as stated in Proposition 2.2. According to his welfare criterion the planner may want to correct for this change by adjusting the amount of public good provided and hence may set different tax rates. Consequently, from the viewpoint of optimal taxation, a change in the method of taxation usually requires an adjustment of the corresponding tax rate. This point should not be overlooked in the ongoing discussions about reforms of capital gains taxation.

2.5 Extensions and concluding remarks

This chapter has investigated the effects of accrual and realization based taxation of capital gains within a simple general equilibrium model with heterogeneous agents. It has been shown that in the presence of accrued capital gains equilibrium asset prices are higher under a realization tax than under an accrual tax. However, due to distributional effects, the impact of the taxation method on welfare is ambiguous, though taxation upon realization cannot be Pareto efficient.

The analysis can be extended in various directions. So far, only the distortion of the liquidation decision is incorporated but not the distortion of the portfolio choice. To take this secondary lock-in effect into account one should look at a situation with alternative saving opportunities and where both taxation methods coexist. Within such a setting, instead of just comparing the two taxation methods, the homogenization of a hybrid system towards either a pure accrual or realization tax should be studied, something that is definitely closer to what can be observed in reality.

It would also be more realistic to allow for arbitrary initial endowments. Nevertheless, to investigate the interesting cases with different amounts of accrued capital gains among consumers, one has to introduce some source of heterogeneity between the agents. And the easiest way to do so is by assuming different initial endowments. The consideration of multiple consumers does not seem very promising either, as it would considerably complicate the analysis while presumably change the results only slightly if at all.

The relation between uncertainty and the taxation of capital gains is mostly studied under an accrual tax system and the results depend upon how tax revenues are spent in the corresponding model.[25] Kovenock and Rothschild (1987) explicitly assume a realization tax and show that due to the lock-in effect the effective tax rate is higher in the presence of risk. It would be interesting to carry forward their analysis within the framework presented here.

One of the most promising extensions and a step towards reality certainly would be to look at a model with endogenous production, i.e. endogenous asset supply.[26] As long as production is exogenously given, the lock-in effect is at most able to distort the decisions on liquidation and portfolio selection but not on real investment. Hence the welfare analysis presented in this paper can at best provide some hints, how a more complete picture might look like and has to be subjected to further research.

2.6 Appendix: Comparative statics

The subsequent analysis makes intensive use of the fact that, due to the normality assumption, for an interior solution to consumer i's problem of utility maximization the following inequalities hold:

$$\begin{aligned} \frac{\partial^2 U^i}{\partial {c_1^i}^2} - \left(\frac{\partial U^i/\partial c_1^i}{\partial U^i/\partial c_2^i}\right)\frac{\partial^2 U^i}{\partial c_2^i \partial c_1^i} &< 0, \\ \frac{\partial^2 U^i}{\partial c_1^i \partial c_2^i} - \left(\frac{\partial U^i/\partial c_1^i}{\partial U^i/\partial c_2^i}\right)\frac{\partial^2 U^i}{\partial {c_2^i}^2} &> 0. \end{aligned} \tag{2.17}$$

This can be seen by applying the implicit function theorem on consumer i's first order condition for a maximum and using the property, that his second order condition for a maximum is fulfilled at an interior solution, as well as taking into account, that consumption is assumed to be a normal commodity in both periods.

[25] See Sandmo (1985) for a survey of this topic and, for example, Konrad (1991) or Christiansen (1995) for some later general equilibrium results.

[26] For the framework presented here, this would require to explicitly model the decision on how much to invest in period 0, i.e. to 'endogenize' the 'price' P_0.

2.6.1 The case of no taxation

Applying the implicit function theorem on the first order conditions (2.5) yields for $i \in \{1,2\}$

$$\frac{dS^i}{dP_1} = -\frac{\frac{P_2}{{P_1}^2} + \frac{\partial\left(\frac{\partial U^i/\partial c_1^i}{\partial U^i/\partial c_2^i}\right)}{\partial P_1}}{\frac{\partial\left(\frac{\partial U^i/\partial c_1^i}{\partial U^1/\partial c_2^i}\right)}{\partial S^i}}\,. \tag{2.18}$$

Using the quotient rule and inequalities (2.17) as well as taking the budget constraints (2.3) and (2.4) respectively into account, one can determine the sign of the following terms

$$\begin{aligned}
\frac{\partial\left(\frac{\partial U^i/\partial c_1^i}{\partial U^i/\partial c_2^i}\right)}{\partial S^i} &= \frac{\left(\frac{\partial^2 U^i}{\partial {c_1^i}^2}\frac{dc_1^i}{dS^i} + \frac{\partial^2 U^i}{\partial c_1^i \partial c_2^i}\frac{dc_2^i}{dS^i}\right)\frac{\partial U^i}{\partial c_2^i} - \frac{\partial U^i}{\partial c_1^i}\left(\frac{\partial^2 U^i}{\partial c_2^i \partial c_1^i}\frac{dc_1^i}{dS^i} + \frac{\partial^2 U^i}{\partial {c_2^i}^2}\frac{dc_2^i}{dS^i}\right)}{(\partial U^i/\partial c_2^i)^2} \\
&= \frac{\frac{dc_1^i}{dS^i}\left[\frac{\partial^2 U^i}{\partial {c_1^i}^2} - \left(\frac{\partial U^i/\partial c_1^i}{\partial U^i/\partial c_2^i}\right)\frac{\partial^2 U^i}{\partial c_2^i \partial c_1^i}\right] + \frac{dc_2^i}{dS^i}\left[\frac{\partial^2 U^i}{\partial c_1^i \partial c_2^i} - \left(\frac{\partial U^i/\partial c_1^i}{\partial U^i/\partial c_2^i}\right)\frac{\partial^2 U^i}{\partial {c_2^i}^2}\right]}{\partial U^i/\partial c_2^i} \\
&> 0\,,
\end{aligned}$$

$$\begin{aligned}
\frac{\partial\left(\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1}\right)}{\partial P_1} &= \frac{\left(\frac{\partial^2 U^1}{\partial {c_1^1}^2}\frac{dc_1^1}{dP_1} + \frac{\partial^2 U^1}{\partial c_1^1 \partial c_2^1}\frac{dc_2^1}{dP_1}\right)\frac{\partial U^1}{\partial c_2^1} - \frac{\partial U^1}{\partial c_1^1}\left(\frac{\partial^2 U^1}{\partial c_2^1 \partial c_1^1}\frac{dc_1^1}{dP_1} + \frac{\partial^2 U^1}{\partial {c_2^1}^2}\frac{dc_2^1}{dP_1}\right)}{(\partial U^1/\partial c_2^1)^2} \\
&= \frac{\frac{dc_1^1}{dP_1}\left[\frac{\partial^2 U^1}{\partial {c_1^1}^2} - \left(\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1}\right)\frac{\partial^2 U^1}{\partial c_2^1 \partial c_1^1}\right] + \frac{dc_2^1}{dP_1}\left[\frac{\partial^2 U^1}{\partial c_1^1 \partial c_2^1} - \left(\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1}\right)\frac{\partial^2 U^1}{\partial {c_2^1}^2}\right]}{\partial U^1/\partial c_2^1} \\
&< 0\,,
\end{aligned}$$

$$\begin{aligned}
\frac{\partial\left(\frac{\partial U^2/\partial c_1^2}{\partial U^2/\partial c_2^2}\right)}{\partial P_1} &= \frac{\left(\frac{\partial^2 U^2}{\partial {c_1^2}^2}\frac{dc_1^2}{dP_1} + \frac{\partial^2 U^2}{\partial c_1^2 \partial c_2^2}\frac{dc_2^2}{dP_1}\right)\frac{\partial U^2}{\partial c_2^2} - \frac{\partial U^2}{\partial c_1^2}\left(\frac{\partial^2 U^2}{\partial c_2^2 \partial c_1^2}\frac{dc_1^2}{dP_1} + \frac{\partial^2 U^2}{\partial {c_2^2}^2}\frac{dc_2^2}{dP_1}\right)}{(\partial U^2/\partial c_2^2)^2} \\
&= \frac{\frac{dc_1^2}{dP_1}\left[\frac{\partial^2 U^2}{\partial {c_1^2}^2} - \left(\frac{\partial U^2/\partial c_1^2}{\partial U^2/\partial c_2^2}\right)\frac{\partial^2 U^2}{\partial c_2^2 \partial c_1^2}\right] + \frac{dc_2^2}{dP_1}\left[\frac{\partial^2 U^2}{\partial c_1^2 \partial c_2^2} - \left(\frac{\partial U^2/\partial c_1^2}{\partial U^2/\partial c_2^2}\right)\frac{\partial^2 U^2}{\partial {c_2^2}^2}\right]}{\partial U^2/\partial c_2^2} \\
&> 0\,,
\end{aligned}$$

and hence the sign of the expressions in (2.18) as stated in (2.6).

2.6.2 The case of accrual taxation

Applying the implicit function theorem on the first order conditions (2.10) yields for $i \in \{1,2\}$

$$\frac{dS^i}{dP_1} = -\frac{\frac{P_2(1-\tau)}{{P_1}^2} + \frac{\partial\left(\frac{\partial U^i/\partial c_1^i}{\partial U^i/\partial c_2^i}\right)}{\partial P_1}}{\frac{\partial\left(\frac{\partial U^i/\partial c_1^i}{\partial U^1/\partial c_2^i}\right)}{\partial S^i}}, \tag{2.19}$$

$$\frac{dS^i}{d\tau} = -\frac{\frac{P_2-P_1}{P_1} + \frac{\partial\left(\frac{\partial U^i/\partial c_1^i}{\partial U^i/\partial c_2^i}\right)}{\partial \tau}}{\frac{\partial\left(\frac{\partial U^i/\partial c_1^i}{\partial U^1/\partial c_2^i}\right)}{\partial S^i}}, \tag{2.20}$$

$$\frac{dS^1}{dP_0} = -\frac{\frac{\partial\left(\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1}\right)}{\partial P_0}}{\frac{\partial\left(\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1}\right)}{\partial S^1}}. \tag{2.21}$$

Using the budget constraints (2.8) and (2.9) respectively, one can determine the sign of the following terms

$$\begin{aligned}
\frac{\partial\left(\frac{\partial U^i/\partial c_1^i}{\partial U^i/\partial c_2^i}\right)}{\partial S^i} &= \frac{\left(\frac{\partial^2 U^i}{\partial {c_1^i}^2}\frac{dc_1^i}{dS^i} + \frac{\partial^2 U^i}{\partial c_1^i \partial c_2^i}\frac{dc_2^i}{dS^i}\right)\frac{\partial U^i}{\partial c_2^i} - \frac{\partial U^i}{\partial c_1^i}\left(\frac{\partial^2 U^i}{\partial c_2^i \partial c_1^i}\frac{dc_1^i}{dS^i} + \frac{\partial^2 U^i}{\partial {c_2^i}^2}\frac{dc_2^i}{dS^i}\right)}{(\partial U^i/\partial c_2^i)^2} \\
&= \frac{\frac{dc_1^i}{dS^i}\left[\frac{\partial^2 U^i}{\partial {c_1^i}^2} - \left(\frac{\partial U^i/\partial c_1^i}{\partial U^i/\partial c_2^i}\right)\frac{\partial^2 U^i}{\partial c_2^i \partial c_1^i}\right] + \frac{dc_2^i}{dS^i}\left[\frac{\partial^2 U^i}{\partial c_1^i \partial c_2^i} - \left(\frac{\partial U^i/\partial c_1^i}{\partial U^i/\partial c_2^i}\right)\frac{\partial^2 U^i}{\partial {c_2^i}^2}\right]}{\partial U^i/\partial c_2^i} \\
&> 0,
\end{aligned}$$

$$\begin{aligned}
\frac{\partial\left(\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1}\right)}{\partial P_1} &= \frac{\left(\frac{\partial^2 U^1}{\partial {c_1^1}^2}\frac{dc_1^1}{dP_1} + \frac{\partial^2 U^1}{\partial c_1^1 \partial c_2^1}\frac{dc_2^1}{dP_1}\right)\frac{\partial U^1}{\partial c_2^1} - \frac{\partial U^1}{\partial c_1^1}\left(\frac{\partial^2 U^1}{\partial c_2^1 \partial c_1^1}\frac{dc_1^1}{dP_1} + \frac{\partial^2 U^1}{\partial {c_2^1}^2}\frac{dc_2^1}{dP_1}\right)}{(\partial U^1/\partial c_2^1)^2} \\
&= \frac{\frac{dc_1^1}{dP_1}\left[\frac{\partial^2 U^1}{\partial {c_1^1}^2} - \left(\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1}\right)\frac{\partial^2 U^1}{\partial c_2^1 \partial c_1^1}\right] + \frac{dc_2^1}{dP_1}\left[\frac{\partial^2 U^1}{\partial c_1^1 \partial c_2^1} - \left(\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1}\right)\frac{\partial^2 U^1}{\partial {c_2^1}^2}\right]}{\partial U^1/\partial c_2^1} \\
&\gtreqless 0,
\end{aligned}$$

$$\frac{\partial\left(\frac{\partial U^2/\partial c_1^2}{\partial U^2/\partial c_2^2}\right)}{\partial P_1} = \frac{\left(\frac{\partial^2 U^2}{\partial {c_1^2}^2}\frac{dc_1^2}{dP_1} + \frac{\partial^2 U^2}{\partial c_1^2 \partial c_2^2}\frac{dc_2^2}{dP_1}\right)\frac{\partial U^2}{\partial c_2^2} - \frac{\partial U^2}{\partial c_1^2}\left(\frac{\partial^2 U^2}{\partial c_2^2 \partial c_1^2}\frac{dc_1^2}{dP_1} + \frac{\partial^2 U^2}{\partial {c_2^2}^2}\frac{dc_2^2}{dP_1}\right)}{(\partial U^2/\partial c_2^2)^2}$$

$$
\begin{aligned}
&= \frac{\frac{dc_1^2}{dP_1}\left[\frac{\partial^2 U^2}{\partial {c_1^2}^2} - \left(\frac{\partial U^2/\partial c_1^2}{\partial U^2/\partial c_2^2}\right)\frac{\partial^2 U^2}{\partial c_2^2 \partial c_1^2}\right] + \frac{dc_2^2}{dP_1}\left[\frac{\partial^2 U^2}{\partial c_1^2 \partial c_2^2} - \left(\frac{\partial U^2/\partial c_1^2}{\partial U^2/\partial c_2^2}\right)\frac{\partial^2 U^2}{\partial {c_2^2}^2}\right]}{\partial U^2/\partial c_2^2} \\
&> 0\,,
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial\left(\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1}\right)}{\partial \tau} &= \frac{\left(\frac{\partial^2 U^1}{\partial {c_1^1}^2}\frac{dc_1^1}{d\tau} + \frac{\partial^2 U^1}{\partial c_1^1 \partial c_2^1}\frac{dc_2^1}{d\tau}\right)\frac{\partial U^1}{\partial c_2^1} - \frac{\partial U^1}{\partial c_1^1}\left(\frac{\partial^2 U^1}{\partial c_2^1 \partial c_1^1}\frac{dc_1^1}{d\tau} + \frac{\partial^2 U^1}{\partial {c_2^1}^2}\frac{dc_2^1}{d\tau}\right)}{(\partial U^1/\partial c_2^1)^2} \\
&= \frac{\frac{dc_1^1}{d\tau}\left[\frac{\partial^2 U^1}{\partial {c_1^1}^2} - \left(\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1}\right)\frac{\partial^2 U^1}{\partial c_2^1 \partial c_1^1}\right] + \frac{dc_2^1}{d\tau}\left[\frac{\partial^2 U^1}{\partial c_1^1 \partial c_2^1} - \left(\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1}\right)\frac{\partial^2 U^1}{\partial {c_2^1}^2}\right]}{\partial U^1/\partial c_2^1} \\
&\gtreqless 0\,,
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial\left(\frac{\partial U^2/\partial c_1^2}{\partial U^2/\partial c_2^2}\right)}{\partial \tau} &= \frac{\left(\frac{\partial^2 U^2}{\partial {c_1^2}^2}\frac{dc_1^2}{d\tau} + \frac{\partial^2 U^2}{\partial c_1^2 \partial c_2^2}\frac{dc_2^2}{d\tau}\right)\frac{\partial U^2}{\partial c_2^2} - \frac{\partial U^2}{\partial c_1^2}\left(\frac{\partial^2 U^2}{\partial c_2^2 \partial c_1^2}\frac{dc_1^2}{d\tau} + \frac{\partial^2 U^2}{\partial {c_2^2}^2}\frac{dc_2^2}{d\tau}\right)}{(\partial U^2/\partial c_2^2)^2} \\
&= \frac{\frac{dc_1^2}{d\tau}\left[\frac{\partial^2 U^2}{\partial {c_1^2}^2} - \left(\frac{\partial U^2/\partial c_1^2}{\partial U^2/\partial c_2^2}\right)\frac{\partial^2 U^2}{\partial c_2^2 \partial c_1^2}\right] + \frac{dc_2^2}{d\tau}\left[\frac{\partial^2 U^2}{\partial c_1^2 \partial c_2^2} - \left(\frac{\partial U^2/\partial c_1^2}{\partial U^2/\partial c_2^2}\right)\frac{\partial^2 U^2}{\partial {c_2^2}^2}\right]}{\partial U^2/\partial c_2^2} \\
&< 0\,,
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial\left(\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1}\right)}{\partial P_0} &= \frac{\left(\frac{\partial^2 U^1}{\partial {c_1^1}^2}\frac{dc_1^1}{dP_0} + \frac{\partial^2 U^1}{\partial c_1^1 \partial c_2^1}\frac{dc_2^1}{dP_0}\right)\frac{\partial U^1}{\partial c_2^1} - \frac{\partial U^1}{\partial c_1^1}\left(\frac{\partial^2 U^1}{\partial c_2^1 \partial c_1^1}\frac{dc_1^1}{dP_0} + \frac{\partial^2 U^1}{\partial {c_2^1}^2}\frac{dc_2^1}{dP_0}\right)}{(\partial U^1/\partial c_2^1)^2} \\
&= \frac{\frac{dc_1^1}{dP_0}\left[\frac{\partial^2 U^1}{\partial {c_1^1}^2} - \left(\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1}\right)\frac{\partial^2 U^1}{\partial c_2^1 \partial c_1^1}\right] + \frac{dc_2^1}{dP_0}\left[\frac{\partial^2 U^2}{\partial c_1^2 \partial c_2^2} - \left(\frac{\partial U^2/\partial c_1^2}{\partial U^2/\partial c_2^2}\right)\frac{\partial^2 U^2}{\partial {c_2^2}^2}\right]}{\partial U^1/\partial c_2^1} \\
&< 0\,,
\end{aligned}
$$

and hence the sign of the expressions in (2.19), (2.20), and (2.21) as stated in (2.6), (2.11), and (2.12) respectively.

2.6.3 The case of realization taxation

Applying the implicit function theorem on the first order condition (2.15) for consumer 1 and (2.10) for consumer 2 respectively yields

$$
\frac{dS^1}{dP_1} = -\frac{(1-\tau)\frac{P_2-\tau(P_2-P_0)}{(P_1-\tau(P_1-P_0))^2} + \frac{\partial\left(\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1}\right)}{\partial P_1}}{\frac{\partial\left(\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1}\right)}{\partial S^1}}\,, \tag{2.22}
$$

$$\frac{dS^2}{dP_1} = -\frac{\frac{P_2(1-\tau)}{{P_1}^2} + \frac{\partial\left(\frac{\partial U^2/\partial c_1^2}{\partial U^2/\partial c_2^2}\right)}{\partial P_1}}{\frac{\partial\left(\frac{\partial U^2/\partial c_1^2}{\partial U^1/\partial c_2^2}\right)}{\partial S^2}}, \tag{2.23}$$

$$\frac{dS^1}{d\tau} = -\frac{\frac{(P_2-P_0)[P_1-\tau(P_1-P_0)]-(P_1-P_0)[P_2-\tau(P_2-P_0)]}{[P_1-\tau(P_1-P_0)]^2} + \frac{\partial\left(\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1}\right)}{\partial\tau}}{\frac{\partial\left(\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1}\right)}{\partial S^1}}, \tag{2.24}$$

$$\frac{dS^2}{d\tau} = -\frac{\frac{P_2-P_1}{P_1} + \frac{\partial\left(\frac{\partial U^2/\partial c_1^2}{\partial U^2/\partial c_2^2}\right)}{\partial\tau}}{\frac{\partial\left(\frac{\partial U^2/\partial c_1^2}{\partial U^1/\partial c_2^2}\right)}{\partial S^2}}, \tag{2.25}$$

$$\frac{dS^1}{dP_0} = -\frac{\frac{\tau(1-\tau)(P_2-P_1)}{[(1-\tau)P_1+\tau P_0]^2} + \frac{\partial\left(\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1}\right)}{\partial P_0}}{\frac{\partial\left(\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1}\right)}{\partial S^1}}. \tag{2.26}$$

Using the budget constraints (2.14) and (2.9) respectively, one can determine the sign of the following terms

$$\begin{aligned}
\frac{\partial\left(\frac{\partial U^i/\partial c_1^i}{\partial U^i/\partial c_2^i}\right)}{\partial S^i} &= \frac{\left(\frac{\partial^2 U^i}{\partial {c_1^i}^2}\frac{dc_1^i}{dS^i} + \frac{\partial^2 U^i}{\partial c_1^i \partial c_2^i}\frac{dc_2^i}{dS^i}\right)\frac{\partial U^i}{\partial c_2^i} - \frac{\partial U^i}{\partial c_1^i}\left(\frac{\partial^2 U^i}{\partial c_2^i \partial c_1^i}\frac{dc_1^i}{dS^i} + \frac{\partial^2 U^i}{\partial {c_2^i}^2}\frac{dc_2^i}{dS^i}\right)}{(\partial U^i/\partial c_2^i)^2} \\
&= \frac{\frac{dc_1^i}{dS^i}\left[\frac{\partial^2 U^i}{\partial {c_1^i}^2} - \left(\frac{\partial U^i/\partial c_1^i}{\partial U^i/\partial c_2^i}\right)\frac{\partial^2 U^i}{\partial c_2^i \partial c_1^i}\right] + \frac{dc_2^i}{dS^i}\left[\frac{\partial^2 U^i}{\partial c_1^i \partial c_2^i} - \left(\frac{\partial U^i/\partial c_1^i}{\partial U^i/\partial c_2^i}\right)\frac{\partial^2 U^i}{\partial {c_2^i}^2}\right]}{\partial U^i/\partial c_2^i} \\
&> 0,
\end{aligned}$$

$$\begin{aligned}
\frac{\partial\left(\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1}\right)}{\partial P_1} &= \frac{\left(\frac{\partial^2 U^1}{\partial {c_1^1}^2}\frac{dc_1^1}{dP_1} + \frac{\partial^2 U^1}{\partial c_1^1 \partial c_2^1}\frac{dc_2^1}{dP_1}\right)\frac{\partial U^1}{\partial c_2^1} - \frac{\partial U^1}{\partial c_1^1}\left(\frac{\partial^2 U^1}{\partial c_2^1 \partial c_1^1}\frac{dc_1^1}{dP_1} + \frac{\partial^2 U^1}{\partial {c_2^1}^2}\frac{dc_2^1}{dP_1}\right)}{(\partial U^1/\partial c_2^1)^2} \\
&= \frac{\frac{dc_1^1}{dP_1}\left[\frac{\partial^2 U^1}{\partial {c_1^1}^2} - \left(\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1}\right)\frac{\partial^2 U^1}{\partial c_2^1 \partial c_1^1}\right] + \frac{dc_2^1}{dP_1}\left[\frac{\partial^2 U^1}{\partial c_1^1 \partial c_2^1} - \left(\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1}\right)\frac{\partial^2 U^1}{\partial {c_2^1}^2}\right]}{\partial U^1/\partial c_2^1} \\
&< 0,
\end{aligned}$$

$$\frac{\partial\left(\frac{\partial U^2/\partial c_1^2}{\partial U^2/\partial c_2^2}\right)}{\partial P_1} = \frac{\left(\frac{\partial^2 U^2}{\partial {c_1^2}^2}\frac{dc_1^2}{dP_1} + \frac{\partial^2 U^2}{\partial c_1^2 \partial c_2^2}\frac{dc_2^2}{dP_1}\right)\frac{\partial U^2}{\partial c_2^2} - \frac{\partial U^2}{\partial c_1^2}\left(\frac{\partial^2 U^2}{\partial c_2^2 \partial c_1^2}\frac{dc_1^2}{dP_1} + \frac{\partial^2 U^2}{\partial {c_2^2}^2}\frac{dc_2^2}{dP_1}\right)}{(\partial U^2/\partial c_2^2)^2}$$

$$
\begin{aligned}
&= \frac{\frac{dc_1^2}{dP_1}\left[\frac{\partial^2 U^2}{\partial {c_1^2}^2} - \left(\frac{\partial U^2/\partial c_1^2}{\partial U^2/\partial c_2^2}\right)\frac{\partial^2 U^2}{\partial c_2^2 \partial c_1^2}\right] + \frac{dc_2^2}{dP_1}\left[\frac{\partial^2 U^2}{\partial c_1^2 \partial c_2^2} - \left(\frac{\partial U^2/\partial c_1^2}{\partial U^2/\partial c_2^2}\right)\frac{\partial^2 U^2}{\partial {c_2^2}^2}\right]}{\partial U^2/\partial c_2^2} \\
&> 0\,,
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial\left(\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1}\right)}{\partial \tau} &= \frac{\left(\frac{\partial^2 U^1}{\partial {c_1^1}^2}\frac{dc_1^1}{d\tau} + \frac{\partial^2 U^1}{\partial c_1^1 \partial c_2^1}\frac{dc_2^1}{d\tau}\right)\frac{\partial U^1}{\partial c_2^1} - \frac{\partial U^1}{\partial c_1^1}\left(\frac{\partial^2 U^1}{\partial c_2^1 \partial c_1^1}\frac{dc_1^1}{d\tau} + \frac{\partial^2 U^1}{\partial {c_2^1}^2}\frac{dc_2^1}{d\tau}\right)}{(\partial U^1/\partial c_2^1)^2} \\
&= \frac{\frac{dc_1^1}{d\tau}\left[\frac{\partial^2 U^1}{\partial {c_1^1}^2} - \left(\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1}\right)\frac{\partial^2 U^1}{\partial c_2^1 \partial c_1^1}\right] + \frac{dc_2^1}{d\tau}\left[\frac{\partial^2 U^1}{\partial c_1^1 \partial c_2^1} - \left(\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1}\right)\frac{\partial^2 U^1}{\partial {c_2^1}^2}\right]}{\partial U^1/\partial c_2^1} \\
&\gtreqless 0\,,
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial\left(\frac{\partial U^2/\partial c_1^2}{\partial U^2/\partial c_2^2}\right)}{\partial \tau} &= \frac{\left(\frac{\partial^2 U^2}{\partial {c_1^2}^2}\frac{dc_1^2}{d\tau} + \frac{\partial^2 U^2}{\partial c_1^2 \partial c_2^2}\frac{dc_2^2}{d\tau}\right)\frac{\partial U^2}{\partial c_2^2} - \frac{\partial U^2}{\partial c_1^2}\left(\frac{\partial^2 U^2}{\partial c_2^2 \partial c_1^2}\frac{dc_1^2}{d\tau} + \frac{\partial^2 U^2}{\partial {c_2^2}^2}\frac{dc_2^2}{d\tau}\right)}{(\partial U^2/\partial c_2^2)^2} \\
&= \frac{\frac{dc_1^2}{d\tau}\left[\frac{\partial^2 U^2}{\partial {c_1^2}^2} - \left(\frac{\partial U^2/\partial c_1^2}{\partial U^2/\partial c_2^2}\right)\frac{\partial^2 U^2}{\partial c_2^2 \partial c_1^2}\right] + \frac{dc_2^2}{d\tau}\left[\frac{\partial^2 U^2}{\partial c_1^2 \partial c_2^2} - \left(\frac{\partial U^2/\partial c_1^2}{\partial U^2/\partial c_2^2}\right)\frac{\partial^2 U^2}{\partial {c_2^2}^2}\right]}{\partial U^2/\partial c_2^2} \\
&< 0\,,
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial\left(\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1}\right)}{\partial P_0} &= \frac{\left(\frac{\partial^2 U^1}{\partial {c_1^1}^2}\frac{dc_1^1}{dP_0} + \frac{\partial^2 U^1}{\partial c_1^1 \partial c_2^1}\frac{dc_2^1}{dP_0}\right)\frac{\partial U^1}{\partial c_2^1} - \frac{\partial U^1}{\partial c_1^1}\left(\frac{\partial^2 U^1}{\partial c_2^1 \partial c_1^1}\frac{dc_1^1}{dP_0} + \frac{\partial^2 U^1}{\partial {c_2^1}^2}\frac{dc_2^1}{dP_0}\right)}{(\partial U^1/\partial c_2^1)^2} \\
&= \frac{\frac{dc_1^1}{dP_0}\left[\frac{\partial^2 U^1}{\partial {c_1^1}^2} - \left(\frac{\partial U^1/\partial c_1^1}{\partial U^1/\partial c_2^1}\right)\frac{\partial^2 U^1}{\partial c_2^1 \partial c_1^1}\right] + \frac{dc_2^1}{dP_0}\left[\frac{\partial^2 U^2}{\partial c_1^2 \partial c_2^2} - \left(\frac{\partial U^2/\partial c_1^2}{\partial U^2/\partial c_2^2}\right)\frac{\partial^2 U^2}{\partial {c_2^2}^2}\right]}{\partial U^1/\partial c_2^1} \\
&\gtreqless 0\,,
\end{aligned}
$$

and hence the sign of the expressions in (2.22), (2.23), (2.24), (2.25), and (2.26) as stated in (2.6), (2.11), and (2.16) respectively.

2.7 Appendix: An example with quasi-linear logarithmic preferences

The considerations concerning optimal taxation briefly discussed in Section 2.4.4 are illustrated by an example using quasi-linear logarithmic preferences. To this end, consider a slight extension of the model presented above: A social planner spends the tax revenues to provide a certain amount g of a public good at the end of period 2, which enters the utility function of the consumers:

$$U^i(c_1^i, c_2^i, g) = \ln(c_1^i) + \ln(c_2^i) + \theta g.$$

The taste parameter θ expresses the valuation of the public good. By Lemma 2.1, $g = \tau(P_2 - P_0)$ under both methods of taxation.

2.7.1 Individual utility maximization

The consumers take the tax rate τ and hence the amount g as given and maximize their utilities $U^i(c_1^i, c_2^i, g)$ deciding on their savings S^i in period 1. The relevant budget constraints for consumer 1 are described by (2.8) under accrual taxation and by (2.14) under realization taxation, whereas for consumer 2 they look the same under both methods of taxation and are given by (2.9). From the first order conditions for the consumers' utility maxima one can compute their optimal savings for a given asset price in period 1, $S^i(P_1)$. Under an accrual tax, consumer 1's asset demand equals

$$S^1_{\text{acc}} = \frac{1}{2}\frac{P_1 - \tau(P_1 - P_0)}{P_1},$$

under a realization tax it is given by

$$S^1_{\text{real}} = \frac{1}{2}.$$

Consumer 2's asset demand is the same under both methods of taxation:

$$S^2 = \frac{1}{2}\frac{W}{P_1}.$$

For assumption 2.2 to hold, it is sufficient to assume that the parameters of the model fulfil the following inequality:

$$P_0 < W < P_2.$$

Making use of the market clearing condition (2.2), the equilibrium asset price is then computed to be, in fact, higher under realization than accrual taxation, as stated in Proposition 2.1:

$$P^*_{1\,\text{acc}} = \frac{W + \tau P_0}{1 + \tau} \quad < \quad W = P^*_{1\,\text{real}}.$$

Substituting these prices into the corresponding asset demand functions and budget constraints one derives the equilibrium values for savings and consumption. For consumer 1, this yields

$$\begin{aligned}
{S^{1}}^{*}_{\text{acc}} &= \tfrac{1}{2}\tfrac{W+\tau(2P_0-W)}{W+\tau P_0} \\
&\leq \tfrac{1}{2} &&= {S^{1}}^{*}_{\text{real}}, \\
c^{1*}_{1\text{acc}} &= \tfrac{1}{2}W - \tfrac{\tau}{1+\tau}(W-P_0) \\
&\leq \tfrac{1}{2}W - \tfrac{\tau}{2}(W-P_0) &&= c^{1*}_{1\text{real}}, \\
c^{1*}_{2\text{acc}} &= \tfrac{1}{2}(1-\tau)P_2\tfrac{W+\tau P_0-\tau(W-P_0)}{W+\tau P_0} + \tfrac{1}{2}\tau\tfrac{W+\tau(2P_0-W)}{1+\tau}, \\
& \quad \tfrac{1}{2}(1-\tau)P_2 + \tfrac{1}{2}\tau P_0 &&= c^{1*}_{2\text{real}},
\end{aligned}$$

and for consumer 2,

$$\begin{aligned}
{S^{2}}^{*}_{\text{acc}} &= \tfrac{1}{2}\tfrac{W+\tau W}{W+\tau P_0} \\
&\geq \tfrac{1}{2} &&= {S^{2}}^{*}_{\text{real}}, \\
c^{2*}_{1\text{acc}} &= \tfrac{1}{2}W &&= c^{2*}_{1\text{real}}, \\
c^{2*}_{2\text{acc}} &= \tfrac{1}{2}(1-\tau)P_2\tfrac{W+\tau W}{W+\tau P_0} + \tfrac{1}{2}\tau W \\
&\geq \tfrac{1}{2}(1-\tau)P_2 + \tfrac{1}{2}\tau W &&= c^{2*}_{2\text{real}}.
\end{aligned}$$

Using those results one observes that for the consumers' equilibrium utility levels

$$\begin{aligned}
{U^{2}}^{*}_{\text{acc}} &\geq {U^{2}}^{*}_{\text{real}}, \\
{U^{1}}^{*}_{\text{acc}} &\leq {U^{1}}^{*}_{\text{real}},
\end{aligned} \tag{2.27}$$

hold as stated in Proposition 2.2. For consumer 2 this is obvious from his respective equilibrium consumption levels, for consumer 1 one can verify it by comparing the products $c^1_{1\,\text{acc}}c^1_{2\,\text{acc}}$ and $c^1_{1\,\text{real}}c^1_{2\,\text{real}}$ respectively; note that $U^i(c^i_1, c^i_2, g) = \ln(c^i_1 c^i_2) + \theta g$.

Moreover, for $t \in \{1, 2\}$ and $M \in \{\text{acc}, \text{real}\}$ the consumers' equilibrium consumption levels satisfy $c^2_{t\,M} \geq c^1_{t\,M}$ in this example. Hence, in equilibrium the utility level of consumer 2 is at least as high as that of consumer 1 under either method of taxation:

$$ {U^{2}}^{*}_{M} \geq {U^{1}}^{*}_{M}. \tag{2.28}$$

2.7.2 Socially optimal method of taxation

Of course, consumer i's equilibrium utility level does depend not only on the method of taxation $M \in \{\text{acc}, \text{real}\}$ but also on the tax rate $\tau \in]0,1[$. His indirect utility is henceforth denoted by ${U^i_M}^*(\tau)$. Now consider the problem of a social planner who maximizes social welfare by a choice in his two-dimensional policy space, i.e. deciding on the method of taxation and the tax rate. For the sake of concreteness, assume a weighted utilitarian type of (indirect) social welfare function V with

$$V(M,\tau) = \alpha {U^1}^*_M(\tau) + (1-\alpha){U^2}^*_M(\tau), \tag{2.29}$$

where $0 \leq \alpha \leq 1$. The socially optimal policy can be found by computing the optimal tax rates under either method of taxation τ^*_M and choosing the method of taxation that yields the highest level of welfare resulting from taxation at the corresponding optimal rate: $M^* \in \text{argmax}_M(V(M,\tau^*_M))$.

However, corresponding to the discussion in Section 2.4.3, first look at a situation in which the planner faces an exogenous revenue requirement, i.e. he has to provide a fixed amount g of the public good, and therefore sticks to the tax rate τ with $g = \tau(P_2 - P_0)$. From inequalities (2.27) it is obvious that a cut-off value $0 < \alpha_0 < 1$ exists such that the planner chooses accrual taxation for all $0 \leq \alpha \leq \alpha_0$ and a realization tax for all $\alpha_0 < \alpha \leq 1$. Put differently, whenever he puts enough weight on the utility of consumer 1, he chooses taxation upon realization.

2.7.3 Socially optimal tax rate

Now consider the planner's problem to maximize social welfare by choosing the optimal tax rate for a given method of taxation. The purpose of this section is to illustrate the fact discussed in Section 2.4.4, that in general the optimal tax rates differ, although for a given tax rate revenue is the same under both methods (see Lemma 2.1).

To this end, suppose $\alpha = 0$, i.e. the planner only cares about the well-being of consumer 2; by inequality (2.28), here this is equivalent to the assumption of a Maxmax social welfare function.[27] Using the optimal consumption levels of consumer 2, one derives the first order condition for the optimal tax rate under taxation upon realization

$$\frac{\partial V_{\text{real}}}{\partial \tau} = \frac{\partial {U^2}^*_{\text{real}}}{\partial \tau} = \frac{W - P_2}{(1-\tau)P_2 + \tau W} + \theta(P_2 - P_0) = 0. \tag{2.30}$$

[27]Analogously, $\alpha = 1$ is equivalent to the assumption of a Maxmin (Rawlsian) social welfare function.

The second order condition for a maximum is fulfilled ($\frac{\partial^2 V_{\text{real}}}{\partial \tau^2} < 0$) and the optimal realization tax rate is given by

$$\tau^*_{\text{real}} = \frac{P_2}{P_2 - W} - \frac{1}{\theta(P_2 - P_0)}. \tag{2.31}$$

If one assumes the taste parameter to be

$$\theta = \frac{2(P_2 - W)}{(P_2 - P_0)(P_2 + W)},$$

then the optimal realization tax rate equals $\tau^*_{\text{real}} = \frac{1}{2}$.

Now it is shown that the optimal accrual tax rate τ^*_{acc} must differ from $\frac{1}{2}$. Again, using the optimal consumption levels of consumer 2, the first order condition for the optimal tax rate under taxation upon accrual equals

$$\frac{\partial V_{\text{acc}}}{\partial \tau} = \frac{\partial U^{2*}_{\text{acc}}}{\partial \tau} = \frac{W - P_2\left[\frac{W+\tau W}{W+\tau P_0} + \tau \frac{W-P_0}{(W+\tau W)^2}\right]}{(1-\tau)P_2\frac{W+\tau W}{W+\tau P_0} + \tau W} + \frac{2(P_2 - W)}{P_2 + W} = 0. \tag{2.32}$$

However, substituting $\tau = \frac{1}{2}$ into (2.32) and rearranging terms, the expression on the left hand side proves to be negative and therefore $\tau^*_{\text{acc}} < \frac{1}{2} = \tau^*_{\text{real}}$, since $\frac{\partial^2 V_{\text{acc}}}{\partial \tau^2} < 0$. Put differently, shifting from accrual to realization taxation the planner tries to compensate consumer 2 for his reduction in private utility by a higher level of public good provision, which can be achieved by choosing a higher tax rate.[28]

[28]Of course, this is no general result but hinges on the special structure of the chosen example.

Part II

Public goods provision in large economies

The work presented in Part II of this book is the result of an ongoing joint project together with Felix Bierbrauer from the *Max Planck Institute for Research on Collective Goods* in Bonn. In the center of our research agenda we put the intent to combine the literature on public goods provision in the tradition of Clarke-Groves with the literature on optimal income taxation originating from Mirrlees (1971). The objective is based on the idea that the decisions about financing and providing a public good usually cannot be separated from each other but should be considered simultaneously.

The standard model in the literature on optimal income taxation is the continuum economy and the problem at hand is a screening problem: The planner tries to tax the agents according to their individual skills. On the other hand, the standard model in the literature on public goods provision is the finite economy and the problem at hand is a problem of information aggregation: The planner aims at providing the appropriate amount of public goods according to the aggregated individual valuations. Given that, if one tries to address both problems in a single model, it is convenient to include both relevant heterogeneities: skill and taste. However, one has to make a decision wether to use the framework of a finite or a continuum economy. Our choice is the continuum economy, because it captures best the fact we aim to model: the observation, that in a large economy a single individual has only a marginal influence on aggregate data.

Since this individual influence is only marginal, one may think that in a continuum economy the classical free rider problem is solved trivially, i.e. the social planner gets the relevant information for free and is able to provide the efficient amount of public goods according to the Samuelson rule. However, Chapter 3* illustrates that this point of view is not plausible in an environment with aggregate uncertainty where individuals differ with respect to the two dimensions, taste and skill. Instead, we argue that the process of information aggregation comes at a cost, since the social planner must provide additional incentives in order to learn the true state of the economy. Put differently, the Samuelson rule has to be modified to make public goods provision robust to marginal individual influence. It is shown that the necessary modification crucially depends on the extent of skill heterogeneity among the individuals.

In order to derive those results, we model the marginal individual influence as follows: For the purpose of information aggregation, we restrict the communication between the social planner and the agents to finite samples.

*Chapter 3 is based on our paper *Public goods provision in a continuum economy with two-dimensional heterogeneity.* I presented an earlier version of this paper at the Public Economic Theory Meeting (PET) 2005 in Marseille as well as at the Spring Meeting of Young Economists (SMYE) 2005 in Geneva.

Reporting their characteristics, sample members can influence the planner's state perception. As the sample size grows, this influence vanishes. The provision of public goods will be called robust to sampling, if it is immune against this vanishing influence. Chapter 4[**] introduces this notion of robustness for a more general framework. Moreover, it identifies this concept with a second one, the so-called requirement of informative subscription, which is an adaption from the literature on voting. This identification provides a useful tool for a great variety of problems of information aggregation in continuum economies with aggregate uncertainty.

[**]Chapter 4 is based on our paper *Robustness to sampling*.

Chapter 3

Public goods provision in a continuum economy with aggregate uncertainty

We consider the classical free-rider problem in public goods provision in a continuum economy. Individuals differ with respect to their valuation of a public good and with respect to their skills. We argue that the standard notion of incentive compatibility is not suitable for problems of information aggregation in continuum economies with aggregate uncertainty. We thus introduce a refinement of this solution concept, called robustness to sampling. When applied to the free-rider problem in a quasilinear environment it yields the following results: Skill heterogeneity implies that the optimal incentive compatible mechanism for public good provision is not robust. The extent of information aggregation undertaken by an optimal robust mechanism is inversely related to the polarization of preferences which results from the properties of the skill distribution.

3.1 Introduction

Public good provision faces the classical problem of *free-riding*. Individuals like to enjoy public goods but they do not like to contribute to the cost of provision. Still, one would like to provide public goods in such a way that the decision on provision reflects the *aggregate valuation* by the individuals who are supposed both to enjoy and to pay for it. However, the information on the aggregate valuation is dispersed among individuals. Consequently, allocation mechanisms which are to determine the quantity of a public good need to take the strategic reactions of individuals into account. Such an allocation mechanism hence has to fulfill the task of *information aggregation* under *incentive constraints*.

This chapter addresses the problem of information aggregation under incentive constraints in the context of a *large economy*, represented by a continuum of individuals. To this end, it introduces a concept for robust implementation in dominant strategies and applies it to a simple version of the classical free-rider problem. In Chapter 4 this concept, called *robustness to sampling*, is established for a more general framework and discussed in greater detail.

The reason for using the modelling device of a *continuum* to represent a large economy – as opposed to a large but *finite economy* – is analytical simplicity. As we argue in more detail below (see Section 3.2), a continuum economy has several convenient features which make it an attractive model of large economic systems. However, the existing literature on *implementation in continuum economies* has so far not developed an approach which allows to incorporate a free-rider problem. It implicitly takes the view, that, when it comes to the problem of information aggregation, strategic behavior is not an issue. As no single individual has an impact on the aggregate valuation of a public good, there are no gains to strategic behavior and hence information aggregation becomes a trivial task.

In Section 3.3, we provide a detailed discussion of the implications of this view in the context of a simple public goods problem and argue that it rests on unsatisfactory behavioral assumptions. The main point of the criticism is the following: even in a continuum economy, if individuals participate in a process of information aggregation and moreover have specific interests concerning the desired outcome of this procedure, then one should not impose behavioral assumptions on them which run counter to these interests.

In order to ensure that an allocation mechanism is not vulnerable to strategic behavior if a problem of information aggregation arises, Section 3.4 introduces the requirement of robustness to sampling. Underlying is the idea, that an allocation mechanism is trustworthy when it comes to information aggregation in a large economy, only if it survives giving individuals a marginal impact on the *perceived state* of the world. The robustness requirement is incorporated into a mechanism design problem with the objective to determine an optimal level of public goods provision as a function of the aggregate valuation, which is taken to be an unknown variable. Within this framework, the individual impact is modelled by *sampling*: First, on a stage of information aggregation, only data from a *finite sample* of individuals is used to estimate the aggregate valuation and to choose an appropriate provision level. Second, on a financing stage, the cost of providing the chosen amount of public goods has to be covered by the design of a contribution scheme that is incentive compatible for all agents in the economy.

Our application contributes to the literature on the free-rider problem in quasilinear environments (see Section 3.2 below for an account of this literature). However, it differs from the standard framework in the following way: A second source of individual heterogeneity is introduced into the model. As usual, individuals differ regarding their taste parameters which capture their marginal valuation of the public good. In addition, we allow for skill heterogeneity. For less skilled individuals it is harder to generate income and hence they suffer from a larger utility loss if forced to contribute to the cost of public good provision. The interaction of taste and skill parameters determines an individual's *effective valuation* of the public good.

As an example, think of a taxi driver, who works at night in the dangerous districts of a town. He would ceteris paribus be very happy to see more policemen in the street. On the other hand, he presumably risks being attacked only because he has a hard time generating income. Hence, having to give away money in order to contribute to this public good is really harmful for him. The balance of these two concerns pins down the quantity of the public good the taxi driver wants to have.

The presence of both skill and taste parameters allows to generate a rich set of observations, even in the very simple model we investigate. The most interesting ones, presented in Section 3.5, are as follows: Information aggregation causes *additional* incentive problems – i.e. on top of *individual incentive compatibility* in the financing scheme – if and only if there is skill heterogeneity. Moreover, by how much an optimal decision on public good provision may reflect the actual aggregate valuation, depends on the polarization of skills within the groups of individuals with the same tastes. The higher this *within-group* polarization, i.e. the stronger the effect of skill heterogeneity on the effective valuations, the less information aggregation is undertaken by an optimal rule for providing public goods.

In addition, as shown in Section 3.6, the proposed concept of robustness to sampling may serve as a theoretical foundation for the requirement of *informative voting*, which is known from the field of political economy. This condition requires that, whenever facing two alternatives, any individual with a low taste parameter is – irrespective of his skill – willing to vote for a lower amount of public goods, and any individual with a high valuation is – irrespective of his skill – willing to vote for a higher provision level. Our analysis demonstrates that the concepts of informative voting and robustness to sampling coincide in the limit as the sample size grows.

Section 3.7 contains concluding remarks on possible extensions of this chapter's analysis. All longer proofs can be found in the Appendix 3.8.

3.2 Relationship to the literature

In this section we comment on both related strands of the literature: the literature dealing with the characterization of implementable allocations in continuum economies as well as the literature addressing the free-rider problem of public goods provision in quasilinear environments with a countable number of individuals.

3.2.1 Implementation in continuum economies

In this paragraph, we first discuss what are, according to our view, the major advantages of using the model of a continuum economy, as opposed to a large but finite economy. We then debate why the literature on implementation in continuum economies has not taken problems of information aggregation into account. Finally, we explain how the solution concept that we propose tries to fill this gap.

A main convenience of a continuum of individuals for economic modelling is the following property: via some large numbers effect, the continuum assumption allows to separate *idiosyncratic uncertainty*, also referred to as *individual specific risk* or *assignment uncertainty*, from the behavior of aggregates.[1] Such randomness on the individual level may be combined with a *deterministic aggregate*, i.e. the randomness induced by individual-specific shocks washes out in the aggregate. Alternatively, there may be *aggregate uncertainty*, that is, the individual randomness washes out *conditional on a given state of the world*, the latter being a random variable as well. In either case, the central idea is that uncertainty in the behavior of aggregates can be separated from uncertainty referring to individual data.

As noted by Hammond (1979), this separation proves particularly useful in the description of *anonymous allocation mechanisms*. That is, how an individual with characteristics x is treated by an allocation mechanism depends on x and on some state s of the economy. Such a separation remains mathematically sound even if a state s is defined by a cross-section distribution of characteristics, e.g. by the fraction of individuals with property x in the whole system. Obviously, this kind of separation is not feasible in a finite economy as any change in the characteristics of one particular individual always affects the composition of the economy as a whole.

Under such an anonymous allocation mechanism, strategic interdependence is drastically reduced. The payoff of a specific individual depends on the reported characteristics of others only via aggregate data. These proper-

[1]See Alòs-Ferrer (2002) for a discussion of prominent applications.

ties imply that anonymous allocation mechanisms are an attractive tool for a characterization of outcomes which result from decentralized interactions, for instance via market transactions, tax systems or voting procedures.

As an example, the usefulness of such mechanisms has been demonstrated by Guesnerie (1995) in an attempt to provide a foundation for the consideration of tax systems in public economic theory.[2] Guesnerie (1995) identifies tax systems with the class of allocation mechanisms with the property of being decentralizable via some (possibly non-linear) price system. He uses the continuum assumption as a justification for a model which exhibits at the same time uncertainty about individual characteristics and a commonly known cross-section distribution of characteristics. The latter property permits to work with a deterministic public sector budget constraint. In addition it is assumed that individual characteristics are realizations of an independently and identically distributed (iid) random process. Guesnerie (1995) argues that, under those assumptions, it entails no loss of generality to restrict attention to tax systems, in the sense, that there do not exist allocation mechanism outside the class of anonymous allocation mechanisms with superior welfare properties.

This example demonstrates what we consider the main advantage of the continuum assumption: simplicity. It allows to make simultaneous use of the possibilities to work with a deterministic budget constraint, to separate individual and aggregate data in the description of an anonymous allocation mechanism, and finally to justify anonymity itself. Moreover, if the overall aim is an analysis of decentralized allocation mechanisms with a large number of participants these assumptions seem very plausible.

Our criticism of the literature on continuum economies is as follows: There is no study of implementation problems in which a problem of information aggregation causes incentive problems.[3] Something like a free-rider problem is entirely ignored by this literature. Typically, the cross-section distribution of characteristics in the economy is assumed to be commonly known and the analysis is hence limited to pure *screening* problems, i.e. the problem to determine for each agent his position in the predetermined cross-section.[4] This creates a concern of robustness. Whenever there is aggregate uncer-

[2]Further contributions to this line of research are Hammond (1987), Dierker and Haller (1990) or Hellwig (2003b, 2004b).

[3]An exception is Bierbrauer (2005). This paper combines a problem of information aggregation with a problem of optimal income taxation.

[4]This is true for all mentioned references from the area of public economic theory. Further studies which are based on a known cross-section distribution of characteristics are Makowski and Ostroy (1992) and Mas-Colell and Vives (1993).

tainty, i.e. the cross-section distribution of characteristics is not commonly known, an allocation mechanism has to fulfill two tasks: screening and information aggregation. That is, one has to figure out how the composition of the economy actually looks like while identifying the characteristics of each particular individual. However, if only the screening part of an allocation problem is taken into account, this implies a vulnerability to the possibility that individuals exploit their impact on the perceived state of the world.

Our objective is to fix this robustness problem in a constructive manner. We attempt to maintain all the convenient properties of the continuum economy mentioned above, while, at the same time, providing a more plausible treatment of the information aggregation part of the problem.

To this end we strengthen the solution concept of *individual incentive compatibility (I-IC)* by an additional requirement that we term robustness to sampling (RS). The term sampling refers to a hypothetical situation in which a randomly chosen large but finite subset of individuals – a large *random sample* of individuals – is, prior to the working of the final allocation mechanism, asked to report their characteristics. The reported data is only used in order to obtain an estimate of the distribution of characteristics in the economy. Hence, sample members have a small but strictly positive impact on the perceived state of the world. We say that RS holds if any individual, conditional on being in the sample, is willing to reveal the own characteristics.

We interpret this solution concept as a *refinement of I-IC*. It serves to eliminate those equilibria from the game induced by an individually incentive compatible allocation mechanism, that turns out to be weakly dominated in the presence of sampling. The overall idea is, at an intuitive level, best understood in comparison with the problem of voting over two alternatives in a continuum economy. Suppose there are two alternatives A and B and a continuum of individuals who have to decide on the chosen alternative via majority voting. One might argue that individual voting behavior is indeterminate. As no single individual is able to affect the 'numbers' of A-votes and B-votes, respectively, any voting behavior is a best response. Nevertheless, in the voting literature one often finds the following somewhat vague statement:[5] *Only sincere voting (i.e. an individual votes for A if and only if A is preferred) survives the elimination of weakly dominated strategies.* That is, it is required that individual behavior is such as if there was for each individual some chance of being decisive for the outcome of the voting procedure.

The analogy to our solution concept is as follows: Suppose an individual

[5] See e.g. Gersbach (2002) or Meirowitz (2004).

is indifferent between a variety of announcements because it has, in a continuum economy, no impact on the perceived state of the economy. Then our refinement requires to eliminate those best responses that turn out to be weakly dominated as soon as the individual is granted some influence on the perceived state of the economy.

3.2.2 The free-rider problem in finite economies with quasilinear preferences

Articles classified under this heading have the following setup in common: There is a number $n \in \mathbb{N}$ of individuals with quasilinear utility functions and *private information* on their valuation of a public good. The aggregate valuation of the public good is a random variable defined as the sum of individual valuations. The quantity decision on a public good as well as individual contributions to the cost of provision are determined via a direct revelation mechanism. The main question underlying the analysis is to what extent the so-called *free-rider problem* in public good provision may be resolved. I.e. how are the welfare costs of having to finance public good provision in such a way that any individual is willing to reveal his willingness to pay.

The focus of the early literature in that field has been the question whether one may have simultaneously ex-post efficiency and incentive compatibility. This question has been addressed with two different solution concepts, implementation in dominant strategies as well as the weaker notion of *Bayesian implementation*.[6] The more recent literature starts out from the observation that, even in a Bayesian framework, the requirements of ex-post efficiency and incentive compatibility imply that participation constraints cannot be respected.[7] Hence, the requirement of ex-post efficiency has to be dropped. This branch of literature has instead moved to an investigation of surplus-maximizing or second best allocation mechanisms under incentive as well as participation constraints.[8]

[6]Green and Laffont (1977) establish an impossibility result for implementation in dominant strategies, while Arrow (1979) and d'Aspremont and Gérard-Varet (1979) establish the possibility of implementing an efficient allocation as a Bayesian Nash equilibrium.

[7]This result is due to Güth and Hellwig (1986). These authors extend the famous Myerson and Satterthwaite (1983)-Theorem to a public goods problem with $n \in \mathbb{N}$ participants.

[8]Mailath and Postlewaite (1990) and Hellwig (2003a) analyze this question for a non-excludable public good and discuss how the welfare properties of an optimal allocation depend on the number of individuals who consume the public good. Schmitz (1997) and Norman (2004) observe that exclusion becomes a valuable instrument in this second-best setting and analyze how an optimal mechanism makes use of it.

As contrasted with this literature, we are considering the provision of a non-excludable public good in a continuum economy. This implies that there is neither a chance to respect participation constraints nor to make payment obligations a function of individual valuations of the public good: As, for any individual, the influence on the quantity decision is negligible, individuals are concerned only with a minimization of their contribution to the cost of provision. This implies that *equal cost sharing* turns out to be the only incentive compatible payment scheme. Moreover, if there is equal cost sharing and participation constraints are imposed, then the maximal amount of public good provision is determined by those individuals with the lowest valuation of the public good. This implies that the quantity decision on the public good may react to changes in the aggregate valuation only if there is the possibility to extract payments from individuals irrespective of their own valuation.[9]

These observations are the starting point of our analysis. As we want to study the public goods problem in a continuum economy, we have to accept the restriction of equal cost sharing as well as the fact that our analysis applies only to those situations in which there is a right of coercion. However, accepting the need of coercion does not yet answer the question how to employ it. The fact that all individuals, irrespective of their characteristics, are treated equally by the payment scheme implies a rich pattern of opposing views on the desirable level of public good provision. We impose the requirement that any procedure of information aggregation, even in a continuum economy, has to respect these interests. This allows to derive a set of incentive constraints, which are imposed in addition to equal cost sharing. We thus finally arrive at a second best problem of public good provision for a continuum economy.

3.3 The problem of information aggregation

This section illustrates by means of an example why the concept of *individual incentive compatibility* yields unreasonable outcomes if used as an instrument of information aggregation in a large economy. The chosen application is a problem of public goods provision under private information of individuals on their valuation of the public good. As will become clear, with *quasilinear preferences*, this public goods problem is essentially a problem of information aggregation under incentive constraints. I.e. the only question of interest is how to elicit the aggregate valuation of the public good.

[9] This observation has been made earlier by Olsen (1965), who defines a *large group* by the necessity of coercion for public good provision.

3.3.1 Basic assumptions

Individual characteristics

The economy under consideration consists of a continuum of individuals $i \in I := [0,1]$. All individuals are identical but with respect to two parameters, namely their individual *ability* to earn income, also referred to as their *skill* w^i, and their individual *valuation* of the public good, also referred to as their *taste* θ^i. For simplicity, assume that the taste parameter may take only two different values:

$$\theta^i \in \Theta := \{\theta_L, \theta_H\} \quad \text{with} \quad 0 \leq \theta_L < \theta_H,$$

where θ_L stands for a low valuation of the public good and θ_H for a high one. The productivity parameter belongs to the compact interval

$$w^i \in \Omega := [\underline{w}, \bar{w}] \quad \text{with} \quad 0 < \underline{w} < \bar{w},$$

where $\underline{w}$ denotes the lowest possible earning ability and $\bar{w}$ the highest one.

Each individual derives utility from the consumption of a public good, but he does not like to contribute to its provision. In particular, we use the following additively separable and quasi-linear function to express agent i's utility:

$$U^i = \theta^i Q - \frac{t^i}{w^i} \ .$$

Q denotes the quantity of a non-excludable public good and t^i captures i's contribution to the cost of public good provision. Note that lower levels of skill w^i translate into a larger utility loss due to a given payment obligation. This reflects the idea that, for less able individuals, it is harder to generate the income needed to meet a given payment obligation.

The function U^i is the cardinal representation of preferences which is relevant for welfare assessments. An individual's ranking of alternatives can be equivalently expressed by the monotone transformation

$$w^i U^i = \theta^i w^i Q - t^i \ .$$

We refer to the product $\theta^i w^i$ as individual i's *effective valuation* of the public good.

Informational structure

The parameters w^i and θ^i are both private information of individual i. We assume the w^is as well as the θ^is to be *independently and identically distributed (iid)* and both distributions to be independent of each other. The skill distribution is represented by a function $F : \Omega \to [0,1]$ with density f. Moreover, we define

$$p := \mathrm{Prob}\{\theta^i = \theta_H\}$$

as the probability for individual i to have a high valuation for the public good.

In the following we assume that a *law of large numbers (LLN)* applies[10] in the sense that the value $F(w)$ and the probability p can be interpreted as the fractions of individuals with earning ability $w^i \leq w$ and high valuation for the public good in the population respectively. While the distribution F is known to the policy maker, i.e. there is no uncertainty about skill distribution on the aggregate level, p is a random variable of its own. I.e. there is uncertainty with respect to the aggregate valuation of the public good. The planner only has a *prior belief* on the distribution of p. For ease of calculation, this prior is assumed to be a uniform distribution of p on $[0,1]$. Note that, since we are interested in robust mechanisms and therefore consider implementation in dominant strategies as equilibrium concept, we do not need to impose any *common prior assumption*.

To sum up, the information structure exhibits a deterministic aggregate regarding the productivity parameter and aggregate uncertainty with respect to the taste parameter. That is, individual differences in productivity entirely disappear on the aggregate level while the aggregate distribution of taste is not taken to be known. Thus, the unknown parameter p is henceforth also referred to as the *state of the economy*. It is the relevant object for the process of information aggregation.

3.3.2 Individual incentive compatibility (I-IC)

We now describe what *individual incentive compatibility* (*truthful implementation in dominant strategies*) means for the problem at hand. In Chapter 4 this concept is defined in the context of a general allocation problem.

Individually incentive compatible allocations

An *anonymous allocation rule* (Q,t) consists of a *provision rule* for the public good and a *payment scheme* to cover the cost of provision. The provision

[10] The problem, that the LLN may not hold for a continuum of iid random variables, i.e. a stochastic process, is discussed, for example, in Judd (1985), Feldman and Gilles (1985), Sun (1998), Alòs-Ferrer (2002), and Al-Najjar (2004). These authors offer several opportunities how to deal with it. A more rigorous description of the informational structure is provided in Chapter 4.

rule Q assigns to alternative values of p a quantity of the public good,

$$Q : [0,1] \to \mathbb{R}_+, \quad p \mapsto Q(p).$$

The payment scheme t specifies for each individual i a payment obligation as a function of the distribution of characteristics in the economy p and individual i's characteristics (θ^i, w^i). The payment scheme is anonymous in the sense that individuals with the same characteristics have the same payment obligation, in every state p of the economy. Put differently, individual payments do not depend on the index i. Formally the payment scheme is described as a function

$$t : [0,1] \times \Theta \times \Omega \to \mathbb{R}, \ (p, \theta, w) \mapsto t(p, \theta, w).$$

Definition 3.1 (I-IC) *An anonymous allocation rule is called individually incentive compatible (I-IC) if* $\forall p \in [0,1]$, $\forall(\theta, w) \in \Theta \times \Omega$, *and* $\forall(\hat{\theta}, \hat{w}) \in \Theta \times \Omega$,

$$\theta w Q(p) - t(p, \theta, w) \geq \theta w Q(p) - t(p, \hat{\theta}, \hat{w}).$$

The I-IC property requires that for any given state p of the economy no individual has an incentive to hide his true characteristics.

Remark 3.1 *In the continuum economy, the concept of I-IC is equivalent to the concept of truthful implementation in dominant strategies. An anonymous allocation rule is said to be truthfully implementable in dominant strategies (straightforward incentive compatible)*[11] *if* $\forall i \in [0,1]$, $\forall(\theta^i, w^i) \in \Theta \times \Omega$, $\forall(\hat{\theta}^i, \hat{w}^i) \in \Theta \times \Omega$, *and* $\forall(\theta_{-i}, w_{-i}) \in (\Theta \times \Omega)^{I \setminus \{i\}}$

$$\theta^i w^i Q(p) - t(p, \theta, w) \geq \theta w Q(\hat{p}) - t(\hat{p}, \hat{\theta}, \hat{w}),$$

where $p = p((\theta^i, \theta_{-i}), (w^i, w_{-i}))$ *and* $\hat{p} = p((\hat{\theta}^i, \theta_{-i}), (\hat{w}^i, w_{-i}))$. *However, in the continuum, under an anonymous allocation rule, no single individual has an impact on the (perceived) state of the economy, i.e.* $p = \hat{p}$.

In any state of the world p, incentive compatibility requires for each individual the minimal payment obligation. Formally, we arrive at the following observation, which results immediately from the definition of I-IC.

Lemma 3.1 *The following statements are equivalent.*

1. (Q, t) *is I-IC.*

[11] See Mas-Colell et al. (1995, Chapter 23).

2. (Q, t) *satisfies* $\forall p \in [0,1]$, $\forall(\theta, w) \in \Theta \times \Omega$ *and* $\forall(\hat{\theta}, \hat{w}) \in \Theta \times \Omega$,

$$t(p, \theta, w) = t(p, \hat{\theta}, \hat{w}).$$

As the economy is large, no single individual has an impact on the level of public good provision. This implies that incentive compatibility can be achieved only via an equal treatment of all individuals by the payment scheme. This is the classical *free-rider problem*. As access to the public good is free, no one is willing to pay more than he is forced to.[12]

Consequently, any *I-IC* payment scheme is constant, in the sense that, for given p, all individuals are treated equally. The converse statement is also true. That is, any anonymous provision rule $Q : p \mapsto Q(p)$ gives rise to an *I-IC* allocation rule if accompanied by constant payments, i.e. a payment scheme that does only depend on p.

The budget constraint and equal cost sharing

The costs of public good provision are represented by a twice continuously differentiable, strictly increasing and strictly convex cost function $K : \mathbb{R}_+ \to \mathbb{R}_+$, which satisfies $K(0) = 0$ as well as the boundary conditions

$$\lim_{x \to 0} K'(x) = 0 \quad \text{and} \quad \lim_{x \to \infty} K'(x) = \infty.$$

The payment scheme has to be such that the costs of public good provision are covered. Combining the requirements of I-IC and of *budget balance* (*feasibility*) yields the following observation:

Lemma 3.2 *The only admissible payment scheme under I-IC is equal cost sharing, i.e.*

$$\forall p \in [0,1], \forall(\theta, w) \in \Theta \times \Omega : \quad t(p, \theta, w) = K(Q(p)).$$

We summarize the observations made in Lemma 3.1 and Lemma 3.2 as follows:

Proposition 3.1 *An anonymous allocation rule* (Q, t) *is I-IC and budget balancing if and only if the payment scheme exhibits equal cost sharing.*

Due to Proposition 3.1, in what follows, we restrict the search for an appropriate feasible I-IC anonymous allocation rule (Q, t) to the problem of finding an appropriate provision rule Q accompanied by equal cost sharing.

[12] Note that if budget balance has to be achieved and there are limits to coercion due to participation constraints as in Mailath and Postlewaite (1990) or in Hellwig (2003a) and there are individuals who do not value the public good at all – i.e. with effective valuation of 0 – one will end up with $Q \equiv 0$ under any admissible, incentive compatible provision rule.

3.3.3 Information Aggregation under I-IC

The resolution of aggregate uncertainty requires to figure out the true state of the economy, i.e. the fraction p of individuals with a high taste parameter. By definition, the I-IC constraints ensure that no individual has an incentive to hide the own characteristics for a given cross-section distribution of characteristics. However, under aggregate uncertainty, an incentive compatible allocation rule is used in addition for the purpose of information aggregation as the mechanism designer has to deduce the actual distribution of characteristics from the profile of individual announcements. To make this more explicit, it is instructive to think of the *revelation game* in the following sequential manner:

Sequence 3.1

Stage 1: *The mechanism designer specifies a provision rule Q accompanied by equal cost sharing. I.e. there is a distinct level of public good provision $Q(p)$ and a distinct payment obligation $K(Q(p))$ for each possible state p of the economy.*

Stage 2: *The mechanism designer collects all individual data and uses this information to deduce the actual value $\hat{p}$ for the state of the economy.*

Stage 3: *According to the specified rule, the amount $Q(\hat{p})$ of the public good is provided. Individuals make the corresponding contributions $K(Q(\hat{p}))$.*

Multiple equilibria

As outlined in Section 3.2, the problem of aggregate uncertainty in large economies has not been addressed yet by the literature. An exception is Hammond (1979). Based on his work, one could take the view that the problem of information aggregation is resolved trivially as a corollary of I-IC. Suppose, with no loss of generality, that, on stage 2, the planner can communicate with all agents, asks them to report their characteristics[13], and is able to measure the fraction

$$\hat{p} = \mu(\{i \in I \mid \hat{\theta}^i = \theta_H\})$$

of high valuation reports among the population[14]. The allocation rule, proposed on stage 1, exhibits equal cost sharing and specifies the level of public

[13]The revelation principle tells that any achievable allocation can be attained by such a direct mechanism.

[14]The problem of measurability is discussed, for example, by Judd (1985) or Al-Najjar (2004), showing that an appropriate measure μ can be constructed. More details are provided in Chapter 4.

good provision as a function of the perceived state of the economy $\hat{p}$. Since the economy is large, no single agent i is able to influence the planner's perception $\hat{p}$ by his report $\hat{\theta}^i$. Hence, no individual has a payoff relevant move. This implies that revealing the own characteristics is a best response. Of course, the same can be said about any reporting behavior.

In other words, there are *multiple equilibria* and one faces a problem of *equilibrium selection*. From the angle just outlined, this problem could be solved trivially by just breaking the agents' indifference in favor of *truth-telling*. Hence, the planner would learn the true state of the world p and the problem of information aggregation would disappear. It would be resolved as a byproduct of I-IC.

However, we doubt this view and aim at offering an alternative one. To illustrate the problem, we first explore the consequences of the perspective that the need of information aggregation does not create additional incentive problems, once I-IC has been taken into account and the costs of public good provision are covered via equal cost sharing.

The optimal utilitarian allocation under I-IC

Given this notion of implementability we can solve for the provision rule which is optimal from a utilitarian perspective. For given p and an arbitrary provision rule Q, utilitarian welfare under equal cost sharing is given by

$$(p\theta_H + (1-p)\theta_L)\, Q(p) - \lambda K(Q(p)) \quad \text{with} \quad \lambda := \int_{\underline{w}}^{\bar{w}} \frac{f(w)}{w} dw.$$

The parameter λ can be interpreted as *the shadow cost of public funds*.[15] We may equivalently assume that the utilitarian objective is to maximize

$$\bar{v}(p)Q(p) - K(Q(p)) \quad \text{with} \quad \bar{v}(p) := \frac{p\theta_H + (1-p)\theta_L}{\lambda}$$

and refer to the term $\bar{v}(p)$ as the *effective utilitarian valuation* or as the *effective aggregate valuation* of the public good if the state of the economy is p.

Now suppose that I-IC is a sufficient condition to guarantee that anyone will truthfully report his taste. Using his prior beliefs on p being uniformly

[15]From a utilitarian welfare perspective, equal cost sharing as implied by I-IC, is a distortion. A utilitarian planner would prefer that only individuals with earning ability $\bar{w}$ pay for the public good.

distributed, the planner assesses expected utilitarian welfare to equal

$$EW := \lambda \int_0^1 \{\bar{v}(p)Q(p) - K(Q(p))\}\, dp.$$

The provision rule which maximizes this expression is denoted Q^* and will serve as a benchmark in the following. In combination with equal cost sharing it forms the optimal feasible allocation rule that satisfies the I-IC constraints. Q^* is characterized by the following version of the *Samuelson rule* for public good provision.

Proposition 3.2 *The optimal feasible provision rule $Q^* : p \mapsto Q^*(p)$ under I-IC is characterized by a continuum of first order conditions,*

$$K'(Q^*(p)) = \bar{v}(p) \quad \textit{for} \quad p \in [0,1].$$

Q^* is strictly increasing and continuously differentiable with derivative $Q^{*\prime}$.

Proof. The properties of Q^* follow immediately from the planners maximization problem taking into account the assumptions on K.

□

The need of a refinement

We now illustrate that the admissibility of the optimal utilitarian allocation Q^* requires that agents use actions which are not robust in the following sense: Agents are assumed to behave in a way that would be incompatible with implementation in dominant strategies as soon as they were given an arbitrary small influence on the perceived state of the world. Hence, for the problem at hand, the notion of individual incentive compatibility lacks plausibility.

To see this, note first that the utility realized by an individual under provision rule Q^* is a function of the individual's characteristics (θ, w) and the unknown parameter p. This indirect utility function is henceforth written as

$$U^*(p,\theta,w) = \theta Q^*(p) - \frac{K(Q^*(p))}{w}.$$

Its partial derivative with respect to p equals

$$U_1^*(p,\theta,w) = \frac{1}{w}Q^{*\prime}(p)[\theta w - \bar{v}(p)].$$

Consequently, U^* is increasing in p as long as $\theta w > \bar{v}(p)$, i.e. the individual's effective valuation of the public good exceeds the effective valuation of a utilitarian planner. Analogously, U^* is decreasing in p if θw falls short of the utilitarian valuation.

Now suppose, for the sake of concreteness, that p is such that[16]

$$\theta_L \bar{w} > \bar{v}(p) > \frac{\theta_L}{\lambda}.$$

This implies that there exists a critical value $\hat{w} \in]\underline{w}, \bar{w}[$ such that all individuals with $\theta^i = \theta_L$ and $w^i < \hat{w}$ have an effective valuation $\theta_L w^i$ which falls short of the utilitarian planner's effective valuation $\bar{v}(p)$ and hence, according to their indirect utility function, would prefer a slightly lower perceived value of p. Analogously, individuals with $\theta^i = \theta_L$ and $w^i > \hat{w}$ have an effective valuation $\theta_L w^i$ exceeding the one of the utilitarian planner. Therefore, they would prefer a slightly larger perceived value of p.

Despite those conflicting interests, under the truth-telling assumption, individuals in both sets are assumed to behave the same way, namely to reveal their low valuation of the public good. In particular, an individual with $\theta_L w > \bar{v}(p)$ is assumed not to exaggerate when reporting the own taste for the public good, even though this individual would be happy if the utilitarian planner could be induced to believe that the effective aggregate valuation of the public good was in fact higher.

In more abstract terms, it has been noted above that, in a revelation game with equal cost sharing, any kind of behavior constitutes a best response as no individual is able to affect with the own announcement the planner's perception of the aggregate valuation $\bar{v}(p)$. At the same time individuals are not indifferent regarding this perception. They just have no direct influence on it. Still, they might want to marginally increase their indirect utility by strategic reports in favor of their preferred state perception. Moreover, being involved in the revelation game forces any individual to subscribe either to the group of individuals with a high taste parameter or to the group of individuals with a low taste parameter. Hence, the assumption that under provision rule Q^* any individual tells the truth, amounts to the postulate that some individuals do not subscribe to the group whose size they would be willing to support, but instead join the 'wrong' group. In that sense, a solution that takes only the I-IC constraints into account is not robust to the (illusion of) indirect individual influence on the perceived aggregate valuation.

Note that this is not only a concern of whether or not it is acceptable to break individual indifference in favor of truth-telling in the presence of

[16] As $\bar{v}(p)$ is a convex combination of $\frac{\theta_H}{\lambda}$ and $\frac{\theta_L}{\lambda}$, for any $x \in [\frac{\theta_L}{\lambda}, \frac{\theta_H}{\lambda}]$ there exists p such that $\bar{v}(p) = x$.

multiple best responses. The point is, that individual indifference is taken too literally. I.e. indifference is broken in favor of truth-telling even though individuals should not be regarded as indifferent. Hence, in order to solve the problem of multiple equilibria, we suggest to switch from the behavioral assumption of truth-telling to the more plausible one that individuals behave as if they had some (arbitrarily small) influence.

These observations serve as a starting point for the next section, which introduces a refinement of individual incentive compatibility for the problem at hand. Chapter 4 discusses this concept, called robustness to sampling, in more detail and for a more general framework.

3.4 Robustness to sampling as a refinement

The basic idea behind the refinement of individual incentive compatibility we propose is to postulate that the equilibrium allocation be robust to the following *perturbation* of the revelation game: The allocation rule must be implementable even if the individuals were provided with some influence on the planner's perceived state of the economy. Our approach of granting them some influence is to restrict the communication between the planner and the agents for the purpose of information aggregation to a finite random sample $S_N \subset I$ of individuals with sample size $N \in \mathbb{N}$.

The corresponding revelation game is defined in Section 3.4.1 as a preliminary step for the definition of our refinement concept. The planner uses the sample data on taste parameters to update his prior beliefs about the actual state of the economy. Based on this estimation he decides on a level of public good provision Q. It is assumed that the mechanism designer's estimation procedure is commonly known. This implies that all sampled agents can predict how their own taste report affects the provision level. Consequently, the former indifference is broken such that their taste announcement always contributes to their most preferred state of the world. Hence, in order to ensure that sampled individuals still are willing to reveal their true taste parameter, an additional set of incentive constraints comes into play, labelled robustness to sampling of size N (RS_N). These additional constraints allow to formulate the problem P_N of finding an optimal robust provision rule based on sampling of size N.

Making an appeal to purely formal analogy, in Section 3.4.2 the RS_N conditions are extended to a revelation game in the continuum in which not only a subset of sampled agents but all individuals are involved in the process of information aggregation. The resulting constraints prevent individuals from strategic reports and ensure sincere ones instead. Thus, these conditions are referred to as the requirement of informative voting (IV). The corresponding problem P of finding the optimal provision rule subject to IV is called the *informative voting problem.* Our definition of informative voting is more than a formal analogy to robustness to sampling of size N in the finite sample: As will be shown in Section 3.6, the solutions of the *finite problems* P_N converge to the solution of the informative voting problem P. Put differently, the optimal provision rule passing our refinement test can be understood as the limit outcome of vanishing individual influence on aggregate data. In this sense, the concept of robustness to sampling serves as a theoretic foundation for the concept of informative voting.

3.4.1 Finite samples

Suppose that, for the purpose of information aggregation, the mechanism designer communicates with only a finite number N of individuals by drawing a random sample $S_N \subset I$.

Information aggregation based on sampling

In order to get information about the unknown state of the economy p, the social planner evaluates the profile of reported taste parameters $\theta_N := (\theta^1, \dots, \theta^N)$, where θ^i denotes the taste announced by the sample member at position $i \in \{1, \dots, N\}$. He uses the number m of high taste reports,

$$m = \#\{i \in S_N \mid \theta^i = \theta_H\},$$

to update his prior beliefs about the true state of the economy p applying Bayes' rule. Under the assumption of p being uniformly distributed on $[0, 1]$, this yields the following *conditional density function* ϕ_N, referred to as perceived state of the economy.

Lemma 3.3 *Suppose the social planner has prior beliefs of* p *being uniformly distributed on* $[0, 1]$*. After observing* m *high taste reports in a sample of size* N*, he will update his beliefs such that the perceived state of the world is given by the conditional density* ϕ_N *with*

$$\phi_N(p \mid m) = (N + 1)\binom{N}{m}p^m(1 - p)^{N-m}. \tag{3.1}$$

The proof is provided in Appendix 3.8.

Implementable provision rules based on sampling

Having clarified how the mechanism designer uses the sample information to update his prior beliefs on the likelihood of different states of the economy, we now characterize conditions for an allocation rule to be truthfully implementable under sampling of size N.

An anonymous *allocation rule based on sampling of size* N, (Q_N, t_N), consists of a provision rule for the public good

$$Q_N : \{0, 1, \ldots, N\} \to \mathbb{R}^0_+, \quad m \mapsto Q_N(m)$$

and a payment scheme

$$t_N : \{0, 1, \ldots, N\} \times \Theta \times \Omega \to \mathbb{R},\ (m, \theta, w) \mapsto t_N(m, \theta, w).$$

Based on the sample observation m, the mechanism designer updates his beliefs, provides an appropriate amount $Q_N(m)$ of public goods, and chooses an anonymous payment scheme $t_N(m, \theta, w)$ in order to cover the cost of provision.

Note that the assumption of anonymity with respect to the payment scheme implies a *no discrimination of sampled individuals*-property. That is, at the financing stage, individuals who happen to be in the sample are not treated differently as individuals with the same characteristics who happened not to be in the sample. In particular, this precludes the possibility to design a specific scheme of payments for sample members to let them internalize the impact of their statements on the whole system. In a word, to specify the individual payments, again the planner must impose I-IC constraints:[17] $\forall m \in \{0, \ldots, N\}$, $\forall(\theta, w) \in \Theta \times \Omega$, and $\forall(\hat{\theta}, \hat{w}) \in \Theta \times \Omega$,

$$\theta w Q_N(m) - t_N(m, \theta, w) \geq \theta w Q_N(m) - t_N(m, \hat{\theta}, \hat{w}).$$

As before, one verifies immediately that a feasible allocation rule satisfying the I-IC constraints must exhibit equal cost sharing, i.e. for all $m \in \{0, \ldots, N\}$ and for all $(\theta, w) \in \Theta \times \Omega$

$$t_N(m, \theta, w) = K(Q_N(m)).$$

To sum up, whenever we consider allocation rules based on sampling of size N, we refer to the following sequence of actions:

[17] However, here the condition is not used to collect information about the state of the economy, but only to identify the individuals' positions on the skill distribution in order to fix their contributions to the public good. See Chapter 4 for an alternative interpretation of I-IC.

Sequence 3.2

Stage 1: *The mechanism designer specifies a provision rule Q_N accompanied by equal cost sharing. I.e. there is a distinct level of public good provision $Q(m)$ and a distinct payment obligation $K(Q(m))$ for each possibly perceived state m of the economy.*

Stage 2: *The mechanism designer collects individual data from the members of sample S_N and uses this information to update his beliefs on p according to the perceived state of the economy $\phi_N(\cdot \mid \hat{m})$.*

Stage 3: *According to the specified rule, the amount $Q(\hat{m})$ of the public good is provided. Individuals make the corresponding contributions $K(Q(\hat{m}))$.*

Unlike the parameter p in the continuum, under sampling the parameter m, i.e. the perceived state of the world, depends on the taste report θ of each single individual within the sample; we account for that fact denoting $m(\theta_N)$. In order to ensure that, on stage 2, the sample members truthfully report their taste parameters, the planner must impose an additional constraint. Let θ_{N-i} be the vector of characteristics for all sample members with exception of individual i. Then, $\forall i \in S_N$, $\forall(\theta^i, w^i) \in \Theta \times \Omega$, $\forall(\hat{\theta}^i, \hat{w}^i) \in \Theta \times \Omega$, and $\forall(\theta_{N-i}, w_{N-i}) \in (\Theta \times \Omega)^{N-1}$

$$\theta w Q(m) - K(Q(m)) \geq \theta w Q(\hat{m}) - K(Q(\hat{m})), \tag{3.2}$$

with $m = m(\theta^i, \theta_{N-i})$ and $\hat{m} = m(\hat{\theta}^i, \theta_{N-i})$ must hold.

If an agent i with characteristics (θ^i, w^i) happens to be in the sample, misreporting his true type, he can increase m by 1 if $\theta^i = \theta_L$ or decrease m by 1 if $\theta^i = \theta_H$, respectively. Reformulating (3.2), this gives rise to the following definition:

Definition 3.2 ($\mathbf{RS}_N$) *A provision rule Q_N is called robust to sampling of size N (RS_N) if the following inequalities hold for all $m \in \{0, \ldots, N-1\}$ and for all $w \in \Omega$:*

$$\begin{aligned} \theta_L w Q_N(m) - K(Q_N(m)) &\geq \theta_L w Q_N(m+1) - K(Q_N(m+1)) \,, \\ \theta_H w Q_N(m) - K(Q_N(m)) &\leq \theta_H w Q_N(m+1) - K(Q_N(m+1)) \,. \end{aligned} \tag{3.3}$$

The RS_N constraints require that, whatever the distribution of characteristics in the sample, an individual with a low taste parameter should prefer a lower perceived state of the world m and an individual with a high taste parameter should prefer a larger level of m. The RS_N requirement ensures *ex post incentive compatibility* in the following sense: No sample member would ever want to revise his taste announcement after having learned the distribution of the announcements induced by all other $N-1$ sample members.

Remark 3.2 *The normal form representation of the revelation game based on sampling as stated in Sequence 3.2 can be described as the following perturbation of the original revelation game as stated in Sequence 3.1: Again, the planner asks all individuals in the economy to report their characteristics. Now, however, he commits to use only the taste reports from the members of a random sample S_N for the purpose of information aggregation.*

In Appendix 3.8 we show that a feasible allocation rule induces a dominant strategy equilibrium of this game if and only if it consists of a RS_N provision rule accompanied by equal cost sharing. In other words, the concept of robust implementation based on sampling of size N can be interpreted as applying the concept of truthful implementation in dominant strategies on the perturbed revelation game described above.

We provide some alternative formulation of RS_N in the following Lemma.

Lemma 3.4 *The following statements are equivalent:*

1. *Q_N satisfies RS_N.*

2. *For all $m, m' \in \{0, \ldots, N\}$ with $m < m'$ and for all $w \in \Omega$:*

$$\begin{aligned} \theta_L w Q_N(m) - K(Q_N(m)) &\geq \theta_L w Q_N(m') - K(Q_N(m')) \,, \\ \theta_H w Q_N(m) - K(Q_N(m)) &\leq \theta_H w Q_N(m') - K(Q_N(m')) \,. \end{aligned} \tag{3.4}$$

3. *For all $m, m' \in \{0, \ldots, N\}$ with $m < m'$:*

   i) Q_N is increasing: $Q_N(m) \leq Q_N(m')$.

   ii) If $Q_N(m) < Q_N(m')$, then

$$\theta_H \underline{w} \;\geq\; \frac{K(Q_N(m')) - K(Q_N(m))}{Q_N(m') - Q_N(m)} \;\geq\; \theta_L \bar{w} \,.$$

The proof is provided in Appendix 3.8.

The mechanism design problem based on sampling

In this paragraph we formulate the utilitarian planner's problem of finding an optimal robust provision rule based on sampling of size N. A utilitarian planner chooses the provision rule Q_N which maximizes expected utilitarian welfare from an *ex-ante perspective*, that is on stage 1 before m is observed. We first derive an explicit expression for this utilitarian objective based on the assumption about the planner's prior beliefs and the conditional updating stated in Lemma 3.3.

Lemma 3.5 *Suppose the utilitarian planner views p as the realization of a random variable that is uniformly distributed on [0,1]. Then, ex-ante expected utilitarian welfare provided by a provision rule based on sampling of size N, Q_N, equals*

$$EW_N = \lambda \frac{1}{N+1} \sum_{m=0}^{N} \left\{ \bar{v}\left(\frac{m+1}{N+2}\right) Q_N(m) - K(Q_N(m)) \right\} . \qquad (3.5)$$

The proof is provided in Appendix 3.8. In particular, a planner who observes a sample in which m individuals have a high valuation of the public good ends up – after his prior beliefs are updated – with an effective valuation of the public good given by

$$\bar{v}\left(\frac{m+1}{N+2}\right) = \frac{m+1}{N+2}\frac{\theta_H}{\lambda} + \frac{N-m+1}{N+2}\frac{\theta_L}{\lambda} .$$

Note that the effective valuation is strictly increasing in m, and for all $m \in \{0, \ldots, N\}$ it exceeds $\bar{v}(0)$ and falls short of $\bar{v}(1)$.

Now we are ready to state the planner's problem of finding an optimal robust provision rule based on sampling of size N.

Definition 3.3 *The following problem is called the finite problem based on sampling of size N, henceforth referred to as P_N:*

Maximize EW_N as given by equation (3.5) by the choice of a robust provision rule Q_N, i.e. fulfilling the RS_N constraints as stated in (3.4), accompanied by equal cost sharing, i.e. fulfilling the budget constraint as well as I-IC.

The solution to this problem is denoted by Q_N^{**}, the induced optimal welfare level by EW_N^{**}.

We omit to explicitly solve the problem here. As will become clear, however, the structure of its solution is very similar to that of the solution to the so-called informative voting problem we will introduce in the next section.

3.4.2 Informative voting (IV)

We now turn back to the originally public goods problem as sketched in Sequence 3.1, i.e. the planner communicates with all individuals in the economy. Recall our actual object of interest: We aim at sorting out provision rules that are based on agents taking actions that would counteract their interests if they had some influence on the planner's decision.

The notion of informative voting

This will be achieved in the following by imposing additional constraints on the public goods problem that exhibit perfect formal analogy to the RS_N constraints for finite samples as stated in equation (3.4).

Definition 3.4 (IV) *A provision rule Q is said to fulfill the constraints of informative voting (IV) if the following inequalities hold for any pair $p, p' \in [0,1]$ with $p < p'$ and any $w \in \Omega$:*

$$\begin{aligned} \theta_L w Q(p) - K(Q(p)) &\geq \theta_L w Q(p') - K(Q(p')) \ , \\ \theta_H w Q(p) - K(Q(p)) &\leq \theta_H w Q(p') - K(Q(p')) \ . \end{aligned} \tag{3.6}$$

Hence, under the IV constraints no individual ever has an incentive to support a higher perceived state of the economy p' if he has a low valuation θ_L and no individual ever has an incentive to support a lower perceived state of the economy p if he has a high valuation θ_H. The notion of informative voting (IV) reflects an interpretation based on the idea that any individual subscribes to one of two groups, either to those individuals with $\theta^i = \theta_L$ or to the group with $\theta^i = \theta_H$. As there are only two such groups a decision on subscription may alternatively be framed as a voting decision for a binary choice between the alternatives $Q(p)$ and $Q(p')$. Informative voting hence ensures that each individual supports the group which shares the own taste characteristic.[18]

As will become clear, the connection between IV and RS_N is not only a formal but a substantial one. In Section 3.6 we will show that the solutions Q_N^{**} of the problems P_N converge, in some sense, to the solution Q^{**} of the problem P induced by informative voting as defined in the next paragraph. This result ensures that, in order to investigate the impact of vanishing individual influence, as well as imposing the RS_N constraints for growing sample sizes one can rely on the IV constraints as the relevant ones. Consequently,

[18] A more extensive discussion of the notion of informative voting can be, for instance, found in Austen-Smith and Banks (1996).

in the following we will first concentrate on the analysis of the IV constraints and their impact on the public goods problem at hand. Afterwards we will come back to the finite problems P_N and formally establish the mentioned convergence.

The following Lemma is a straightforward adaption of Lemma 3.4 to the present context.

Lemma 3.6 *A provision rule Q satisfies IV, if and only if the following two properties hold for any pair $p, p' \in [0,1]$ with $p' > p$:*

i) Q is increasing: $Q(p) \leq Q(p')$.

ii) If $Q(p') > Q(p)$, then

$$\theta_H \underline{w} \;\; \geq \;\; \frac{K(Q(p')) - K(Q(p))}{Q(p') - Q(p)} \;\; \geq \;\; \theta_L \bar{w} \; .$$

A first step on the way to establish the more substantial connection between IV and RS_N is the following observation. Let Q be any provision rule and denote by $Q_{|N}$ its 'restriction' to the domain $\{0, 1, \ldots, N\}$ of a provision rule based on sampling of size N, i.e. $Q_{|N}(m) := Q(\frac{m}{N})$.

Lemma 3.7 *If a provision rule Q satisfies IV, then its restriction $Q_{|N}$ will be RS_N for any $N \in \mathbb{N}$.*

Proof. The result follows immediately from Definition 3.4 using Lemma 3.4.

□

Statement of the informative voting problem

We now are ready to state the refined public goods problem of a utilitarian planner, called the informative voting problem. Note that by the imposition of the IV constraints the planner can be sure that anyone will truthfully report his taste. Hence, at the interim stage 2, he learns the true state p of the economy. Using his prior beliefs on p being uniformly distributed, he is able to asses the ex-ante expected utilitarian welfare.

Definition 3.5 *The following problem is called the informative voting problem, henceforth referred to as P: Maximize expected utilitarian welfare*

$$EW \;\; := \;\; \lambda \int_0^1 \{\bar{v}(p)Q(p) - K(Q(p))\}\, dp$$

by the choice of a provision rule Q, which satisfies the IV constraints as stated in (3.6), accompanied by equal cost sharing, i.e. fulfilling the budget constraint as well as I-IC.

The solution to this problem is denoted by Q^{**}, the induced optimal welfare level by EW^{**}.

For the remainder of this chapter, problem P takes the center stage of our examination. In Section 3.5 we will characterize its solution showing that informative voting matters only in the presence of a second source of heterogeneity that influences, besides taste, the individual's effective valuation of the public good. Finally, in Section 3.6, we explore the deeper relationship between the informative voting problem P and the finite problems P_N.

3.5 The optimal provision rule under IV

In this section we solve the informative voting problem in the sense that we characterize the structure of its possible solutions. As will become clear, this structure depends on the polarization of skill levels. The lower the polarization, the more information aggregation is possible and the more the optimal provision rule Q^{**} under IV resembles the solution Q^* under I-IC only.

3.5.1 The impact of skill heterogeneity

This section investigates the role of skill heterogeneity with respect to the impact our refinement concept has on the optimal provision rule. To this end, we employ some convenient terminology. In line with the idea of informative voting, that each individual has to subscribe to one of the groups of agents with either high or low valuation for the public good, we group together individuals with the same taste parameter. Accordingly, we refer to the extent of taste heterogeneity, i.e. the difference in effective valuations $\theta_H w - \theta_L w$ for a given skill parameter w, as *between-group polarization*. Likewise, we refer to the extent of skill heterogeneity, i.e. the maximum difference in effective valuations $\theta \bar{w} - \theta \underline{w}$ for a given taste parameter θ, as *within-group polarization*.

In order to do comparative statics with respect to within-group polarization, we fix the taste parameters θ_L and θ_H. It will be shown that the within-group polarization determines the extent to which an optimal provision rule aggregates information.[19] In an extreme case of within-group polar-

[19] These results are similar in spirit to those of signalling games in which the extent

ization, any provision rule that satisfies IV is constant, i.e. provides the same amount of public goods in all states of the economy. By contrast, if there is no within-group polarization, then the optimal provision rule Q^* according to the Samuelson rule satisfies IV. However, as soon as there is skill heterogeneity among the agents, the Samuelson rule has to be modified in order to fulfill the constraints of informative voting.

As a byproduct of our considerations, we derive very similar results for the finite problems P_N as well. They will be useful to finally establish the equivalence between the concepts of IV and RS_N in the limit as the sample size N grows (see Section 3.6).

Extreme within-group polarization

The following observation is an immediate consequence of Lemma 3.4 and Lemma 3.6 respectively. It shows that skill heterogeneity may imply that information aggregation via sampling becomes quite impossible.

Corollary 3.1 *Suppose $\theta_H \underline{w} < \theta_L \bar{w}$. Then the following statements hold:*

1. *A provision rule Q_N based on sampling satisfies RS_N if and only if*

$$Q_N(0) = Q_N(1) = \ldots = Q_N(N).$$

2. *A provision rule Q satisfies IV if and only if it is constant:*

$$\forall p \in [0,1] : Q(p) = \bar{Q}.$$

The parameter constellation $\theta_H \underline{w} < \theta_L \bar{w}$ is extreme in the sense that the within-group polarization is that intense that the given level of between group polarization is insufficient to separate the effective valuations in one group from the effective valuations within the other group. Put differently, irrespective of taste parameters, an individual from the upper end of the skill distribution has a higher effective valuation of the public good as compared to an individual from the lower end of the skill distribution. The RS_N or IV constraints respectively require both to prefer a larger provision level in case of a high taste parameter and to prefer a smaller provision level with a low taste. However, under $\theta_H \underline{w} < \theta_L \bar{w}$, if an individual with very low skills wants to have more of the public good, then the same is true for a high ability individual, even, if according to RS_N or IV, this individual should prefer a

of information aggregation depends on the intensity of preference polarization between a receiver and one or several senders. Examples include Crawford and Sobel (1982), Schultz (1996) or Grossman and Helpman (2001, Chapter 4).

lower provision level. Hence, there is no way to aggregate information via a decentralized procedure such as sampling under a RS_N requirement.

In particular, the impossibility of information aggregation holds irrespective of the sample size N. This proves that imposing the requirements of RS_N or IV may heavily restrict the set of implementable allocations.[20] The following Lemma makes this statement more precise and establishes the close connection between the problems P and P_N solving them for the given parameter constellation.

Denote by Q^u the quantity of public goods an *uninformed* planner would provide, i.e. if he could not communicate with the agents at all but was constraint to choose a fix amount according to his prior beliefs irrespective of the true state of the world. Hence, the *uninformed optimum* Q^u is the solution of maximizing

$$\lambda \int_0^1 \{\bar{v}(p)Q - K(Q)\}\, dp$$

by the choice of Q and given by the first order condition

$$\bar{v}(\frac{1}{2}) = K'(Q^u).$$

Lemma 3.8 *If $\theta_H \underline{w} < \theta_L \bar{w}$, then for all $N \in \mathbb{N}$, for all $m \in \{0, 1, \ldots, N\}$, and for all $p \in [0, 1]$*

$$Q_N^{**}(m) = Q^{**}(p) = Q^u.$$

The proof is provided in Appendix 3.8.

Lemma 3.8 completely characterizes what can be achieved in the case of extreme within-group polarization. If $\theta_H \underline{w} < \theta_L \bar{w}$, there is no way to aggregate information such that informative voting holds. The planner cannot do better than an entirely uninformed one. Hence, in what follows, we may assume that $\theta_H \underline{w} \geq \theta_L \bar{w}$.

Moreover, Lemma 3.8 shows that for the given parameter constellation the solutions of the finite problems P_N are identical to the restrictions of Q^{**} and hence converge to the solution of problem P. As will be proven in Section 3.6 the convergence holds for any constellation of parameters. Hence, the requirement of informative voting can be regarded as the limit outcome of robust sampling for growing sample sizes.

---

[20]Recall from Section 3.3, that – irrespective of the properties of the skill distribution – imposing only I-IC yields an optimal provision rule $Q^* : p \mapsto Q^*(p)$ with a continuum of different provision levels, one for each $p \in [0, 1]$.

No within-group polarization

The reasoning above suggests that within-group polarization may cause incentive problems. Indeed, the following proposition shows that heterogeneity in skills is both necessary and sufficient for the optimal utilitarian provision rule Q^*, derived in Proposition 3.2, to be affected by the requirement of IV.

Proposition 3.3 *Q^* satisfies IV if and only if $\underline{w} = \bar{w}$.*
For the restrictions $Q^*_{|N}$ to finite samples the following statements hold:

i) If $\underline{w} = \bar{w}$, then $Q^*_{|N}$ will satisfy RS_N for any $N \in \mathbb{N}$.

ii) If $\underline{w} < \bar{w}$, then there will exist $N_0 \in \mathbb{N}$ such that for all $N \in \mathbb{N}$ with $N_0 \leq N$ the restriction $Q^*_{|N}$ does not satisfy RS_N.

The proof is provided in Appendix 3.8. Proposition 3.3 basically says that, as soon as there is skill heterogeneity among the agents in the economy, the provision of public goods must not follow the Samuelson rule Q^*, if implementation under informative voting is aimed at. Instead, the Samuelson rule for robust public goods provision has to be modified. The next section clarifies how this modification should look like.

3.5.2 A taxonomy of possible solutions

So far, we have characterized the solution to the informative voting problem P for the specific parameter constellations of no as well as extreme within-group polarization. If there is no skill heterogeneity, i.e. $\underline{w} = \bar{w}$, then the IV constraints will impose no further restriction, i.e. $Q^{**} = Q^*$. On the other hand, if there is extreme within-group polarization such that $\theta_H \underline{w} < \theta_L \bar{w}$, information aggregation will become impossible, i.e. $Q^{**} = Q^u$.

We now investigate how the solution Q^{**} might look like for the case of intermediate within-group polarization where parameters are such that $\underline{w} < \bar{w}$ and $\theta_H \underline{w} \geq \theta_L \bar{w}$. To this end, we first discuss which functional forms a solution might possibly have. The latter clarification is required because, in general, a solution to the informative voting problem will not be a continuous function of p, but exhibit jumps and *mass points*. Hence, we cannot rely on standard first order conditions as provided by optimal control theory or the calculus of variations. Instead we are forced to use a more direct approach.

Additional notation

We first introduce some additional notation that will prove useful in the following. Denote by $\bar{Q}_L$ the utility maximizing provision level for an individual with effective valuation $\theta_L\bar{w}$, i.e. the provision level for which

$$\theta_L\bar{w} = K'(\bar{Q}_L).$$

Analogously, denote by $\underline{Q}_H$, the utility maximizing provision level for an individual with effective valuation $\theta_H\underline{w}$, i.e. the provision level for which

$$\theta_H\underline{w} = K'(\underline{Q}_H).$$

Note that for $\underline{w} < \bar{w}$ we have $\underline{w} < \frac{1}{\lambda} < \bar{w}$ and thus $Q^*(0) < \bar{Q}_L$ as well as $\underline{Q}_H < Q^*(1)$.

For any provision level Q with $Q^*(0) \leq Q < \bar{Q}_L$, denote by $\hat{Q}$ the provision level which satisfies $Q < \hat{Q}$ and moreover

$$\theta_L\bar{w}Q - K(Q) \;\; = \;\; \theta_L\bar{w}\hat{Q} - K(\hat{Q}) \; .$$

That is, an individual with effective valuation $\theta_L\bar{w}$ is indifferent between the provision levels Q and $\hat{Q}$. Likewise, for any Q with $\underline{Q}_H < Q \leq Q^*(1)$, denote by $\tilde{Q}$ the provision level with $\tilde{Q} < Q$ and

$$\theta_H\underline{w}Q - K(Q) \;\; = \;\; \theta_H\underline{w}\tilde{Q} - K(\tilde{Q}) \; .$$

Categories for robust solutions

We now will show that a solution to the informative voting problem belongs to one of the following four classes of provision rules:

Provision rules with four mass points. Provision rules in this class are characterized by four mass points Q_4^s, $\hat{Q}_4^s$, $\tilde{Q}_4^l$ and Q_4^l. Q_4^s satisfies $Q_4^s < \bar{Q}_L$ and is linked with $\hat{Q}_4^s$ via a binding IV constraint for an individual with effective valuation $\theta_L\bar{w}$. Likewise, Q_4^l exceeds $\underline{Q}_H$ and is linked with $\tilde{Q}_4^l$ via a binding IV constraint for an individual with effective valuation $\theta_H\underline{w}$. In addition it is required that $\hat{Q}_4^s \leq \tilde{Q}_4^l$. Moreover, there is a range of values of p for which the provision level is equal to $Q^*(p)$, i.e. the provision level that would be chosen in the absence of IV-constraints. Formally Q_4 is defined by

$$Q_4(p) \;\; := \;\; \begin{cases} Q_4^s & \text{for} \;\; 0 \leq p \leq \hat{p} \; , \\ \hat{Q}_4^s & \text{for} \;\; \hat{p} < p < \hat{p}' \; , \\ Q^*(p) & \text{for} \;\; \hat{p}' \leq p \leq \tilde{p}' \; , \\ \tilde{Q}_4^l & \text{for} \;\; \tilde{p}' < p < \tilde{p} \; , \\ Q_4^l & \text{for} \;\; \tilde{p} \leq p \leq 1 \; , \end{cases}$$

where the critical indices are implicitly defined by the following equations:[21]

$$\bar{v}(\hat{p}) = \theta_L \bar{w}\ , \quad Q^*(\hat{p}') = \hat{Q}_4^s\ , \quad Q^*(\tilde{p}') = \tilde{Q}_4^l\ , \quad \bar{v}(\tilde{p}) = \theta_H \underline{w}\ .$$

The situation is illustrated in Figure 3.1.

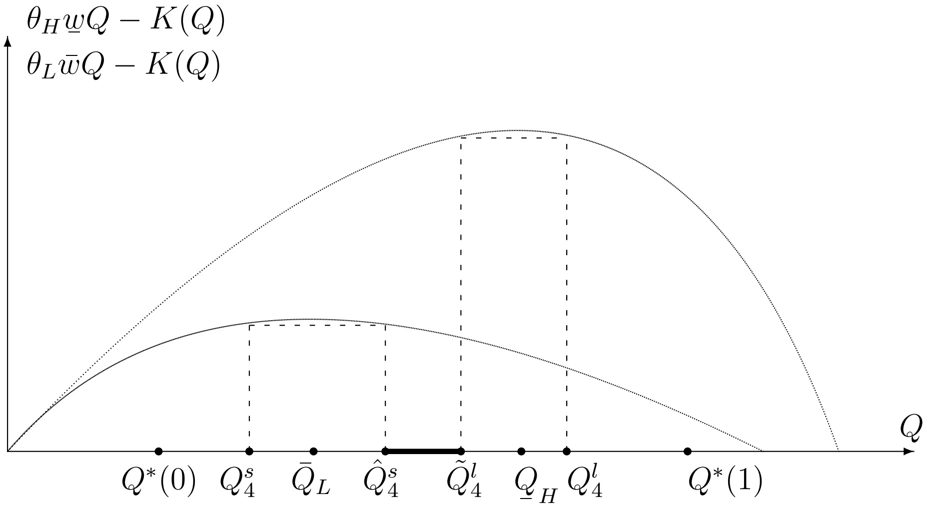

Figure 3.1: Provision rule with four mass points

Provision rules with three mass points. A provision rule with three mass points can be considered as the degenerate case of a provision rule with four mass points, which arises if $\hat{Q}_4^s = \tilde{Q}_4^l$.

This type of provision rule is characterized by three provision levels Q_3^s, Q_3^m and Q_3^l, where Q_3^m is linked via a binding IV constraint for $\theta_L \bar{w}$ with Q_3^s and via a binding IV constraint for $\theta_H \underline{w}$ with Q_3^l, i.e. $Q_3^m = \hat{Q}_3^s = \tilde{Q}_3^l$. Formally Q_3 is defined by

$$Q_3(p) \quad := \quad \begin{cases} Q_3^s & \text{for} \quad 0 \leq p \leq \hat{p}\ , \\ Q_3^m & \text{for} \quad \hat{p} < p < \tilde{p}\ , \\ Q_3^l & \text{for} \quad \tilde{p} \leq p \leq 1\ , \end{cases}$$

where the critical indices $\hat{p}$ and $\tilde{p}$ are defined implicitly by the equations

$$\bar{v}(\hat{p}) = \theta_L \bar{w} \quad \text{and} \quad \bar{v}(\tilde{p}) = \theta_H \underline{w}.$$

The situation is depicted in Figure 3.2.

[21] This already presumes an optimal choice of the critical indices. To see this, note that a utilitarian planner will choose e.g. $\hat{p}$ according to the following criterion: Let $Q(p) = Q_4^s$ if and only if $\bar{v}(p)Q_4^s - K(Q_4^s)$ exceeds $\bar{v}(p)\hat{Q}_4^s - K(\hat{Q}_4^s)$. Given the binding IV constraint which links Q_4^s and $\hat{Q}_4^s$, this is equivalent to $Q(p) = Q_4^s$ if and only if $\bar{v}(p) \leq \theta_L \bar{w}$.

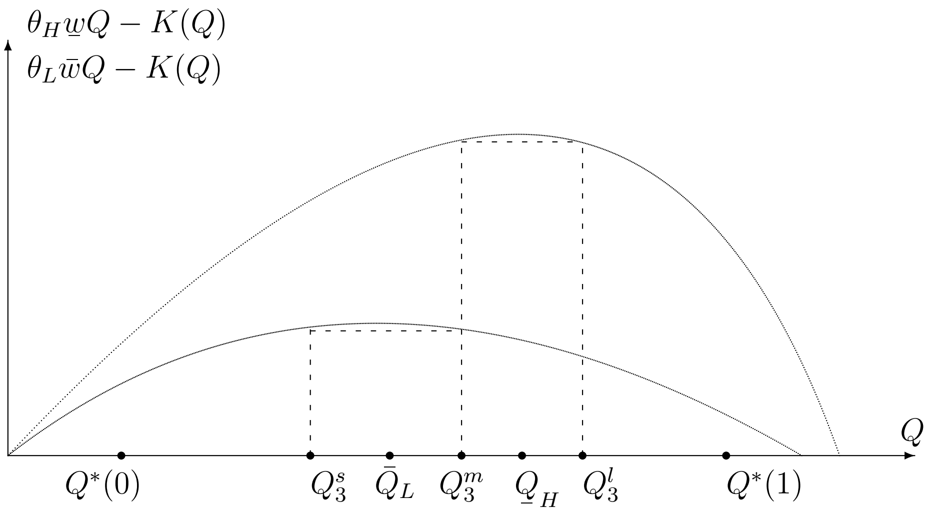

Figure 3.2: Provision rule with three mass points

Provision rules with two mass points. Such a provision rule is characterized by two provision levels Q_2^s and Q_2^l, which satisfy $Q_2^s < \bar{Q}_L$ and $\underline{Q}_H < Q_2^l$. These are possibly linked by a binding IV constraint for an individual with effective valuation $\theta_L \bar{w}$ or via a binding IV constraint for an individual with $\theta_L \bar{w}$.[22] Formally, a provision rule characterized by two mass points Q_2^s and Q_2^l is written as

$$Q_2(p) \quad := \quad \begin{cases} Q_2^s & \text{for} \quad 0 \leq p \leq \hat{p} \ , \\ \\ Q_2^l & \text{for} \quad \hat{p} < p \leq 1 \ , \end{cases}$$

where the critical index $\hat{p}$ is defined implicitly be the equation

$$\bar{v}(\hat{p}) Q_2^s - K(Q_2^s) \quad = \quad \bar{v}(\hat{p}) Q_2^l - K(Q_2^l) \ .$$

In the context of two mass points, the IV constraints can be equivalently represented as the condition that $\hat{p}$ has to be such that

$$\theta_H \underline{w} \quad \geq \quad \bar{v}(\hat{p}) \quad \geq \quad \theta_L \bar{w} \ .$$

The situation of a solution with two mass points is illustrated in Figure 3.3.

[22]Note that as long as $\theta_L \bar{w} < \theta_H \underline{w}$, two different provision levels are linked by at most one of these IV constraints.

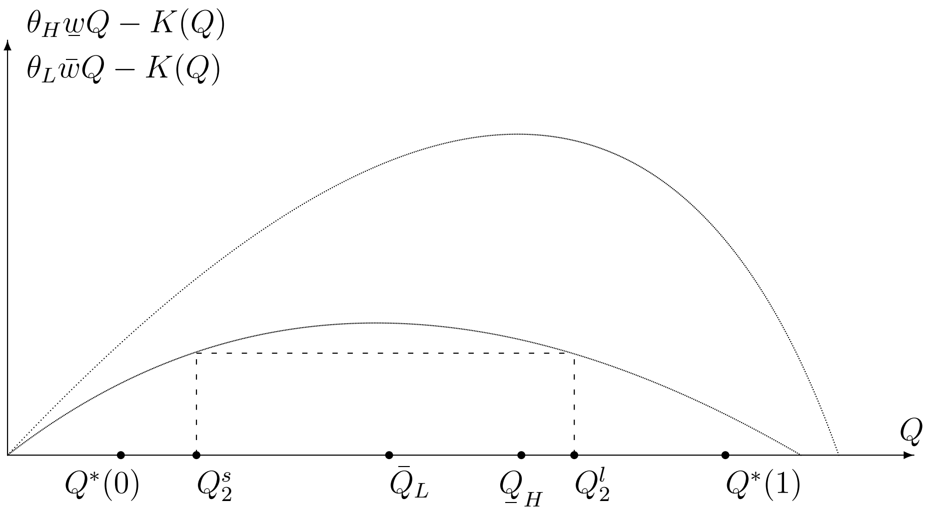

Figure 3.3: Provision rule with two mass points and binding IV for type $\theta_L \bar{w}$

Constant provision rules. Such a provision rule is characterized by a single provision level, namely the one an uninformed planner would choose; formally:

$$Q_1(p) := Q^u \quad \text{for} \quad 0 \leq p \leq 1.$$

We are now able to state the main result of this section: In order to solve the informative voting problem P, one may restrict attention to the classes of provision rules introduced so far.

Proposition 3.4 *Suppose there is skill heterogeneity ($\underline{w} < \bar{w}$). A provision rule which solves the informative voting problem P is either constant or belongs to one of the classes with two, three or four mass points.*

The proof is sketched in Appendix 3.8. It is based on the insight that, under IV constraints, the image of a provision rule contains at most one element smaller than $\bar{Q}_L$ and at most one larger than $\underline{Q}_H$. If there were, to the contrary, two provision levels below $\bar{Q}_L$, then an individual with effective valuation $\theta_L \bar{w}$ would prefer the larger of these two. But IV rules out this possibility. Furthermore, optimality prohibits to have a provision rule whose image lies entirely between $\bar{Q}_L$ and $\underline{Q}_H$. A provision rule with four mass points is superior to such a truncated provision rule. Finally, it is established that an optimal provision rule has exactly one element smaller than $\bar{Q}_L$ and exactly one element larger than $\underline{Q}_H$ as also partial truncations can be excluded. These considerations single out the above candidates.

3.5.3 Solving the informative voting problem

Proposition 3.4 teaches that, in the presence of skill heterogeneity, the optimal provision rule under IV differs strictly from Q^*, the one suggested by the Samuelson rule under I-IC only. Moreover, it gives a hint on how the Samuelson rule has to be modified for robust implementation.

To actually solve the informative voting problem P, one has to compare the welfare levels that can be realized within any of the four classes of provision levels introduced above. In general this requires to solve various separate optimization problems and to rank the resulting welfare levels.

There is however a general intuition, to which class the optimal provision rule belongs, depending on the parameters of the model. Reconsider Figure 3.1 and note that if $\bar{Q}_L$ is close to $Q^*(0)$ and $\underline{Q}_H$ is close to $Q^*(1)$, then a provision rule with four mass points is close to Q^*, which is optimal if IV is not required. This suggests that if the polarization of views on the optimal level of public good provision is relatively mild – in the sense that all individuals with taste parameter θ_L want to have a provision level in a neighborhood of $Q^*(0)$ and all individuals with θ_H want to have a provision level similar to $Q^*(1)$ – then one ends up with a provision rule which exhibits four mass points and hence approximates the optimal I-IC provision rule Q^*.

However, if the difference between $\bar{Q}_L$ and $\underline{Q}_H$ – or equivalently the difference between $\theta_H \underline{w}$ and $\theta_L \bar{w}$ – shrinks, so does the range over which a provision rule with four mass points coincides with Q^*. There will be a critical parameter constellation such that the constraint $\hat{Q}^s_4 \leq \tilde{Q}^l_4$ binds and one ends up with three mass points.

If the polarization increases further, one ends up with $\bar{Q}_L$ being very close to $\underline{Q}_H$. Such a skill distribution implies, for each given taste parameter, a wide range of preferred provision levels. For individuals with taste parameter θ_L, $\bar{Q}_L$ marks the endpoint of this range $[\theta_L \underline{w}, \theta_L \bar{w}]$, for individuals with θ_H, $\underline{Q}_H$ gives the initial point of $[\theta_H \underline{w}, \theta_H \bar{w}]$.

If the difference $\underline{Q}_H - \bar{Q}_L$ has become very small, then a provision rule with three mass points needs to have all three provision levels very close to each other. Hence, there is only very little use of information as a provision rule with three mass points becomes similar to one with $Q(p) = const$, for all p. In such a case a provision rule with only two mass points, which are however to a larger extent differentiated from each other, is superior. Figure 3.4 illustrates this situation showing a provision rule with three mass points and one with two mass points. As $\theta_H \underline{w} - \theta_L \bar{w}$ shrinks, so does $Q^l_3 - Q^s_3$. A provision rule with two mass points eventually becomes more attractive.

Finally, if the difference $\underline{Q}_H - \bar{Q}_L$ has become negative, i.e. we are in the case of extreme polarization where $\theta_H \underline{w} < \theta_L \bar{w}$, information aggregation is

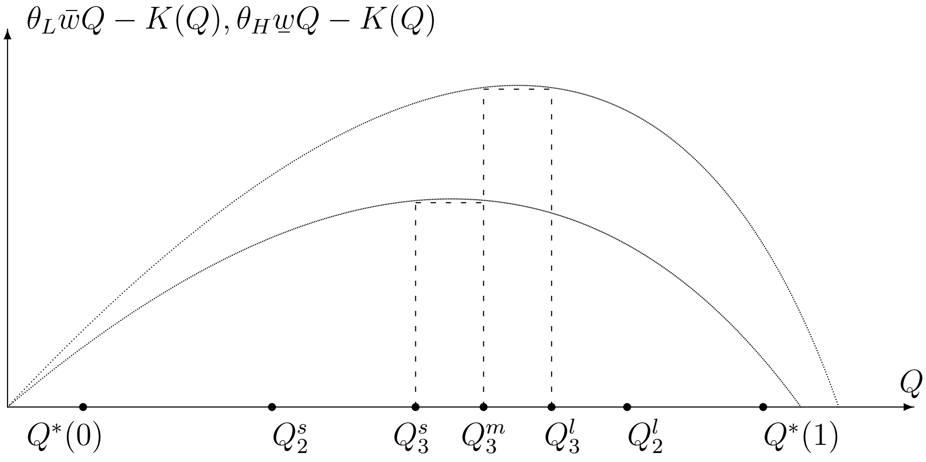

Figure 3.4: Three or two mass points?

quite impossible. As discussed in Lemma 3.8, the planner cannot do better than an uninformed one choosing the constant provision level Q^u.

We refrain from providing a general proof of these intuitive statements. This would require an awkward exercise, which distinguishes a variety of assumptions on the parameters $\theta_L, \theta_H, \bar{w}, \underline{w}$ and λ, i.e. the skill distribution F. We only provide an example which allows to verify the intuition developed above.

Example. Suppose $K(Q) = \frac{1}{2}Q^2$, $\theta_L = 1$, $\theta_H = 3$, and $\lambda = 1$. Let $\underline{w} = 1 - x$ and $\bar{w} = 1 + x$. In this example x is a measure of the welfare burden imposed by the requirement of IV. This welfare burden vanishes as $x \to 0$ implying that $\underline{w} \to \bar{w}$. As $x \to \frac{1}{2}$ one converges to the case with $\theta_L\bar{w} = \theta_H\underline{w}$ which precludes any information aggregation. One may verify that for sufficiently small x, a provision rule with four mass points is optimal. For $x \geq 2^{-\frac{3}{2}}$, an optimal provision rule with four mass points is transformed into the degenerate case with only three mass points. Finally, for x close to $\frac{1}{2}$ a provision rule with only two mass points is superior.

3.6 Robustness to sampling as a foundation for informative voting

In this section we explore the connection between the notions of robustness to sampling of size N on the one hand and informative voting on the other hand. As we will show, the relationship is not only a formal analogy but a substantial equivalence in the following sense: The expected welfare EW_N^{**} induced by the solutions Q_N^{**} of the problems P_N based on sampling of size N converges to the he expected welfare EW^{**} induced by a solution Q^{**} of the informative voting problem P as the sample size N increases. Put differently, in the limit both concepts are equivalent with respect to the induced expected welfare achievable. In addition, if there is a unique solution Q^{**} of problem P, the solutions Q_N^{**} of the problems P_N themselves must 'converge' to this solution Q^{**}, i.e. in the limit both concepts are equivalent with respect to the optimal provision rules induced. This justifies to impose the informative voting constraints if robust implementation is aimed at in the sense that the resulting allocation rule takes into account 'vanishing', i.e. marginal individual influence. Hence, the concept of robustness to sampling can be regarded as a theoretic foundation for the concept of informative voting.

As a byproduct of our analysis we derive a *Condorcet Jury Theorem* for the problem of information aggregation in continuum economies: As soon as information aggregation is possible, i.e. the solution Q^{**} is not constant, more information is better, i.e. for a utilitarian planner, there is no optimal finite sample size.

3.6.1 Large sample properties

We start our analysis with the observation, that the expected welfare EW^{**} induced by a solution Q^{**} of the informative voting problem P serves, for any sample size N, as an upper bound for the expected welfare EW_N^{**} that can be achieved by a solution Q_N^{**} of the problem P_N. To establish this statement, we employ a construction that will also prove useful as we proceed. For given $N \in \mathbb{N}$, consider the following piecewise constant continuation of Q_N^{**}:

$$\begin{aligned} \overline{Q}_N^{**} : [0,1] &\rightarrow \{Q_N^{**}(m)\}_{m=0}^N \quad \text{with} \\ \overline{Q}_N^{**}(p) &:= Q_N^{**}(m) \quad \text{for} \quad \frac{m}{N+1} \leq p < \frac{m+1}{N+1}, \\ \overline{Q}_N^{**}(1) &:= Q_N^{**}(N). \end{aligned} \tag{3.7}$$

The welfare level induced by $\overline{Q}_N^{**}$ is denoted $\overline{EW}_N^{**}$.

Lemma 3.9 *For any $N \in \mathbb{N}$, the following inequalities hold:*

$$EW_N^{**} \leq \overline{EW}_N^{**} \leq EW^{**}.$$

The first inequality will be strict if and only if $Q_N^{**} \not\equiv Q^u$.

The proof is provided in the Appendix 3.8. Lemma 3.9 is the key in order to derive the main results of this section.

Proposition 3.5 *The maximum expected utilitarian welfare under RS_N converges to the maximum expected utilitarian welfare under IV, i.e.*

$$\lim_{N \to \infty} EW_N^{**} = EW^{**}.$$

The proof is also provided in the Appendix 3.8. Proposition 3.5 basically tells that the concepts of robustness to sampling and informative voting are equivalent with respect to achievable expected utilitarian welfare. One can get arbitrarily close to any expected welfare level that can be achieved imposing the IV constraints to the original revelation game as stated in Sequence 3.1 by choosing a sufficiently large sample size N and imposing the RS_N constraints on the perturbed revelation game as stated in 3.2.

In case that there is a *unique* solution Q^{**} of the informative voting problem P, this equivalence is even stronger: For growing sample sizes N, the optimal provision rules Q^{**} under IV and Q_N^{**} under RS_N 'coincide' in the limit. The following corollary makes this statement more precise.

Corollary 3.2 *Suppose there is a unique solution Q^{**} to problem P, and let $\overline{Q}_N^{**}$ be defined as in (3.7). Then, for all $p \in [0, 1]$*

$$\lim_{N \to \infty} \overline{Q}_N^{**}(p) = Q^{**}(p).$$

The proof is provided in the Appendix 3.8.

Hence we have arrived at a theoretic foundation for the imposition of informative voting. The maximum expected welfare under IV and, given *uniqueness*, the optimal provision rule itself can be regarded as the limit outcome of vanishing individual influence that has been granted to the agents by using finite samples for the purpose of information aggregation. In this sense, the optimal provision rule that satisfies the requirement of informative voting is, at the same time, the optimal one which is robust to sampling.

3.6.2 A Condorcet Jury Theorem

It has been shown in Lemma 3.9 that $EW_N^{**} \leq EW^{**}$, i.e. a planner who uses the reports of all individuals for the purpose of information aggregation but is constrained by the requirement of IV will never do worse than a planner who just uses the reports of a finite sample of individuals but is constrained only by the RS_N conditions. As stated in the following Proposition, whenever information aggregation is possible, he can do even better.

Proposition 3.6 *Suppose that Q^u does not solve problem P.*[23] *Then for any $N \in \mathbb{N}$, the following inequality holds:*

$$EW_N^{**} < EW^{**}.$$

The proof is provided in the Appendix 3.8.

As an obvious consequence of this result and its interplay with Proposition 3.5, we can – for the problem of information aggregation in continuum economies with aggregate uncertainty – formulate some version of a *Condorcet Jury Theorem.* I.e. a statement that larger sample sizes are preferred if informative voting of sampled individuals is ensured by appropriate incentive conditions.[24]

Corollary 3.3 (Condorcet Jury Theorem) *Suppose that Q^u does not solve problem P. Then for any given $N \in \mathbb{N}$ there exists $N' \in \mathbb{N}$ with $N < N'$ such that*

$$EW_N^{**} < EW_{N'}^{**}.$$

Put differently, whenever information aggregation is possible, there is no optimal sample size a utilitarian planner could choose for the purpose of information aggregation. Despite the growing number of incentive constraints that must be imposed to ensure truthful reporting behavior, he will always find it preferable to base the decision on public goods provision on a larger number of individual reports. Additional scope for information aggregation generates a strictly larger expected welfare level.

The intuition for this result may be expressed as follows: The higher degree of freedom in choosing different provision levels overcompensates for the disadvantage of being forced to respect a larger number of incentive constraints. This stems from the fact that the planner's estimate of the actual state of the economy becomes more precise and that more differentiated

[23] A sufficient condition for this assumption to hold is, for example, given by $\theta_L \bar{w} \leq \theta_H \underline{w}$ and $\hat{Q}^*(0) \leq Q^*(1)$ and $Q^*(0) \leq \tilde{Q}^*(1)$.

[24] A discussion of this Theorem and related results is provided by Piketty (1999).

information allows for a better adjustment of the provision level to the actual state of the economy. Moreover, this demand for better information is not undermined by the additional RS_N constraints, which come into play with a larger sample. One can think of that in the following way: By choosing an appropriate larger sample the planner can mimic the small sample outcome. Thus there is no welfare burden implied by additional RS_N constraints.

3.7 Concluding remarks

This chapter has investigated the classical free-rider problem in public goods provision in a continuum economy, where individuals differ with respect to their valuation of a public good and with respect to their skills. It has been illustrated that, in the presence of aggregate uncertainty, the standard notion of individual incentive compatibility is not suitable for problems of information aggregation. Thus a refinement of this solution concept, called robustness to sampling, has been introduced and identified with the requirement of informative voting. Its application to the free-rider problem in a quasilinear environment has established the following results: Skill heterogeneity implies that the optimal individually incentive compatible rule for public goods provision is not robust. The extent of information aggregation undertaken by an optimal robust provision rule is inversely related to the polarization of effective valuations which results from the extent of skill heterogeneity.

In this section we briefly sketch some prospects for further research. The following is the most obvious one. Our solution concept of *individual incentive compatibility (I-IC)* cum *robustness to sampling (RS)*, presented in more detail in Chapter 4, is designed to deal with abstract allocation problems which incorporate problems of information aggregation as well as screening problems. The application we discussed – public good provision in a quasilinear continuum economy – is one in which the screening part was trivially reduced to the requirement of equal cost sharing. Hence, this problem became essentially one of pure information aggregation.

We plan to enrich this setting by the introduction of a private consumption good and leisure as arguments in the utility functions of the individuals. This will give rise to a more complex screening problem. In particular, it becomes possible to redistribute consumption goods and to discriminate contributions to the cost of public good provision according to individual characteristics. One thus combines a problem of information aggregation under incentive constraints with a problem of *incentive compatible redistribution*.[25]

[25] This problem of *incentive compatible redistribution* is underlying the theory of optimal

The analysis of this chapter has focussed on large random samples of individuals, which are investigated for the sake of information aggregation. We did not analyze the corresponding *small sample properties*. Further research might clarify, whether a model based on small random samples could be a useful tool for an analysis of decision making in committees. If the answer is yes, this raises a further issue. Our sampling approach takes only those sample characteristics into account that give rise to aggregate uncertainty. In the public goods problem, the quantity of provision depends on the number of individuals with a high valuation. It does not depend on whether these sample members stem from the upper or the lower end of the skill distribution. While, due to a large numbers effect, we believe this to be an innocent assumption in large samples, we do not have an idea of the welfare burden this procedure implies in small samples.

To see why this concern might be of interest, recall that the requirement of robustness to sampling admits an interpretation in terms of informative voting. Hence, an assessment of this welfare burden could be interpreted as a quantification of the welfare costs associated with the use of simple voting rules for social choice.

3.8 Appendix: Formal proofs

Proof of Lemma 3.3.

In the following we will denote by ϕ the density function associated with the planner's prior beliefs on p. Obviously the assumption of a uniform distribution implies that $\phi(p) = 1$ for all $p \in [0, 1]$. The probability distribution associated with the random variable ν, denoting the number of agents with high taste parameters in a sample of size N, is given by:[26]

$$\begin{aligned} pr(\nu = m) &= \int_0^1 pr(\nu = m \mid p)\phi(p)dp \\ &= \int_0^1 \binom{N}{m} p^m (1-p)^{N-m} dp \quad = \quad \frac{1}{N+1} \,. \end{aligned} \tag{3.8}$$

taxation, originating from Mirrlees (1971). See Hellwig (2004a) for a recent contribution.

[26]The following relation is used repeatedly:

$$\int_0^1 p^m (1-p)^{N-m} dp \quad = \quad \frac{m!(N-m)!}{(N+1)!} \,.$$

This is intuitive: With p uniformly distributed, all possible realizations of ν are equally likely. Now suppose that $\nu = m$ and consider the conditional density ϕ_N thereby induced over p:

$$\phi_N(p \mid \nu = m) \quad = \quad \frac{pr(\nu = m \mid p)\phi(p)}{pr(\nu = m)} = (N+1)\binom{N}{m} p^m (1-p)^{N-m} \ .$$

□

Proof of Remark 3.2.

Obviously, by definition of I-IC and RS_N, a RS_N provision rule accompanied by equal cost sharing is truthfully implementable in dominant strategies within the described revelation game.

Now suppose the feasible allocation rule (Q_N, t_N) to be truthfully implementable in dominant strategies within the described revelation game. Thus, the following inequality must hold $\forall i \in I$, $\forall(\theta^i, w^i) \in \Theta \times \Omega$, $\forall(\hat{\theta}^i, \hat{w}^i) \in \Theta \times \Omega$, and $\forall(\theta_{-i}, w_{-i}) \in (\Theta \times \Omega)^{I \setminus \{i\}}$, irrespective wether the agent is a sample member or not:

$$\theta w Q(m) - t_N(m, \theta, w) \geq \theta w Q(\hat{m}) - t_N(\hat{m}, \hat{\theta}, \hat{w}). \tag{3.9}$$

with $m = m(\theta^i, \theta_{-i})$ and $\hat{m} = m(\hat{\theta}^i, \theta_{-i})$.

If the individual is no sample member, she has no influence on the perceived state of the world, i.e. $m = \hat{m}$. As in Section 3.3.2, in combination with the budget balancing requirement this yields equal cost sharing as the only admissible payment scheme: $t_N = K(Q_N)$.

Thus, condition (3.9) simplifies to the requirement that for any individual within the sample condition (3.2) has to be satisfied. But this constraint is just equivalent to RS_N.

□

Proof of Lemma 3.4.

The equivalence of (1.) and (2.) is obvious.

To see that (2.) implies (3.), note first that adding the inequalities in (3.4) yields, for any $m \in \{0, \ldots, N-1\}$ and any $w \in \Omega$,

$$(\theta_H - \theta_L) w [Q_N(m') - Q_N(m)] \quad \geq \quad 0 \ .$$

This establishes i).

Now suppose that $Q(m) < Q(m')$. Then for any $w \in \Omega$, RS_N requires that

$$\frac{K(Q_N(m')) - K(Q_N(m))}{Q_N(m') - Q_N(m)} \quad \geq \quad \theta_L w \ .$$

This property holds for all $w \in W$ if and only if it holds for the largest skill level $\bar{w}$,

$$\frac{K(Q_N(m')) - K(Q_N(m))}{Q_N(m') - Q_N(m)} \geq \theta_L \bar{w} \ .$$

Likewise we derive the requirement

$$\theta_H \underline{w} \geq \frac{K(Q_N(m')) - K(Q_N(m))}{Q_N(m') - Q_N(m)} \ .$$

This establishes ii).

The proof that (3.) implies (2.) is immediate.

□

Proof of Lemma 3.5.

At the *interim* stage 2, when the planner has observed m, he updates his beliefs as stated in Lemma 3.3. Expected welfare on the interim stage 2 is hence given by

$$\begin{aligned}
EW_N^{int}(m) &= \lambda \mathrm{E}[\bar{v}(p)Q_N(m) - K(Q_N(m)) \mid m] \\
&= \lambda \int_0^1 [\bar{v}(p)Q_N(m) - K(Q_N(m))]\phi(p \mid \nu = m)dp \\
&= \lambda(N+1)\binom{N}{m}\Big(\int_0^1 \Big[\Big(p\frac{\theta_H}{\lambda} + (1-p)\frac{\theta_L}{\lambda}\Big) Q_N(m) - K(Q_N(m))\Big] \\
&\qquad p^m(1-p)^{N-m}dp\Big) \\
&= \lambda \left[\frac{m+1}{N+2}\frac{\theta_H}{\lambda} + \frac{N-m+1}{N+2}\frac{\theta_L}{\lambda}\right] Q_N(m) - K(Q_N(m)).
\end{aligned}$$

From an *ex-ante* perspective at stage 1, the outcome m of the sampling procedure is as well the realization of a random variable, which we denote by ν. Taking expectations over m using equation (3.8), expected welfare ex ante equals

$$\begin{aligned}
EW_N &= \sum_{m=0}^{N} EW^{int}(m)pr(\nu = m) \\
&= \lambda \frac{1}{N+1} \sum_{m=0}^{N} \left\{ \bar{v}\left(\frac{m+1}{N+2}\right) Q_N(m) - K(Q_N(m)) \right\}.
\end{aligned}$$

□

Proof of Lemma 3.8.

The equality $Q^{**}(p) \equiv Q^u$ follows immediately from the definition of Q^u and Corollary 3.1.

Using Corollary 3.1, under the given parameter constellation the problem of finding $Q_N^{**} \equiv \bar{Q}_N$ consists in maximizing

$$\lambda \frac{1}{N+1} \sum_{m=0}^{N} \left\{ \bar{v}\left(\frac{m+1}{N+2}\right) \bar{Q} - K(\bar{Q}) \right\}$$

by the choice of $\bar{Q}$ (see equation (3.5)). The first order condition is derived by a simple calculation yielding

$$\bar{v}(\frac{1}{2}) = K'(\bar{Q}_N).$$

Thus $\bar{Q}_N = Q^u$ by the definition of Q^u.

□

Proof of Proposition 3.3.

Before we start with the proof, note that by a revealed preferences argument for the utilitarian planner the following inequalities hold for all $p, p' \in [0, 1]$ with $p < p'$:

$$\begin{aligned} \bar{v}(p)Q^*(p) - K(Q^*(p)) &\geq \bar{v}(p)Q^*(p') - K(Q^*(p')), \\ \bar{v}(p')Q^*(p') - K(Q^*(p')) &\geq \bar{v}(p')Q^*(p) - K(Q^*(p)). \end{aligned}$$

Note that, by Proposition 3.2, Q^* is strictly increasing. Hence, those two inequalities imply that for all $p, p' \in [0, 1]$ with $p < p'$:

$$\bar{v}(p') \geq \frac{K(Q^*(p')) - K(Q^*(p))}{Q^*(p') - Q^*(p)} \geq \bar{v}(p). \tag{3.10}$$

We now prove the claims of Proposition 3.3 within four steps.

1. We first show that $\underline{w} = \bar{w}$ implies that Q^* is IV.

   $\underline{w} = \bar{w} =: \tilde{w}$ implies $\lambda = \frac{1}{\tilde{w}}$. Hence, for all $p, p' \in [0, 1]$ with $p < p'$:

$$\begin{aligned} \theta_H \tilde{w} \geq \bar{v}(p') &= p'\theta_H \tilde{w} + (1 - p')\theta_L \tilde{w}, \\ \theta_L \tilde{w} \leq \bar{v}(p) &= p\theta_H \tilde{w} + (1 - p)\theta_L \tilde{w}. \end{aligned}$$

Combining these observations with (3.10) yields,

$$\theta_H \tilde{w} \;\geq\; \frac{K(Q^*(p')) - K(Q^*(p))}{Q^*(p') - Q^*(p)} \;\geq\; \theta_L \tilde{w}.$$

Since, by Proposition 3.2, Q^* is strictly increasing, it is IV by Lemma 3.6.

2. We have just seen that $\underline{w} = \bar{w}$ implies that Q^* is IV. Hence, by Lemma 3.7, its restriction $Q^*_{|N}$ is RS_N for any $N \in \mathbb{N}$. This proves i).

3. Now we show statement ii).

   Suppose $\underline{w} < \bar{w}$ which implies $\bar{w} > \frac{1}{\lambda} > \underline{w}$. Because of Lemma 3.4, it is sufficient to show that for all sufficiently large N there exist $m', m'' \in \{0, \ldots, N-1\}$ such that

$$\frac{K(Q^*_{|N}(m'+1)) - K(Q^*_{|N}(m'))}{Q^*_{|N}(m'+1) - Q^*_{|N}(m')} \;>\; \theta_H \underline{w} \;,$$

$$\frac{K(Q^*_{|N}(m''+1)) - K(Q^*_{|N}(m''))}{Q^*_{|N}(m''+1) - Q^*_{|N}(m'')} \;<\; \theta_L \bar{w} \;.$$

   Because of the inequalities in (3.10), a sufficient condition for this claim to hold is that for all sufficiently large N there exist $m', m'' \in \{0, \ldots, N-1\}$ such that

$$\bar{v}\left(\frac{m'+1}{N+2}\right) \;>\; \theta_H \underline{w} \quad \text{and} \quad \bar{v}\left(\frac{m''+2}{N+2}\right) \;<\; \theta_L \bar{w} \;.$$

   To see that these conditions are fulfilled, choose, for all N, $m' = N$ and $m'' = 0$ and note that

$$\lim_{N\to\infty} \bar{v}\left(\frac{N+1}{N+2}\right) \;=\; \frac{\theta_H}{\lambda} \;>\; \theta_H \underline{w} \;,$$

$$\lim_{N\to\infty} \bar{v}\left(\frac{2}{N+2}\right) \;=\; \frac{\theta_L}{\lambda} \;<\; \theta_L \bar{w} \;.$$

4. We have just seen that $\underline{w} < \bar{w}$ implies that there exists $N \in \mathbb{N}$ such that $Q^*_{|N}$ is not RS_N. Hence, by Lemma 3.7, Q^* cannot be IV.

□

Proof of Proposition 3.4.

The required arguments are lengthy but not difficult. Hence only a sketch is provided.

If the optimal IV provision rule Q^{**} is constant, then $Q^{**} \equiv Q^u$ by its optimality. As has been shown in Corollary 3.1, there are parameter constellations such that Q^{**} is indeed constant.

Now suppose that Q^{**} is not constant. Thus, by Corollary 3.1, $\theta_L \bar{w} \leq \theta_H \underline{w}$ and hence $Q^*(0) \leq \bar{Q}_L \leq \underline{Q}_H \leq Q^*(1)$. We show within the next three steps, that Q^{**} is a provision rule with either two, three or four mass points.

Step 1. Denote by $\mathcal{V}_Q$ the image of a provision rule Q, i.e. $x \in \mathcal{V}_Q$ if and only if there exists $p \in [0,1]$ with $Q(p) = x$. Under the IV constraints, there exists at most one element $x \in \mathcal{V}_Q$ with $x < \bar{Q}_L$.

To see this, suppose to the contrary that there exist $x, y \in \mathcal{V}_Q$ with $x < y < \bar{Q}_L$. Under IV, as characterized in Lemma 3.6, this implies that there exist p and $p' > p$ with $Q(p) < Q(p') < \bar{Q}_L$. This yields

$$\theta_L \bar{w} Q(p) - K(Q(p)) < \theta_L \bar{w} Q(p') - K(Q(p')) \; ,$$

a contradiction to the IV requirement for an individual with effective valuation $\theta_L \bar{w}$. Analogously one shows that the image of an admissible provision rule contains at most one element x with $x > \underline{Q}_H$.

Step 2. We now show that a provision rule Q for which there exists $y \in \mathcal{V}_Q$ with $y \in [\bar{Q}_L, \underline{Q}_H]$ is a candidate for a solution only if there exist as well $x, z \in \mathcal{V}_Q$ with $x < \bar{Q}_L$ and $\underline{Q}_H < z$.

To this end, we first argue that a provision rule Q for which there exist neither $x \in \mathcal{V}_Q$ with $x < \bar{Q}_L$ nor $z \in \mathcal{V}_Q$ with $z > \underline{Q}_H$ cannot be optimal.

Such a hypothetical provision rule would satisfy $\mathcal{V}_Q \subset [\bar{Q}_L, \underline{Q}_H]$. But this, for such a provision rule to be *optimal*, would imply even $\mathcal{V}_Q = [\bar{Q}_L, \underline{Q}_H]$. However, this would be the degenerate case of a provision rule with four mass points, which results as the limit outcome as Q_4^s converges to $\bar{Q}_L$ and Q_4^l converges to $\underline{Q}_H$. Under a provision rule characterized by four mass points expected welfare EW satisfies the following equation:

$$\begin{aligned}\frac{EW}{\lambda} &= \hat{p}\left[\bar{v}\left(\frac{\hat{p}}{2}\right)Q_4^s - K(Q_4^s)\right] + (\hat{p}'-\hat{p})\left[\bar{v}\left(\frac{\hat{p}'+\hat{p}}{2}\right)\hat{Q}_4^s - K(\hat{Q}_4^s)\right] \\ &\quad + \int_{\hat{p}'}^{\tilde{p}'}\Big\{\bar{v}(p)Q^*(p) - K(Q^*(p))\Big\}dp + (\tilde{p}-\tilde{p}')\Big[\bar{v}\Big(\frac{\tilde{p}'+\tilde{p}}{2}\Big)\tilde{Q}_4^l - K(\tilde{Q}_4^l)\Big] \\ &\quad +(1-\tilde{p})\left[\bar{v}\left(\frac{1+\tilde{p}}{2}\right)Q_4^l - K(Q_4^l)\right] ,\end{aligned}$$

where $\hat{Q}_4^s$ and $\hat{p}'$ are implicit functions of Q_4^s. Similarly, $\tilde{Q}_4^l$ and $\tilde{p}'$ are implicit functions of Q_4^l. Taking these functional relationships into account one may compute the partial derivatives and verify that

$$\lim_{Q_4^s \to \bar{Q}_L} \frac{\partial EW(Q_4^s, Q_4^l)}{\partial Q_4^s} < 0 \quad \text{and} \quad \lim_{Q_4^l \to \underline{Q}_H} \frac{\partial EW(Q_4^s, Q_4^l)}{\partial Q_4^l} > 0.$$

Thus, $Q_4^s = \bar{Q}_L$ and $Q_4^l = \underline{Q}_H$ cannot be optimal.

We now argue in a similar manner that it cannot be optimal to choose a provision rule such that there exist $y, z \in \mathcal{V}_Q$ with $\bar{Q}_L < y < \underline{Q}_H < z$, but such that there does not exist $x \in \mathcal{V}_Q$ with $x < \bar{Q}_L$.

Note that for such a provision rule to be a solution, $y \leq \tilde{z}$ by IV, and hence $\mathcal{V}_Q = [\bar{Q}_L, \tilde{z}] \cup \{z\}$ by optimality and step 1. Again, this is a degenerate case of a provision rule with four mass points, namely the one that results as Q_4^s converges to $\bar{Q}_L$ and $Q_4^l = z$. As above this hypothetical solution can be ruled out as $\lim_{Q_4^s \to \bar{Q}_L} \frac{\partial EW(Q_4^s, Q_4^l)}{\partial Q_4^s} < 0$.

The analogous argument allows to rule out a provision rule such that there exist $x, y \in \mathcal{V}_Q$ with $x < \bar{Q}_L < y < \underline{Q}_H$ but such that there does not exist $z \in \mathcal{V}_Q$ with $z > \underline{Q}_H$.

Step 3. We now claim that a provision rule, for which there exist $x, y \in \mathcal{V}_Q$ with $\bar{Q}_L < x < y < \underline{Q}_H$, is a candidate for a solution only if the whole interval satisfies $[x, y] \subset \mathcal{V}_Q$.

By step 2, there are $a, b \in \mathcal{V}_Q$ with $a < \bar{Q}_L < \underline{Q}_H < b$. However, for such a provision rule to be a solution, $\hat{a} \leq x < y \leq \tilde{b}$ by IV, and hence $[x, y] \subset [\hat{a}, \tilde{b}] \subset \mathcal{V}_Q$ by optimality.

Steps 1-3 imply that an optimal provision rule under IV that is not constant has to be one with two, three or four mass points.

□

Proof of Lemma 3.9.

By definition, $\overline{Q}^{**}_N$ is monotonically increasing in p and inherits the IV property from the fact that $\{Q^{**}_N(m)\}_{m=0}^N$ is RS_N. This is obvious from the characterization of RS_N in Lemma 3.4 and the characterization of IV in Lemma 3.6. Hence, by the optimality of Q^{**} among the provision rules satisfying IV, $\overline{EW}^{**}_N \leq EW^{**}$.

It remains to show that $EW^{**}_N \leq \overline{EW}^{**}_N$. In order to compute $\overline{EW}^{**}_N$, we first collect a number of observations wich are easily verified by the reader.

1. For all $\underline{p}, \bar{p} \in [0,1]$
$$\int\limits_{\underline{p}}^{\bar{p}} \bar{v}(p)dp = (\bar{p}-\underline{p})\bar{v}\left(\frac{\bar{p}+\underline{p}}{2}\right).$$

2. For all $m \in \{0,1,\dots,N\}$
$$\frac{m+\frac{1}{2}}{N+1} = \frac{m+1}{N+2} + \frac{m-\frac{1}{2}N}{(N+1)(N+2)}.$$

3. For all $x, y \in [0,1]$ with $x+y \in [0,1]$
$$\bar{v}(x+y) = \bar{v}(x) + \frac{\theta_H-\theta_L}{\lambda}y.$$

4. Recall that by equation (3.5)
$$EW^{**}_N = \lambda\frac{1}{N+1}\sum_{m=0}^{N}\left\{\bar{v}\left(\frac{m+1}{N+2}\right)Q^{**}_N(m) - K(Q^{**}_N(m))\right\}.$$

Using these equalities one arrives at

$$\begin{aligned}\overline{EW}^{**}_N &= \lambda\int\limits_0^1\left\{\bar{v}(p)\overline{Q}^{**}_N(p) - K(\overline{Q}^{**}_N(p))\right\}dp \\ &= EW^{**}_N + \frac{\theta_H-\theta_L}{(N+1)^2(N+2)}\sum_{m=0}^{N}(m-\frac{1}{2}N)Q^{**}_N(m).\end{aligned}$$

To complete the proof it remains to show that $\sum_{m=0}^{N}(m-\frac{1}{2}N)Q^{**}_N(m) \geq 0$. This expression will equal

$$\sum_{m=0}^{\frac{1}{2}N}(\frac{1}{2}N-m)(Q^{**}_N(N-m)-Q^{**}_N(m)) \tag{3.11}$$

if N is even and

$$\sum_{m=0}^{\frac{N-1}{2}} (\frac{1}{2}N - m)(Q_N^{**}(N-m) - Q_N^{**}(m)) \tag{3.12}$$

if N is odd. However, as Q_N^{**} is increasing, those sums are non-negative. Moreover, they are positive, and hence $EW_N^{**} < \overline{EW}_N^{**}$, if and only if Q_N^{**} is not constant, i.e. $Q_N^{**} \not\equiv Q^u$.

□

Proof of Proposition 3.5.

Let Q^{**} be a solution of problem P and $Q_{|N}^{**}$ its restriction to the domain $\{0, 1, \ldots, N\}$. By Lemma 3.7, $Q_{|N}^{**}$ is RS_N. Denote by $EW_{|N}^{**}$ the expected welfare level induced by $Q_{|N}^{**}$. Then, since Q_N^{**} is optimal among the RS_N provision rules, $EW_{|N}^{**} \leq EW_N^{**}$.

On the other hand,

$$\begin{aligned} EW_{|N}^{**} &= \lambda \frac{1}{N+1} \sum_{m=0}^{N} \left\{ \bar{v}\left(\frac{m+1}{N+2}\right) Q^{**}(\frac{m}{N}) - K(Q^{**}(\frac{m}{N})) \right\} \\ &= \lambda \frac{1}{N+1} \sum_{m=0}^{N} \left\{ \bar{v}\left(\frac{m}{N}\right) Q^{**}(\frac{m}{N}) - K(Q^{**}(\frac{m}{N})) \right\} \\ &\quad + \frac{\theta_H - \theta_L}{N(N+1)(N+2)} \sum_{m=0}^{N} (N-2m) Q^{**}(\frac{m}{N}). \end{aligned}$$

The first term in this sum is a so-called *Riemann sum*[27] for $\bar{v}(p)Q^{**}(p) - K(Q^{**}(p))$ and thus converges to EW^{**} for growing N. The second term in the sum is bounded from above by the expression $\frac{\theta_H - \theta_L}{(N+2)} Q^{**}(1)$, that is vanishing for growing N. Consequently, $\lim_{N\to\infty} EW_{|N}^{**} = EW^{**}$.

Summing up and using Lemma 3.9, the following chain of inequalities must hold:

$$EW^{**} = \lim_{N\to\infty} EW_{|N}^{**} \leq \lim_{N\to\infty} EW_N^{**} \leq EW^{**}.$$

□

[27] See e.g. Heuser (1994, Chapter X).

Proof of Corollary 3.2.

By Lemma 3.9 and Proposition 3.5,

$$\lim_{N\to\infty} EW_N^{**} = \lim_{N\to\infty} \overline{EW}_N^{**} = EW^{**}.$$

Note that $\overline{Q}_N^{**}$ is IV for all $N \in \mathbb{N}$. Thus, the uniqueness of Q^{**} among the provision rules satisfying IV and inducing EW^{**} implies the stated pointwise convergence.

□

Proof of Proposition 3.6.

Denote by EW^u the expected welfare induced by the constant provision rule Q^u an uninformed planner would choose. If Q^u is not a solution to problem P, then $EW^u < EW^{**}$.

If, for $N \in \mathbb{N}$, a solution to problem P_N satisfies $Q_N^{**} \equiv Q^u$ then $EW_N^{**} = EW^u < EW^{**}$. However, if $Q_N^{**} \not\equiv Q^u$ then $EW_N^{**} < EW^{**}$ by Lemma 3.9.

□

Chapter 4

Robustness to sampling

This chapter addresses the problem of information aggregation under aggregate uncertainty at a more abstract level, that is, for a general allocation problem in a continuum economy. The notion of an *individually incentive compatible* or *decentralizable* allocation rule is extended to ensure that an allocation mechanism is not vulnerable to strategic behavior if a problem of information aggregation arises. To this end, we introduce the requirement of *robustness to sampling*. Underlying is the idea, that an allocation mechanism is trustworthy in a large economy, only if it survives giving individuals a marginal impact on the perceived state of the world. Besides that, we introduce a more direct criterion of equilibrium selection called *informative subscription*. It respects that whenever an agent is literally indifferent about the own announcement in a revelation game, he will use his preferences over the composition of the economy to break this indifference. Arguing that the requirements of robustness to sampling and informative subscription are essentially equivalent in the limit, we provide a simple tool to analyze problems of information aggregation under aggregate uncertainty in continuum economies.

4.1 The environment

There is a continuum of individuals identified with the unit interval $I = [0, 1]$ and equipped with a measure μ. An individual $i \in I$ has a utility function U defined over the quantity $Q \in \mathbb{R}_+ := [0, \infty)$ of a non-excludable public good and a vector of private goods $A \in \mathbb{R}^L$. In addition, utility depends on individual characteristics. We distinguish a taste parameter $\theta^i \in \Theta$, $\Theta \subset \mathbb{R}_+$, to formalize heterogeneity regarding valuations of the public good and a productivity or skill parameter $w^i \in \Omega$, $\Omega \subset \mathbb{R}_+$. For brevity, we denote a pair of individual characteristics (θ^i, w^i) by γ^i and the set $\Theta \times \Omega$ by Γ. The utility level U is hence written as

$$U = U(Q, A, \theta^i, w^i) = U(Q, A, \gamma^i) .$$

We will often impose the assumption that the utility function U is *additively separable* in the utility contribution of the public good, depending on the taste parameter θ^i, and the utility contribution of A, depending on the skill parameter w^i.

Assumption 4.1 *The utility function U is additively separable: for any $(\theta, w) \in \Theta \times \Omega$,*

$$U = v(Q, \theta) + u(A, w) .$$

Typically, the skill parameter captures individual differences in the utility loss associated with the need to generate a given income level. This is illustrated by the following examples.

Example 1. A prominent example can be found in the theory of optimal income taxation, in the tradition of Mirrlees (1971). Suppose that assumption 4.1 holds and let $A = (C, Y) \in \mathbb{R}^2$ be a pair specifying a consumption level of private goods C and a level of effective labor supply or income Y. Hence,

$$u = u(C, Y, w) .$$

Typically it is assumed in the literature on optimal income taxation that the function u satisfies the following *single crossing property*,

$$\forall (Y, C) : \quad \frac{\partial}{\partial w} \left(\frac{dC}{dY} \right)_{|du=0} < 0 .$$

Hence, at any point in the Y-C plane the indifference curve of a less skilled individual is steeper. I.e. the lower the skill parameter, the higher is the required compensation for a marginal increase of effective labor supply.

Example 2. As another example, suppose that assumption 4.1 holds and let

$$u = -\frac{t}{w} .$$

A just captures an individual's contribution t to the cost of public good provision. The utility loss associated with a given payment obligation decreases in the skill parameter (see Chapter 3).

The assignment of characteristics to individuals is formalized by an *assignment function* $\gamma_a : I \to \Gamma$ with image denoted by $(\gamma^i)_{i\in I} = (\theta^i, w^i)_{i\in I}$. It is assumed throughout that there is *assignment uncertainty*. I.e. the function γ_a – or equivalently the profile $(\gamma^i)_{i\in I}$ – is not commonly known. Instead, individual i has private information on the parameters γ^i. It is assumed that almost all conceivable assignments are measurable functions. This implies that expressions such as

$$\mu(\{i \mid \theta^i \leq \theta \text{ and } w^i \leq w\}) \quad \text{or} \quad \mu(\{i \mid \theta^i \leq \theta \})$$

are, for almost any resolution of assignment uncertainty γ_a, well defined.

Assumption 4.2 *Each individual $i \in I$ has private information on γ^i. Almost all assignments γ_a are measurable functions.*

In addition to assignment uncertainty, there is aggregate uncertainty which refers to the distribution of individual characteristics in the economy. From an ex ante perspective there are different states of the world. Each such state corresponds to a *cumulative distribution function (cdf)* $D : \Gamma \to [0, 1]$, which lists for each $\gamma = (\theta, w)$ the fraction of individuals with characteristics $\gamma^i \leq \gamma$,

$$D(\gamma) = \mu(\{i \mid \theta^i \leq \theta \text{ and } w^i \leq w\}) \ .$$

Assumption 4.3 *There is aggregate uncertainty, in the sense that the actual distribution of characteristics D in the economy is not commonly known. There is a commonly known feasible set $\mathcal{D}$ of distributions.*

The following information about the distribution of characteristics in the economy is common knowledge. There is *aggregate stability* regarding the marginal distribution of productivity parameters. That is, any feasible distribution $D \in \mathcal{D}$ gives rise to the same marginal cdf F, with $F(w) = \mu(\{i \mid w^i \leq w \})$, of the skill parameter in the economy.

Assumption 4.4 *There is aggregate stability with respect to the productivity parameter. I.e. there is a commonly known marginal cumulative distribution function $F : \Omega \to [0, 1]$.*

In addition, it is commonly known that the marginal skill distribution and the marginal taste distribution are independent. That is, there exists a bijection between the set $\mathcal{D}$ and the set of feasible marginal cdfs of the taste parameter.

Assumption 4.5 *The marginal taste and the marginal skill distribution are independent. For any $D \in \mathcal{D}$ there exists a cdf $G : \Theta \to [0, 1]$ such that*

$$\forall(\theta, w) : \quad D(\theta, w) \quad = \quad F(w)G(\theta) \ .$$

For the analysis below, it proves convenient to work with the assumption that $\Theta = \{\theta_1, \theta_2, \ldots, \theta_m\}$ is a finite ordered set. This implies that any G is equivalently represented by a probability density g. Each such density can be written as a finite list of probabilities $g = (g_1, \ldots, g_m)$, with the understanding that the share of individuals with taste parameter θ_j in state g is given by g_j. Accordingly, the set $\mathcal{G}$ of possible states is represented by the set of possible probability densities,

$$\mathcal{G} := \{v = (v_1, \ldots, v_m) \mid \forall j: \; 0 \leq v_j \leq 1 \text{ and } \sum_{j=1}^{m} v_j = 1\} .$$

Assumption 4.6 *The set of taste parameters $\Theta = \{\theta_1, \theta_2, \ldots, \theta_m\}$ is a finite ordered set, with $\theta_1 < \theta_2 < \ldots < \theta_m$.*

Finally, in every state of the world $g \in \mathcal{G}$, the profile of taste parameters $(\theta^i)_{i \in I}$ is taken to be the realization of a stochastic process $(\tilde{\theta}^i)_{i \in I}$ of independent and identically distributed (iid) random variables.

Assumption 4.7 *For each g, the profile of taste parameters $(\theta^i)_{i \in I}$ satisfies the Law of Large Numbers for a Large Economy:*

i) For each g, and for each $\theta_j \in \Theta$, and for each $i \in I$, the individual probability that $\tilde{\theta}^i$ equals θ_j is given by $g(\theta_j)$.

$$\forall i \in I: \quad g_j = prob(\tilde{\theta}^i = \theta_j \mid g) .$$

This relationship is stated as a conditional probability as it refers to individual randomness conditional on a certain realization of aggregate uncertainty, i.e. conditional on g.

ii) For each g and any $k, l \in I$ the random variables $\tilde{\theta}^k$ and $\tilde{\theta}^l$ are independent.

iii) For each g, ex post – that is after aggregate as well as individual uncertainty have been resolved – the share of individuals with taste parameter equal to θ_j is given by g_j.

$$\forall j: \quad g_j = \mu(\{i \mid \theta^i = \theta_j\}) .$$

Remark 4.1 *If one assumes that the continuum I is the real interval $[0, 1]$ and that μ is the Lebesgue measure, then the assumption that $(\tilde{\theta}^i)_{i \in I}$ is a process of iid random variables creates a measurability problem.*[1] *Al-Najjar*

[1] Details can be found in Judd (1985) and Feldman and Gilles (1985).

(2004) introduces a model of a large economy in which the Law of Large Numbers as stated in assumption 4.7 holds true. In his model, I is an infinite countable set and μ is an appropriate generalization of the counting measure. Our analysis is based on these results.

4.2 Incentive compatible allocation rules

The set of admissible allocation rules is characterized using a mechanism design approach. More precisely, an allocation rule is said to be *implementable* if it is the equilibrium outcome of some revelation game. In addition, we restrict attention to anonymous allocation rules. An *anonymous allocation rule* consists of two mappings, a provision rule for the public good,

$$Q : \mathcal{G} \to \mathbb{R}_+,\ g \mapsto Q(g),$$

and a private goods allocation rule,

$$A : \mathcal{G} \times \Gamma \to \mathbb{R}^L_+,\ (g, \gamma) \mapsto A(g, \gamma)\ .$$

Definition 4.1 *An anonymous allocation rule $[Q, A]$ is said to be individually incentive compatible (I-IC) if*

$$\forall g, \forall \hat{\gamma}, \forall \gamma : \quad U(Q(g), A(g, \gamma), \gamma) \geq U(Q(g), A(g, \hat{\gamma}), \gamma)\ .$$

It is said to be feasible if in any state g, $[Q(g), \{A(g, \gamma)\}_{\gamma \in \Gamma}]$ belongs to the economy's production set.

Attention is henceforth restricted to allocation rules which are I-IC and feasible. The requirement of incentive compatibility is due to the fact that individuals have private information on their characteristics. This restricts the possibilities for a differential treatment of individuals by an anonymous allocation rule.

Note that the I-IC conditions are stated for a given g. This reflects that, in a large economy, no single individual has a direct impact on the distribution of announced characteristics. I.e. there is no impact on the state of the world as perceived by the mechanism designer. In combination with the postulate of anonymity this implies in particular, that no single individual has an impact on public good provision.

An instructive alternative characterization of I-IC allocation rules has been provided by Hammond (1979). He calls an allocation rule $[Q, A]$ *decentralizable* if there exists a collection of budget sets $\{B(g)\}_{g \in \mathcal{G}}$ such that $\forall \gamma$, $\forall g$: $A(g, \gamma) \in \text{argmax}_{X \in B(g)} U[Q(g), X, \gamma]$. Hammond (1979) shows that an allocation rule is decentralizable if and only if it is I-IC.

Example 3. The equivalence between individual incentive compatibility and decentralizability is used intensively in the theory of optimal income taxation. To illustrate this, consider again the setting of Example 1 and let $A = (C, Y)$. An allocation rule is thus represented by a collection of mappings $[Q, C, Y]$. Now define an *income tax* as follows: $[Q, C, Y]$ is called an income tax, if there exists a function $T : \mathcal{G} \times \mathbb{R}_+ \rightarrow \mathbb{R}$ such that for all g and for all γ: Consumption equals after tax income:

$$C(g, \gamma) \;\; = \;\; Y(g, \gamma) - T(g, Y(g, \gamma)) \; ,$$

individuals choose a utility maximizing level of income subject to the given income tax schedule,

$$Y(g, \gamma) \;\; \in \;\; \text{argmax}_Y \;\; U(Q(g), Y - T(g, Y), Y, \gamma) \; ,$$

and the public sector budget constraint is satisfied,

$$\textstyle\int_\Gamma T(g, Y(g, \gamma)) dD \;\; = \;\; K(Q(g)) \; ,$$

where K is a cost function, which captures the cost of public good provision. In this setting, the results in Hammond (1979) yield the observation that an allocation rule is I-IC and feasible if and only if it is an income tax.[2]

The I-IC constraints ensure that an individual is willing to reveal the own characteristics for a given cross-section distribution of characteristics. However, under aggregate uncertainty, an incentive compatible allocation rule is used in addition for the purpose of information aggregation as the mechanism designer has to deduce the actual distribution of characteristics from the profile of individual announcements. To make this more explicit, it is instructive to think of the revelation game in the following sequential manner:

Sequence 4.1

1. *The allocation mechanism specifies a collection of choice sets $\{B(g)\}_{g \in \mathcal{G}}$ and of public good provision levels $\{Q(g)\}_{g \in \mathcal{G}}$. I.e. there is a distinct level of public good provision $Q(g)$ and a distinct choice set $B(g)$ for each possible state g of the economy.*

2. *The mechanism collects all individual data and uses this information to deduce the state of the world g.*

[2]A formal proof can be found in Guesnerie (1995).

3. *The quantity $Q(g)$ of the public good is provided. Individuals maximize utility by a decentralized choice from the set $B(g)$ and behave in the intended manner due to I-IC.*

4.3 The problem of multiple equilibria

The possibility to use an I-IC allocation rule for the purpose of information aggregation under aggregate uncertainty, relies on the fact that, in a large economy, individuals have no chance to affect directly the mechanism designer's perception of the state of the world.

As the discussion of the specific allocation problem in Chapter 3 has shown, it could well be the case that individuals are entirely indifferent regarding their treatment for a given state of the world – suppose $A(g, \gamma)$ is, for given g, just a constant and hence independent of individual characteristics. Consequently, the utility level an individual realizes depends only on the mechanism designer's state perception. As we argued before, in such a situation the implementability of an allocation rule can be questioned. Individuals might apply the following reasoning: They want to avoid a perceived state which they do not like and hence feel that, instead of revealing themselves, they should contribute to a different state perception.

The present section serves to make this intuitive reasoning more precise. We will show that the underlying problem which gives rise to this 'temptation to deviate' is one of multiple equilibria. This is most easily demonstrated if preferences satisfy the separability assumption 4.1. In this case I-IC holds *only if* each individual possesses multiple best responses.

Lemma 4.1 *Suppose assumption 4.1 holds. An anonymous allocation rule is I-IC if and only if it satisfies the following properties:*

i) *The no discrimination of taste in terms of utility (NDT-U) property:*

$$\forall g, \forall w, \forall \theta, \forall \theta' : \quad u(A(g, \theta, w), w) = u(A(g, \theta', w), w) \ .$$

ii) *The individual revelation of productivity (I-RP) property:*

$$\forall g, \forall \theta, \forall w, \forall w' : \quad u(A(g, \theta, w), w) \geq u(A(g, \theta, w'), w) \ .$$

Proof. To prove the only if-part note that, because preferences satisfy assumption 4.1, the NDT-U property is an implication of I-IC. Obviously I-RP is also an implication of I-IC. To prove the if-part, suppose an allocation rule, such that the NDT-U and the I-RP property hold, is not I-IC. Then there exist (θ, w) and $(\hat{\theta}, \hat{w})$ and g such that $u(A(g, \theta, w), w) < u(A(g, \hat{\theta}, \hat{w}), w)$. Using NDT-U and I-RP one has: $u(A(g, \hat{\theta}, \hat{w}), w) = u(A(g, \theta, \hat{w}), w) \leq u(A(g, \theta, w), w)$. Hence, a contradiction.

□

The lemma follows from the fact that individuals take the state g and hence the level of public good provision as given. Due to the separability assumption, the utility contribution of the public good vanishes from individual incentive compatibility constraints. In particular, this implies that I-IC conditions become independent of taste parameters. Consequently, an I-IC allocation mechanism can use only individual differences in productivity as a screening device and leaves all individuals indifferent regarding possible taste announcements.

To make the implications of this observation for the multiplicity of best responses in the revelation game apparent, the following notation is introduced. For an individual with characteristics γ, denote by $b[g \mid \gamma]$ the *set of best responses* in the revelation game if the distribution of announcements is given by g,

$$b[g \mid \gamma] \ := \ \text{argmax}_{\tilde{\gamma}} \, U(Q(g), A(g, \tilde{\gamma}), \gamma) \ .$$

Making use of this notation, the I-IC condition is equivalently written as $\forall g, \forall \gamma$: $\gamma \in b[g \mid \gamma]$. Under the NDT-U property, however, $b[g \mid \gamma]$ is not a singleton. For an individual with characteristics $\gamma = (\theta, w)$ any announcement $(\hat{\theta}, \hat{w})$ with $\hat{w} = w$ belongs to $b[g \mid \gamma]$. Put differently, with separable preferences, individuals are willing to announce any taste parameter.

Recall that by assumption 4.5 aggregate uncertainty is entirely due to the unknown marginal distribution of taste parameters. The implementability of an allocation rule hence requires that all individuals reveal their taste parameter truthfully and that the true distribution of characteristics can be deduced from the profile of taste announcements. Consequently, assuming that individual incentive compatibility and feasibility are sufficient conditions for implementability is equivalent to the assumption that individual indifference among all taste announcements is always broken in favor of the truth.

We believe that this postulate is not particularly plausible based on the observation, that individuals are not indifferent with respect to the mechanism designer's state perception. They just have no direct influence on it. In case of being literally indifferent among all conceivable taste announcements, individuals might break indifference as well such that they indirectly contribute to a more favorable announced distribution of taste parameters.

This is the perspective that we want to adopt in the following. We say that an allocation rule is robustly implementable only if it survives breaking indifference in the way just outlined. The refinement which we discuss in the next subsections serves to achieve this notion of stability.

4.4 Finite samples

As a preliminary step for the definition of our refinement concept, in this section we consider the following situation: the mechanism designer *first* collects data on taste parameters from a finite random sample S_N of individuals with sample size N. He observes the distribution of characteristics in the sample and updates his prior beliefs about the actual state of the economy. These updated beliefs are interpreted as the perceived state of the world. Based on these beliefs, the mechanism designer decides in a *second* step on a level of public good provision Q and on the menu of private goods B which is offered to the individuals in the economy. *Finally*, individuals enjoy the public good and choose their utility maximizing element from the menu B.

This allows to get rid of the multiple equilibrium problem in the following manner. It is assumed that the mechanism designer's estimation procedure is commonly known. This implies that all sampled individuals can predict how their own statement on their preference parameter affects the provision level Q and the menu B. The refinement idea now works as follows: We require that sample members make a taste announcement which is consistent with their optimizing behavior at the final stage. Under this constraint, sample members are free to choose the taste announcement which affects the mechanism designer's state perception in the most favorable way. As will become clear, this implies that, whenever individuals have a multiplicity of best options at the final stage, indifference is broken such that their announcement contributes to the most preferred state perception.

As a result of these considerations, allocation rules are made more robust as they do not depend any more on a fragile way of breaking individual indifference. However, this comes at a cost. In order to ensure that sampled individuals are willing to reveal their taste parameter, even if they are granted some influence on the mechanism designer's state perception, an additional set of incentive conditions comes into play. As will become clear, I-IC and feasibility are no more sufficient conditions for implementability.

4.4.1 Information aggregation based on sampling

In this paragraph we describe how a certain state perception of the mechanism designer comes about, if the actual state of the economy can not be deduced from the whole profile of announcements $(\hat{\gamma}_i)_{i \in I}$ but instead a finite random sample of taste parameters is evaluated.

Recall that each state of the economy can be represented by a vector $g \in \mathcal{G}$. We assume that the mechanism designer has some prior beliefs on the likelihood of different states. That is, he attaches probability weights to the elements of $\mathcal{G}$. These prior beliefs are formalized with a density function $\phi : \mathcal{G} \to \mathbb{R}_+$, defined on the set of possible states of the world.

The mechanism designer draws a random sample S_N of N individuals and evaluates the profile of taste parameters which is denoted by $\theta_N := \{\theta^1, \theta^2, \ldots, \theta^N\}$, where θ^j is the taste parameter of the sample member at position j, $j \in \{1, \ldots N\}$. A sample realization gives rise to an empirical sample density, denoted by

$$g_N = (g_{N1}, \ldots, g_{Nm}) \text{ , with } \quad g_{Nk} = \frac{\#\{j \in S_N \mid \theta^j = \theta_k\}}{N} \ .$$

The set of possible sample realizations $\mathcal{G}_N$ is represented by the set of m-dimensional vectors $v = (v_1, \ldots, v_m)$ such that each component v_k is a multiple of $1/N$ and such that all components sum up to 1,

$$\mathcal{G}_N := \{v \mid \forall k \in \{1, \ldots, m\}, \exists x \in \{0, 1, \ldots, N\} : \ v_k = \frac{x}{N} \text{ and } \sum_{k=1}^{m} v_k = 1\} \ .$$

Note that $\mathcal{G}_N$ is a finite set. A typical element is denoted by g_N or, if the dependence on a specific underlying profile of taste parameters deserves emphasis, by $g_N(\theta_N)$.

After observing a certain sample distribution of characteristics g_N, the planner updates his prior beliefs on the likelihood of different states. These updated beliefs constitute the perceived state of the economy based on a finite sample of size N. Formally, a perceived state is a conditional density function, which is derived using Bayes' rule,

$$\phi_N(\cdot \mid g_N) : \mathcal{G} \to \mathbb{R}_+ \text{ , with } \quad \phi_N(g \mid g_N) = \frac{\text{prob}(g_N \mid g)\phi(g)}{\int_{\mathcal{G}} \text{prob}(g_N \mid g)\phi(g)dg} \ .$$

We use in the following the notation ϕ_N to denote a typical posterior density of the mechanism designer. If we want to emphasize the dependence on a certain sample distribution we write $\phi_N(\cdot \mid g_N)$ or $\phi_N(\cdot \mid g_N(\theta_N))$.

Note that there exists a bijection between the set $\mathcal{G}_N$ of possible sample distributions and the set of posterior density functions. With a slight abuse of notation, we will write $\phi_N \in \mathcal{G}_N$ to indicate that there exists a sample realization g_N, which gives rise to the state perception ϕ_N.

Remark 4.2 *One can derive an explicit expression for prob($\hat{g}_N \mid g$). Making repeated use of conditional probabilities and the binomial probability distrib-*

ution one verifies that

$$prob(\hat{g}_N = v \mid g) \quad = \quad \prod_{k=1}^{m} \binom{N - N\sum_{j=k+1}^{m} v_j}{Nv_k} g_k^{Nv_k} (1 - g_k)^{(N - N\sum_{j=k}^{m} v_j)} ,$$

with the convention that $\sum_{j=m+1}^{m} v_j = 0$.

4.4.2 Allocation rules based on sampling

Having clarified how the mechanism designer uses the sample information to update his prior beliefs on the likelihood of different states of the economy, we now define the notion of an *allocation rule based on sampling of size* N. It consists of a provision rule for public goods,

$$Q_N : \mathcal{G}_N \to \mathbb{R}_+, \ \phi_N \mapsto Q_N(\phi_N) ,$$

and a private goods allocation rule

$$A_N : \mathcal{G}_N \times \Gamma \to \mathbb{R}_+^L, \ (\phi_N, \gamma) \mapsto A_N(\phi_N, \gamma) .$$

Both functions depend on the perceived state of the world ϕ_N, as opposed to the actual distribution g of characteristics in the economy. I.e. based on the sample observation g_N the mechanism designer forms beliefs ϕ_N and decides on public good provision and the offered menu of private goods.

We require that the pair $[Q_N, A_N]$ is *feasible*, i.e. for each state perception ϕ_N and for each actual state g the collection $\{Q_N(\phi_N), \{A_N(\phi_N, \gamma)\}_{\gamma \in \Gamma}\}$ lies in the economy's production set. In addition, we consider only pairs $[Q_N, A_N]$ which permit a decentralization via a collection of budget sets $\{B(\phi_N)\}_{\phi_N \in \mathcal{G}}$. Equivalently, again I-IC is required, i.e.

$$\forall \phi_N \forall \gamma \forall \hat{\gamma} : \quad U(Q_N(\phi_N), A_N(\phi_N, \gamma), \gamma) \geq U(Q_N(\phi_N), A_N(\phi_N, \hat{\gamma}), \gamma) .$$

To sum up, whenever we consider allocation rules based on sampling we refer to the following sequence of actions.

Sequence 4.2

1. *The allocation mechanism specifies a collection of choice sets* $\{B(\phi_N)\}_{\phi_N \in \mathcal{G}}$ *and of public good provision levels* $\{Q(\phi_N)\}_{\phi_N \in \mathcal{G}}$. *I.e. there is a distinct level of public good provision* $Q(\phi_N)$ *and a distinct choice set* $B(\phi_N)$ *for each possible state perception* ϕ_N.

2. *The mechanism collects individual data from sample members and uses this information to arrive at the actual state perception* ϕ_N. *As a consequence* ϕ_N *is via* g_N *a function of* θ_N, *the profile of taste parameters in the sample.*

3. *The quantity $Q(\phi_N)$ of the public good is provided. Individuals maximize utility by a decentralized choice from the set $B(\phi_N)$ and behave in the intended manner due to I-IC.*

4.4.3 Robust implementation based on sampling

Similarly as in the previous section, the set of optimal announcements at the final allocation stage is denoted by $b_N[\phi_N \mid \gamma]$ and formally defined as

$$b_N[\phi_N \mid \gamma] \;\; := \;\; \text{argmax}_{\tilde{\gamma}} \, U(Q_N(\phi_N), A_N(\phi_N, \tilde{\gamma}), \gamma) \; .$$

As has been clarified before, I-IC is equivalent to the requirement that $\forall \phi_N$, $\gamma \in b_N[\phi_N \mid \gamma]$. In general, however, the set $b_N[\phi_N \mid \gamma]$ is not a singleton. E.g. as discussed in the previous section, if U satisfies additive separability, then I-IC holds only if any announcement $(\hat{\theta}, \hat{w})$ with $\hat{w} = w$ belongs to $b_N[\phi_N \mid \gamma]$.

To eliminate the possibility that individuals break this indifference in such a way as to contribute indirectly to a false state perception, we now introduce a condition, which we call *robustness to sampling of size N (RS_N)*.

Some additional notation is needed. Let θ_{N-i} be the vector of characteristics for all sample members with exception of individual i. For a sample member $i \in S_N$ with characteristics $\gamma = (\theta, w)$, denote by

$$rb_N[\theta_{N-i} \mid (\theta, w)] := \Big\{ (\hat{\theta}, \hat{w}) \mid (\hat{\theta}, \hat{w}) \in b_N[\phi_N(\cdot \mid g_N(\theta_{N-1}, \hat{\theta})) \mid (\theta, w)] \Big\}$$

the *refined best response set*, i.e. the set of best responses at the final stage which are consistent with reporting $\hat{\theta}$ in case of being a sample member at stage 2. Finally, denote by $\Phi_N(\theta_{N-i} \mid (\theta, w))$ the set of state perceptions a sample member i with characteristics (θ, w) has at his disposal if he is constrained not do deviate from this refined best response set,

$$\Phi_N(\theta_{N-i} \mid (\theta, w)) := \Big\{ \phi_N(\cdot \mid g(\theta_{N-i}, \hat{\theta})) \mid \exists \hat{w} : (\hat{\theta}, \hat{w}) \in rb_N[\theta_{N-i} \mid (\theta, w)] \Big\}.$$

A typical element of $\Phi_N(\theta_{N-i} \mid (\theta, w))$ is henceforth by $\hat{\phi}_N$. The element which results if sample member i reveals his taste parameter truthfully, is denoted ϕ_N.

Definition 4.2 *An I-IC and feasible pair $[Q_N, A_N]$, is said to be robust to sampling of size N (RS_N) if*

$$\forall i \in S_N, \forall \theta_{N-i}, \forall (\theta, w), \text{ and } \forall \hat{\phi}_N \in \Phi_N(\theta_{N-i} \mid (\theta, w)) :$$

$$U(Q_N(\phi_N), A_N(\phi_N, \theta, w), \theta, w) \geq U(Q_N(\hat{\phi}_N), A_N(\hat{\phi}_N, \theta, w), \theta, w) \; .$$

Remark 4.3 *The RS_N property can alternatively be written as follows:*

$$\forall i \in S_N, \forall \theta_{N-i}, \forall (\theta, w), \text{ and } \forall \hat{\phi}_N \in \Phi_N(\theta_{N-i} \mid (\theta, w)) :$$

$$U(Q_N(\phi_N), A_N(\phi_N, \theta, w), \theta, w) \geq U(Q_N(\hat{\phi}_N), A_N(\hat{\phi}_N, \hat{\theta}, \hat{w}), \theta, w) \ ,$$

with the understanding that $(\hat{\theta}, \hat{w}) \in rb_N[\theta_{N-i} \mid (\theta, w)]$ is the announcement which induces the state perception $\hat{\phi}_N$. To see that this condition is equivalent, note that by I-IC, for all (θ, w) and for all θ_{N-i}, $(\theta, w) \in rb_N[\theta_{N-i} \mid (\theta, w)]$. Hence, both (θ, w) and $(\hat{\theta}, \hat{w})$ belong to the best response set $rb_N[\theta_{N-i} \mid (\theta, w)]$, and we have:

$$U(Q_N(\hat{\phi}_N), A_N(\hat{\phi}_N, \theta, w), \theta, w) = U(Q_N(\hat{\phi}_N), A_N(\hat{\phi}_N, \hat{\theta}, \hat{w}), \theta, w) \ .$$

The RS_N-constraints are to be read as follows: Suppose an individual happens to be a sample member and is able to foresee how his or her announcement affects the estimate ϕ_N and hence the level of public good provision $Q(\phi_N)$ as well as his own choice from the set $B(\phi_N) := \{A(\phi_N, \gamma)\}_{\gamma \in \Gamma}$. The individual is constraint to give a best response at the final stage. Furthermore, the individual is required to behave at the sampling stage consistent with the best response he finally plans to give. These consistency requirements define the refined best response set $rb_N[\cdot]$. That is, we apply a *lexicographic reasoning*: Only if optimal behavior at the final stage is not uniquely determined we grant a sample member a degree of freedom. He may deviate from the truth at the sampling stage only if this deviation is compatible with some alternative best response taking the induced perceived state of the world as given.

The requirement of RS_N serves as a refinement of I-IC. It restricts the admissible ways to break the indifference individuals may have at the final stage. The refinement says that implementability of an allocation rule is incompatible with the existence of false reports, which are consistent with optimal individual behavior ex post and – conditional on being a sample member – strictly superior. Loosely speaking, one may not break indifference in favor of truth-telling if this counters individual incentives as soon as the individual has some – possibly arbitrarily small – influence on the perception of the state of the world.[3]

[3]To see that those statements are formally sound, note that, if $b_N[\cdot]$ happens to be a singleton for all ϕ_N and hence contains only the truth, then the RS_N-property is implied by I-IC.

4.4.4 RS$_N$ under separable preferences

Under the assumption that preferences are additively separable, the requirement of RS$_N$ can be simplified. Recall that under this specification I-IC holds only if individuals are willing to announce *any* taste parameter. As we will show in the following, the RS$_N$ requirement holds if and only if sampled individuals are willing to announce the true taste parameter.

To see this, it proves convenient to introduce a partition of the set $rb_N[\theta_{N-i} \mid (\theta, w)]$. We refer to the collection of those elements for which $\hat{w} = w$, in the following as $rb^*_N[\theta_{N-i} \mid (\theta, w)]$. Note that under the separability assumption 4.1, $rb^*_N[\theta_{N-i} \mid (\theta, w)]$ is equal to $\{(\hat{\theta}, \hat{w}) \mid \hat{w} = w\}$. The set of elements of $rb_N[\theta_{N-i} \mid (\theta, w)]$, for which $\hat{w} \neq w$, is denoted by $rb'_N[\theta_{N-i} \mid (\theta, w)]$.

Lemma 4.2 *Suppose the separability assumption 4.1 holds. The restriction of the RS$_N$-requirement to the set $rb^*_N[\theta_{N-i} \mid (\theta, w)]$ is sufficient for overall RS$_N$.*

Proof. The proof proceeds by contradiction. Suppose the RS$_N$ requirement holds if applied to the set $rb^*_N[\theta_{N-i} \mid (\theta, w)]$, but $\exists \theta_{N-i}$, $\exists (\theta, w)$, $\exists (\hat{\theta}, \hat{w}) \in rb'_N[\theta_{N-i} \mid (\theta, w)]$ such that the induced state perception $\hat{\phi}_N$ satisfies

$$v(Q(\phi_N), \theta) + u(A(\phi_N, \theta, w), w) < v(Q(\hat{\phi}_N), \theta) + u(A(\hat{\phi}_N, \theta, w), w) \ .$$

However, the state perception $\hat{\phi}_N$ is also induced if the report $(\hat{\theta}, \hat{w})$ is replaced by $(\hat{\theta}, w) \in rb^*_N[\theta_{N-i} \mid (\theta, w)]$. Hence a contradiction to the assumption that the RS$_N$ requirement is fulfilled for reports in the set $rb^*_N[\theta_{N-i} \mid (\theta, w)]$.

□

Denote by $\Phi^*_N(\theta_{N-i})$ the set of state perceptions sample member i can induce via some announcement from $\{(\hat{\theta}, \hat{w}) \mid \hat{w} = w\}$,

$$\Phi^*_N(\theta_{N-i}) \ := \ \{\phi_N(\cdot \mid g_N(\theta_{N-i}, \hat{\theta}))\}_{\hat{\theta} \in \Theta} \ .$$

Again, denote a typical element of Φ^*_N by $\hat{\phi}_N$. Denote by ϕ_N the element which results if sample member i has taste parameter θ and reports this parameter truthfully.

Corollary 4.1 *With separable preferences RS$_N$ holds if and only if*

$$\forall \theta_{N-i}, \forall (\theta, w), \forall \hat{\phi}_N \in \Phi^*_N(\theta_{N-i}) :$$

$$v(Q(\phi_N), \theta) + u(A(\phi_N, \theta, w), w) \geq v(Q(\hat{\phi}_N), \theta) + u(A(\hat{\phi}_N, \theta, w), w) \ .$$

We have thus arrived at a separation of incentive problems. The RS_N-constraints deal with a revelation of those characteristics for which aggregate uncertainty prevails – assuming that the assignment uncertainty regarding individual earning ability has been resolved. The I-IC constraints by contrast address the revelation of skill parameters, which satisfy aggregate stability.

4.4.5 Mechanism design based on finite samples

In this paragraph we define a mechanism design problem under the assumption that only a finite random sample of individuals is used for the purpose of information aggregation.

The investigation of this class of mechanism design problems is not the main focus of this paper's analysis. However, as will become clear, the definition of an optimal allocation rule with the RS_N property is an ingredient for a definition of optimal allocation rules which are *robust to sampling*, without referring to a particular sample size N.

We consider the mechanism design problem of a utilitarian planner who has prior beliefs ϕ on the distribution of taste parameters in the economy. His object of choice is an implementable allocation rule based on sampling of size N. I.e. the mechanism designer chooses a pair of functions $[Q_N, A_N]$ with the properties of feasibility, I-IC and RS_N. The planner's objective is to maximize expected utilitarian welfare from an ex ante perspective.

To complete the statement of the mechanism design problem, this objective function is derived as follows: Ex post, under a given state perception ϕ_N and for a given state of the economy g, utilitarian welfare is given by

$$W_N(\phi_N, g) := \int_\Theta \int_W U(Q_N(\phi_N), A_N(\phi_N, \theta, w)), \theta, w) f(w) dw \; g(\theta) d\theta \; .$$

From the interim perspective, the state of the world is not known, but the planner has already observed the sample realization g_N and has derived at the posterior beliefs ϕ_N. From the interim perspective expected utilitarian welfare equals,

$$EW_N^{int}(\phi_N) \;\; = \;\; \int_{\mathcal{G}} W_N(\phi_N, g) \phi_N(g \mid g_N) dg \; .$$

The planners's objective function is given by expected utilitarian welfare from an ex ante perspective. Accordingly, the planner has not yet observed a certain sample realization and perceives ϕ_N as a random quantity. Expected utilitarian welfare from the ex ante perspective thus reads as

$$EW_N \;\; = \;\; \sum_{\mathcal{G}_N} EW_N^{int}(\phi_N) \; \text{prob}(\phi_N)$$

where prob(ϕ_N) is equal to the ex ante probability to observe the sample realization g_N. Hence,

$$\text{prob}(\phi_N) = \text{prob}(g_N) = \int_{\mathcal{G}} \text{prob}(g_N \mid g)\phi(g)dg \ ,$$

where ϕ are the planner's ex ante beliefs.

We are now able to formally define an optimal allocation rule based on sampling of size N.

Definition 4.3 *The mechanism design problem based on a finite sample of size N is to find a pair of functions $[Q_N, A_N]$, which maximizes the expected welfare function EW_N subject to the constraints of feasibility, I-IC and RS_N. We will refer to the solution of this problem henceforth as $[Q_N^{**}, A_N^{**}]$.*

We call this mechanism design problem in the following the *finite sample problem.*

4.4.6 Some clarifying remarks

Remark 4.4 *A first remark concerns the question whether our sampling approach should be really understood as a sequential game with two distinct stages, an early stage for information aggregation and a final stage for public goods provision and the allocation of private goods. Even though we distinguish these two stages, we conform with the common approach in the theory of mechanism design to use the normal form for a characterization of the set of admissible allocation rules. To see this, frame the final stage as a revelation game – as opposed to a collection of decentralized utility maximization problems. This yields a revelation game with two stages, in which individuals who happen to be in the sample have to report their characteristics twice. However, the consistency requirements which shape the refined best response set ensure that sampled individuals undertake the same report twice.*

Remark 4.5 *Another clarification concerns the estimation procedure. At the sampling stage, the perceived state of the world ϕ_N is determined such that an allocation rule $[Q_N, A_N]$ does not make use of the full sample information. Neither the level of public good provision $Q(\phi_N)$ nor the menu $B(\phi_N)$ depend on the vector $w_N := \{w^i\}_{i \in S_N}$. This implies a loss of generality for two different reasons:*

i) As long as one does not assume that the empirical distributions induced by the sets of parameters $(\theta^i)_{i \in I}$ and $(w^i)_{i \in I}$ are independent, this partial ignorance will affect the statistical accuracy of the estimate of the true distribution of characteristics.

ii) Even if independence is assumed, the incentive constraints at the estimation stage – i.e. the RS_N constraints – depend on the whole vector of individual characteristics, i.e. on w and on θ. Consequently, a more general sampling approach would make the level of public good provision $Q(\cdot)$ and the menu of private goods $B(\cdot)$ a function of the empirical distribution of sample characteristics induced by (w_N, θ_N).

The main line of defense for our sampling approach is twofold. First, it is more tractable. The allocation mechanism reacts identically to all samples which give rise to the same statistical estimate of the true state of the world. Second, our main interest lies in asymptotic results. As the sample size gets very large, the law of large numbers implies that the empirical cross section distribution of $\{w^i\}_{i \in S_N}$ converges to F. Hence, in a large sample our approach should not be misleading.

Remark 4.6 *Finally, we comment a peculiarity of our approach which is a no discrimination of sampled individuals-property. That is, individuals who happen to be in the sample are – at the final screening stage – not treated differently as compared to individuals with the same characteristics who happened not to be in the sample. In particular, this precludes the possibility to design a specific scheme of payments for sample members to let them internalize the impact of their statements on the whole system. In this sense, our approach rules out transferable utility among sampled individuals. Put differently, it forces individuals to take the impact of their announcement on their personal well-being – via its impact on the final screening mechanism – into account. There is no additional instrument that permits an internalization of the impact on the well being of others.*[4]

The reason is that our attempt is to provide a guideline to a problem of equilibrium selection for incentive compatible allocation mechanisms in continuum economies. We do not aim to model a different game that gives rise to a different set of equilibria. This is captured formally by restricting sampled individuals to choose among their best responses at the final allocation stage.

[4]Green and Laffont (1979) analyze a tradeoff between information precision and communication costs in the context of Clarke-Groves mechanisms. These authors also use the idea of sampling and ask for the optimal size of the subset of individuals subjected to the play of a Clarke-Groves mechanism. Inherent in their approach is the possibility of an allocation mechanism which treats individuals in the sample different from those outside the sample. Here, by contrast, the allocation rule treats all individuals with the same characteristics identically.

4.5 Informative subscription

We return in this section to the set of anonymous allocation rules defined in Section 4.2. The objective is to propose a simple rule for breaking individual indifference if the problem of multiple equilibria, discussed in Section 4.3, arises. The main idea is illustrated as follows: Suppose an individual with characteristics (θ, w) is willing to announce (θ', w') in two states of the economy, namely if the distribution of characteristics equals g and if it equals g', where g and g' differ only by the shares of individuals with taste parameters θ' and θ. Under g', the share of individuals with θ' is higher. Imposing the requirement of *informative subscription* means, that the individual has to be weakly better off in state g. Otherwise he was tempted to break indifference in such a way as to contribute indirectly to a more favorable state perception.

In more abstract terms, we apply a lexicographic reasoning. In a revelation game based on an anonymous allocation rule, individuals make a utility maximizing announcement taking the state of the economy as given. If these considerations still allow for a wide range of optimal announcements due to multiple equilibria, then individuals are assumed to take their preferences over the composition of the economy into account. I.e. indifference is broken such that individuals contribute indirectly to the most preferred state perception. Informative subscription thus requires that an individual wants to subscribe to the 'right' group of individuals, namely those who share the own characteristics.

A clarification of the relationship between the requirements of informative subscription and robustness to sampling introduced in the previous Section 4.4, is postponed until the next section. There we will argue, for the case of separable preferences, that robust allocation rules based on sampling of size N are in the limit as $N \to \infty$ essentially – that is, for the purpose of welfare analysis – the same as allocation rules under informative subscription.

4.5.1 Allocation rules under informative subscription

Recall that an anonymous allocation rule is represented by a provision rule for the public good $Q : g \mapsto Q(g)$ and an allocation rule for private goods $A : (g, \theta, w) \mapsto A(g, \theta, w)$. As before, we restrict attention to those allocation rules which are feasible and I-IC. The purpose of this paragraph is to define an additional set of constraints, referred to as *informative subscription (IS)* constraints.

Some additional notation is needed. Suppose the true state of the economy is given by $g = (g_1, \ldots, g_m)$. With reference to g, define $\mathcal{G}^{kl}(g)$ as the subset of $\mathcal{G}$ such that, relative to g, more individuals have taste parameter

θ_l and less individuals have taste parameter θ_k, anything else being equal. Formally,

$$\mathcal{G}^{kl}(g) := \{v \in \mathcal{G} \mid \exists \epsilon \geq 0 : v_k = g_k - \epsilon,\ v_l = g_l + \epsilon,\ v_j = g_j \text{ else}\} .$$

Moreover, we write $g' \ {}^{k}\!\!\leq^{l} g$ to indicate that $g' \in \mathcal{G}^{kl}(g)$. Likewise we write $g' \ {}^{k}\!\!<^{l} g$ if $g' \in \mathcal{G}^{kl}(g) \backslash \{g\}$.

Definition 4.4 *Consider an I-IC and feasible allocation rule* $[Q, A]$. *This allocation rule is said to satisfy the property of informative subscription (IS) if the following statement holds true: for any* $g \in G$, *any* $k, l \in \{1, \ldots, m\}$ *with* $k \neq l$, *any* $g' \ {}^{k}\!\!<^{l} g$ *and any* $w \in \Omega$, *whenever* $(\theta_l, \hat{w}) \in b[\bar{g} \mid (\theta_k, w)]$ *for all* $\bar{g}$ *with* $g' \ {}^{k}\!\!\leq^{l} \bar{g} \ {}^{k}\!\!\leq^{l} g$ *then,*

$$U(Q(g), A(g, \theta_k, w), \theta_k, w) \quad \geq \quad U(Q(g'), A(g', \theta_k, w), \theta_k, w) .$$

The definition says that, whenever an individual with characteristics (θ_k, w) is willing to make the same false announcement $(\theta_l, \hat{w})$ in the true state of the world, g, as well as in any false one, $\bar{g}$, that results if, starting from g, one successively replaces individuals with taste parameter θ_k by individuals with taste parameter θ_l until arriving at g', then it must not prefer to live in state g'.

IS is a *local* requirement because, if $g'' \ {}^{k}\!\!<^{l} g' \ {}^{k}\!\!<^{l} g$, then by transitivity an individual with characteristics (θ_k, w) prefers not only g over g' and g' over g'' but also g over g''. Hence, if the IS property holds for pairs g' and g which are arbitrary close to each other, then it holds for all elements of $\mathcal{G}^{kl}(g)$. I.e. even though the above definition applies to any pair g' and g with $g' \ {}^{k}\!\!<^{l} g$, it is essentially a condition on neighboring pairs.

The postulate that $(\theta_l, \hat{w}) \in b[\bar{g} \mid (\theta_k, w)]$ for all $\bar{g}$ with $g' \ {}^{k}\!\!\leq^{l} \bar{g} \ {}^{k}\!\!\leq^{l} g$ reflects the idea that an individual is tempted to contribute to a false state perception only if he is ensured not to be harmed. Accordingly, a 'temptation to deviate' is considered only if all states $\bar{g}$ between g and g' are such that the false announcement $(\theta_l, \hat{w})$ belongs to the best response set $b[\bar{g} \mid (\theta_k, w)]$. That is, even if the individual is not able to predict exactly the final state perception but only knows the 'direction' to which he contributes, he is not forced to depart from optimizing behavior.

The IS requirement ensures incentive compatibility *ex post* in the following sense: After the true state of the world g has become commonly known, there must not exist individuals with the impression that they should have undertaken a false taste announcement in order to contribute indirectly to a different state perception.

4.5.2 IS under separable preferences

In this paragraph we show that under the assumption of separable preferences, the IS condition can be considerably simplified. Recall that the separability assumption 4.1 implies that the requirement of I-IC can be decomposed into the I-RP property and the NDT-U property. The latter implies that individuals are willing to announce any taste parameter. I.e. for all g, and all (θ, w), the set

$$b^*(\theta, w) \quad := \quad \{(\hat{\theta}, \hat{w}) \mid \theta \in \Theta \text{ and } \hat{w} = w\}$$

is contained in $b[g \mid (\theta, w)]$. Denote by $b'[g \mid (\theta, w)]$ the complement of $b^*(\theta, w)$,

$$b'[g \mid (\theta, w)] \quad := \quad \{(\hat{\theta}, \hat{w}) \in b[g \mid (\theta, w)] \mid \hat{w} \neq w\} \ .$$

In analogy to lemma 4.2, we first show that to ensure IS, it is sufficient to ensure that this property holds for reports in $b^*(\theta, w)$.

Lemma 4.3 *Suppose the separability assumption 4.1 holds. The restriction of the IS requirement to the set $b^*(\theta, w)$ is sufficient for overall IS.*

Proof. The proof proceeds by contradiction. Suppose there exists (θ, w) such that the IS requirement holds if applied to the set $b^*(\theta, w)$, but there exist g, and $k, l \in \{1, \ldots, m\}$ and $g' \overset{k}{<}{}^{l}\, g$ and $(\theta_l, \hat{w})$ such that $(\theta_l, \hat{w}) \in b'[\bar{g} \mid (\theta_k, w)]$ for all $\bar{g}$ with $g' \overset{k}{\leq}{}^{l}\, \bar{g} \overset{k}{\leq}{}^{l}\, g$ and

$$v(Q(g), \theta_k) + u(A(g, \theta_k, w), w) < v(Q(g'), \theta_k) + u(A(g', \theta_k, w), w) \ .$$

However, under separable preferences, for all g, $(\theta_l, w) \in b^*(\theta, w)$. Hence, a contradiction to the assumption that the IS requirement is fulfilled for reports in the set $b^*(\theta, w)$.

□

The observation that the IS property has to be postulated for the reports in $b^*(\theta, w)$ only, yields the following characterization of the IS property under separable preferences.

Corollary 4.2 *Suppose the separability assumption 4.1 holds. Consider an I-IC and feasible allocation rule $[Q, A]$. This allocation rule satisfies the property of informative subscription (IS) if and only if the following statement holds true: for any $g \in G$, any $k, l \in \{1, \ldots, m\}$ with $k \neq l$, any $g' \overset{k}{<}{}^{l}\, g$ and any $w \in \Omega$:*

$$v(Q(g), \theta_k) + u(A(g, \theta_k, w), w) \quad \geq \quad v(Q(g'), \theta_k) + u(A(g', \theta_k, w), w) \ .$$

Analogously to corollary 4.1, we now have a separation of incentive conditions. The IS-constraints deal with a revelation of those characteristics for which aggregate uncertainty prevails – assuming that individuals reveal their skill levels truthfully. The I-IC constraints by contrast address the revelation of skill parameters.

Under separable preferences, the condition of informative subscription has a straightforward interpretation. In any state g, an individual with taste parameter θ_k must not prefer to live in an economy in which less individuals have taste parameter θ_k. I.e. the informative subscription condition ensures that an individual always wants to support the group of individuals who share the own taste parameter.

4.5.3 Mechanism design under IS

We can now define a mechanism design problem for anonymous allocation rules $[Q, A]$ which have to satisfy feasibility, I-IC and IS conditions. As in Subsection 4.4.5 we assume that the mechanism designer has prior beliefs ϕ on the distribution of taste parameters in the economy and aims at a maximization of expected utilitarian welfare from an ex ante perspective. To derive this objective function, denote the level of utilitarian welfare realized ex post, as a function of the actual state g of the economy, by

$$W(g) \ := \ \int_\Theta \int_W U(Q(g), A(g,\theta,w)), \theta, w) f(w) dw \; g(\theta) d\theta \; .$$

The objective function is given by expected utilitarian welfare from an ex ante perspective. Accordingly, the mechanism designer has not yet observed a certain sample realization and perceives g as a random quantity. Expected utilitarian welfare from the ex ante perspective is given by

$$EW \ := \ \int_{\mathcal{G}} W(g) \phi(g) dg \; .$$

Definition 4.5 *The mechanism design problem under the requirement of informative subscription is to find an anonymous allocation rule $[Q, A]$ which maximizes EW subject to the constraints of feasibility, I-IC and IS.*

For brevity, we refer to this mechanism design problem as the *informative subscription problem.*

4.6 The relationship between RS_N and IS

In this section we argue that under the assumption of separable preferences the previously defined mechanism design problems are essentially equivalent as $N \to \infty$. That is, for large N, an allocation rule based on sampling of size N, which satisfies robustness to sampling, can be found such that it induces a level of expected welfare arbitrarily close to that of any given allocation rule which satisfies the requirement of informative subscription.

We proceed in two steps. First, it is shown that if one restricts the domain of an allocation rule $[Q, A]$ which satisfies the IS property, one arrives at a pair $[Q_N, A_N]$ with the RS_N property. This allows to show in a second step, that, for large N, a mechanism designer with a finite sample problem can approximate any outcome under informative subscription. The latter part of the argument is only sketched. We appeal to the Law of Large Numbers to establish this result. However, we do so only informally, without specifying the details of the underlying probability spaces and the appropriate notions of convergence.

Let $[Q, A]$, with $Q : \mathcal{G} \to \mathbb{R}_+$ and $A : \mathcal{G} \times \Gamma \to \mathbb{R}^L$, be a given allocation rule. Note that, for any N, $\mathcal{G}_N \subset \mathcal{G}$. Denote by $[Q_{|N}, A_{|N}]$ the restriction of $[Q, A]$ to the domain $\mathcal{G}_N$. I.e. the functions $Q_{|N} : \mathcal{G}_N \to \mathbb{R}_+$ and $A_{|N} : \mathcal{G}_N \times \Gamma \to \mathbb{R}^L$, coincide with Q and A, respectively, for $g \in \mathcal{G}_N$. The interpretation is that whenever the sample realization g_N is observed and the corresponding state perception ϕ_N is induced, then the allocation rule based on sampling $[Q_{|N}, A_{|N}]$ prescribes the same outcome as $[Q, A]$.

Intuitively, as $N \to \infty$ the set $\mathcal{G}_N$ lies dense in $\mathcal{G}$ and the domain of $[Q, A]$ and the domain of the restriction $[Q_{|N}, A_{|N}]$ are 'essentially' the same. The following lemma shows that, in addition, if $[Q, A]$ is admissible under informative subscription, then $[Q_{|N}, A_{|N}]$ satisfies all relevant constraints of a finite sample problem.

Lemma 4.4 *Suppose that preferences satisfy the separability assumption 4.1. If $[Q, A]$ satisfies feasibility, I-IC and IS, then, for any N, $[Q_{|N}, A_{|N}]$ satisfies feasibility, I-IC and RS_N.*

Proof. By corollary 4.2, the IS property implies that for any $k, l \in \{1, \dots, m\}$ and any pairwise comparison of a pair g', g with $g' \leq^{kl} g$, an individual with taste parameter θ_k prefers state g over g'. Now consider the restriction $[Q_{|N}, A_{|N}]$. For some θ_{N-i} consider two state perceptions ϕ_N and $\hat{\phi}_N$ in $\Phi^*_N(\theta_{N-1})$. Without loss of generality assume that the true sample distribution is g_N and that it induces the state perception ϕ_N, whereas $\hat{\phi}_N$ results from some $\hat{g}_N$. Then there must exist k, l such that $\hat{g}_N \,{}^{k}\!\leq^{l}\, g_N$. The RS_N property requires that sample member i prefers g_N. This is implied by the IS property.

□

We have thus shown that any allocation rule $[Q, A]$ which satisfies the constraints defining the informative subscription problem induces a sequence of restricted allocations rules $([Q_{|N}, A_{|N}])_{N=0}^{\infty}$ which are admissible for the corresponding finite sample problems. Moreover, as $N \to \infty$, the restriction $[Q_{|N}, A_{|N}]$ and the initial allocation rule $[Q, A]$ 'coincide'. We will now heuristically argue that the welfare level $EW_N[Q_{|N}, A_{|N}]$ which results from $[Q_{|N}, A_{|N}]$ under finite sampling converges to the welfare level $EW[Q, A]$ which results from $[Q, A]$ under informative subscription. To this end we make use of the Law of Large Numbers. However, we do not go into mathematical details.[5]

The expected welfare level $EW_N[Q_{|N}, A_{|N}]$ which results from $[Q_{|N}, A_{|N}]$ under finite sampling can be written as

$$EW_N[Q_{|N}, A_{|N}] \;=\; \int_{\mathcal{G}} \Big(\sum_{\mathcal{G}_N} W_N(\phi_N, g)\ \text{prob}(g_N \mid g) \Big) \phi_N(g) dg$$

where

$$W_N(\phi_N, g) := \int_{\Theta} \int_W U(Q_{|N}(\phi_N), A_{|N}(\phi_N, \theta, w)), \theta, w) f(w) dw\ g(\theta) d\theta \ .$$

Intuitively, the Law of Large Numbers, stated in assumption 4.7, implies that as $N \to \infty$, g_N lies almost surely in a small ϵ ball, $\mathcal{B}_\epsilon(g) \subset \mathcal{G}$, surrounding g. That is, as $N \to \infty$, $\text{prob}(g_N \in \mathcal{B}_\epsilon(g) \mid g) = 1$, for ϵ arbitrarily close to zero. Hence, for large N, the expression

$$\sum_{\mathcal{G}_N} W_N(\phi_N, g)\ \text{prob}(g_N \mid g)$$

is arbitrarily close to welfare level $W(g)$ which results (ex post) from $[Q, A]$ if the state of the world is given by g,

$$\lim_{N \to \infty} \sum_{\mathcal{G}_N} W_N(\phi_N, g)\ \text{prob}(g_N \mid g) \;=\; W(g) \ .$$

As a consequence,

$$\lim_{N \to \infty} EW_N[Q_{|N}, A_{|N}] \;=\; \int_{\mathcal{G}} W(g) \phi(g) dg \;=\; EW[Q, A] \ .$$

To sum up, starting out from any $[Q, A]$ with the *IS* property, for sufficiently large N, there is an allocation rule based on sampling of size N, which yields approximatively the same welfare level.

[5]The interested reader is referred to Al-Najjar (2004).

4.7 Conclusion

We have addressed a problem of public goods provision in a continuum economy with private information of individuals on their valuation of a public good and uncertainty about the aggregate valuation. It has been assumed that the decision on public good provision is based on a process of information aggregation. In addition, there is a need to cover the cost of public good provision via some incentive compatible allocation mechanism, which respects the private information of individuals on their taste as well as their skill characteristics.

As has been argued in Section 4.3, the requirement of incentive compatibility will in general give rise to a problem of multiple equilibria for the process of information aggregation. I.e. multiple equilibria do not arise incidentally. For the special case of separable preferences, it has been shown that incentive compatibility holds *only if* there are multiple equilibria. Intuitively, the reason is that, in a large economy, no single individual is ever pivotal for the quantity of the public good that is installed. As a consequence, individual behavior in the underlying revelation game is essentially undetermined.

We have formulated two different approaches to deal with this problem. The first one, discussed in Section 4.4, is based on the idea to separate more explicitly the stages of information aggregation to determine the quantity of public good provision and the financing of this desired quantity. We assume that a large random sample of individuals is used for the process of information aggregation and that sampled individuals foresee the consequences of their preference announcement. Individuals now have an impact on public good provision and their preferences over the level of public good provision govern their behavior in the revelation game. This allows to get rid of the multiple equilibrium problem.

The crucial assumption is that sampled individuals are treated at the final allocation stage not differently as compared to individuals who possess the same characteristics but have not been in the sample. This implies that sample members internalize the consequences of their own announcements for a given final allocation mechanism, which is used to finance public good provision. From a general mechanism design perspective, this assumption clearly involves a loss of generality. There certainly exist welfare superior allocation mechanism, which force sample members to internalize the consequences of their announcements not only on their own well-being but also on the well being of others. Hence, it has to be emphasized, that we ask a different question, namely how a scheme of taxation, which treats all individuals equally for public goods finance, should be designed if there is a need of information aggregation.

The second idea, formulated in Section 4.5, is a simple criterion for breaking individual indifference in the continuum: Whenever an individual is literally indifferent about the own announcement in a revelation game, use the individual's preferences over the composition of the economy to break this indifference. I.e. whenever an individual is indifferent between, say, announcements a and b but would be happy if more individuals in the economy announced b, then assume that the individual in question will announce b as well.

As such, this simple rule is subject to the criticism of ad-hocery. In the final Section 4.6 we have argued, however, that for large random samples, the two approaches are essentially equivalent. That is, the simple rule which we refer to as informative subscription can be interpreted as the limiting outcome of vanishing individual influence on aggregate data – under the premise of a sampling mechanism under which the stages of information aggregation and the final allocation stage are separated from each other.

Bibliography

Al-Najjar, N. (2004). Aggregation and the law of large numbers in large economies. *Games and Economic Behavior*, 47:1–35.

Alòs-Ferrer, C. (2002). Individual randomness in economic models with a continuum of agents. Working Paper 9807, Dpt. of Economics, University of Vienna.

Alstadsæter, A. (2003). The dual income tax and firm's income shifting through the choice of organizational form and real capital investment. Working Paper, Norweigean School fof Economics and Business Administration.

Alworth, J. (1998). Taxation and integrated financial markets: the challanges of derivatives and other financial innovations. *International Tax and Public Finance*, 5:507–534.

Alworth, J., Arachi, G., and Hamaui, R. (2002). Adjusting capital income taxation: Some lessons from the italian experience. Working Paper No. 23/10, Department of Economics, Università di Lecce.

Arrow, K. (1979). The property rights doctrine and demand revelation under incomplete information. In Boskin, M., editor, *Economics and Human Welfare*, pages 23–39. Academic Press, New York.

Auerbach, A. J. (1979). Wealth maximization and the cost of capital. *The Ouarterly Journal of Economics*, 93:433–446.

Auerbach, A. J. (1991). Retrospective capital gains taxation. *American Economic Review*, 81(1):167–178.

Auerbach, A. J. and Bradford, D. F. (2002). Generalized cash-flow taxation. Working Paper, University of California, Berkeley.

Austen-Smith, D. and Banks, J. (1996). Information aggregation, rationality and the Condorcet Jury Theorem. *American Political Science Review*, 90:34–45.

Balcer, Y. and Judd, K. L. (1987). Effects of capital gains taxation on life-cycle investment and portfolio management. *The Journal of Finance*, 42:743–757.

Banks, J., Blundell, R., and Smith, J. P. (2002). Wealth portfolios in the UK and the US. NBER Working Paper 9128.

Bierbrauer, F. (2005). Optimal taxation and public good provision in a two-class economy. Mimeo, Max Planck Institute for Research on Collective Goods, Bonn.

Boadway, R. and Keen, M. (2003). Theoretical perspectives on the taxation of capital income and financial services. In Honohan, P., editor, *Taxation of Financial Intermediation: Theory and Practice for Emerging Economies*, chapter 2, pages 31–80. New York: World Bank and Oxford University Press.

Bradford, D. F. (1995). Fixing realization accounting: Symmetry, consistency and correctness in the taxation of financial instruments. *Tax Law Review*, 50:731–785.

Christiansen, V. (1995). Normative aspects of differential, state-contingent capital income taxation. *Oxford Economic Papers*, 47(2):286–301.

Collins, J. H. and Kemsley, D. (2000). Capital gains and dividend taxes in firm valuation: evidence of triple taxation. *The Accounting Review*, 75:405–427.

Constantinides, G. M. (1983). Capital market equilibrium with personal tax. *Econometrica*, 51(3):611–636.

Crawford, V. and Sobel, J. (1982). Strategic information transmission. *Econometrica*, 50:1431–1451.

Cunningham, N. B. and Schenk, D. H. (1992). Taxation without realization: A 'revolutionary' approach to ownership. *Tax Law Review*, 47:725–814.

Dammon, R. M., Spatt, C. S., and Zhang, H. H. (2001). Optimal consumption and investment with capital gains taxes. *The Review of Financial Studies*, 14:583–616.

d'Aspremont, C. and Gérard-Varet, L. (1979). Incentives and incomplete information. *Journal of Public Economics*, 11:25–45.

Dierker, E. and Haller, H. (1990). Tax systems and direct mechanisms in large finite economies. *Journal of Economics*, 52:99–116.

Disney, R., Henley, A., and Jevons, D. (2003). House price shocks, negative equity and household consumption in the UK in the 1990s. Mimeo, University of Nottingham.

Feldman, M. and Gilles, C. (1985). An expository note on individual risk without aggregate uncertainty. *Journal of Economic Theory*, 35:26–32.

Fjaerli, E. and Lund, D. (2001). The choice between owner's wages and dividends under the dual income tax. *Finnish Economic Papers*, 14(2):104–119.

Gersbach, H. (2002). Democratic mechanisms: Double majority rules and flexible agenda costs. CESifo Working Paper Series, No. 749.

Gordon, R. H. and Bradford, D. F. (1980). Taxation and the stock market valuation of capital gains and dividends. *Journal of Public Economics*, 14:109–136.

Green, J. and Laffont, J. (1979). *Incentives in Public Decision-Making.* North-Holland Publishing Company.

Green, J. and Laffont, J. J. (1977). Characterization of satisfactory mechanisms for the revelation of preferences for public goods. *Econometrica*, 45:472–487.

Grossman, G. and Helpman, E. (2001). *Special Interest Politics.* Cambridge, MA, MIT Press.

Guesnerie, R. (1995). *A Contribution to the Pure Theory of Taxation.* Cambridge University Press.

Güth, W. and Hellwig, M. (1986). The private supply of a public good. *Journal of Economics*, Supplement 5:121–159.

Hammond, P. (1979). Straightforward individual incentive compatibility in large economies. *Review of Economic Studies*, 46:263–282.

Hammond, P. (1987). Markets as constraints: Multilateral incentive compatibility in continuum economies. *Review of Economic Studies*, 54:399–412.

Hellwig, M. (2003a). Public-good provision with many participants. *Review of Economic Studies*, 70:589–614.

Hellwig, M. (2003b). A utilitarian approach to the provision and pricing of excludable public goods. Sonderforschungsbereich 504, Discussion Paper, No. 03-36.

Hellwig, M. (2004a). A contribution to the theory of optimal income taxation. Mimeo, Max Planck Institute for Research on Collective Goods, Bonn.

Hellwig, M. (2004b). Optimal income taxation, public goods provision and public-sector pricing: A contribution to the foundations of public economics. Preprint 2004/14, Max Planck Institute for Research on Collective Goods, Bonn.

Henley, A. (1998). Changes in the distribution of housing wealth in Graet Britain, 1985–91. *Economica*, 65:363–380.

Heuser, H. (1994). *Lehrbuch der Analysis, Teil 1.* Teubner, Stuttgart, 11th edition.

Judd, K. (1985). The law of large numbers with a continuum of iid random variables. *Journal of Economic Theory*, 35:19–25.

Klein, P. (1999). The capital gain lock-in effect and equilibrium returns. *Journal of Public Economics*, 71:355–378.

Konrad, K. A. (1991). Risk taking and taxation in complete capital markets. *Geneva Papers on Risk and Insurance Theory*, 16(2):167–177.

Kovenock, D. J. and Rothschild, M. (1987). Notes on the effect of capital gain taxation on non-austrian assets. In Razin, A. and Sadka, E., editors, *Economic Policy in Theory and Practice*, pages 309–339. McMillan, London.

Land, S. B. (1996). Defeating deferral: A proposal for retrospective taxation. *Tax Law Review*, 52:45–117.

Landsman, W. R. and Shackelford, D. A. (1995). The lock-in effect of capital gains taxes: evidence from the RJR Nabisco leveraged buyout. *National Tax Journal*, 48:245–259.

Lang, M. H. and Shackelford, D. A. (2000). Capitalization of capital gains taxes: evidence from stock price reactions to the 1997 rate reduction. *Journal of Public Economics*, 76:69–85.

Mailath, G. and Postlewaite, A. (1990). Asymmetric bargaining procedures with many agents. *Review of Economic Studies*, 57:351–367.

Makowski, L. and Ostroy, J. (1992). Vickery-Clarke-Groves mechanisms in continuum economies. Characterization and existence. *Journal of Mathematical Economics*, 21:1–35.

Mas-Colell, A. and Vives, X. (1993). Implementation in economies with a continuum of agents. *Review of Economic Studies*, 60:613–629.

Mas-Colell, A., Whinston, M., and Green, J. (1995). *Microeconomic Theory*. Oxford University Press, New York.

Meade, J. E. (1978). *The Structure and Reform of Direct Taxation*. Allen and Unwin, London.

Meirowitz, A. (2004). Polling games and information revelation in the downsian framework. Mimeo, Princeton University.

Mirrlees, J. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies*, 38:175–208.

Musgrave, R. A. and Musgrave, P. B. (1984). *Public Finance in Theory and Practice*. McGraw-Hill, 4 edition.

Myerson, R. and Satterthwaite, M. (1983). Efficient mechanisms for bilateral trading. *Journal of Economic Theory*, 28:265–281.

Norman, P. (2004). Efficient mechanisms for public goods with use exclusion. *Review of Economic Studies*, 71:1163–1188.

OECD (1994). *Taxation of New Financial Instruments*. OECD.

Olsen, M. (1965). *The Logic of Collective Action*. Cambridge, MA, Harvard University Press.

Piketty, T. (1999). The information-aggregation approach to political institutions. *European Economic Review*, 43:791–800.

Poterba, J. M. (1987). How burdensome are capital gains taxes? *Journal of Public Economics*, 33:157–172.

Rendleman, R. J. and Shackelford, D. A. (2003). Diversification and the taxation of capital gains and losses. NBER Working Paper 9674, Cambridge, MA.

Sandmo, A. (1985). The effects of taxation on savings and risk taking. In Auerbach, A. and Feldstein, M., editors, *Handbook of Public Economics*, volume 1, chapter 5, pages 265–311. North-Holland, Amsterdam.

Schmitz, P. (1997). Monopolistic provision of excludable public goods under private information. *Public Finance/ Finance Publiques*, 52:89–101.

Schultz, C. (1996). Polarization and inefficient policies. *Review of Economic Studies*, 63:331–343.

Shakow, D. J. (1986). Taxation without realization: a proposal for accrual taxation. *University of Pennsylvania Law Review*, 134:1111–1205.

Shuldiner, R. (1992). A general approach to the taxation of financial instruments. *Texas Law Review*, 71(2):243–350.

Sørensen, P. B. (1994). From the global income tax to the dual income tax: recent reforms in the nordic countries. *International Tax and Public Finance*, 1:57–79.

Stiglitz, J. E. (1983). Some aspects of the taxation of capital gains. *Journal of Public Economics*, 21:257–294.

Sun, Y. (1998). A theory of hyperfinite processes: the complete removal of individual uncertainty via exact LLN. *Journal of Mathematical Economics*, 29:419–503.

Vickrey, W. (1939). Averiging of income for income tax purposes. *The Journal of Political Economy*, 47(3):379–397.

von Weizsäcker, H. and Winkler, G. (1990). *Stochastic Integrals*. Vieweg & Sohn.

Warren, A. C. (1993). Financial contract innovation and income tax policy. *Harvard Law Review*, 107:460–492.

Warren, A. C. (2004). US income taxation of new financial products. *Journal of Public Economics*, 88(5):899–923.

Welcker, J., Kloy, J. W., and Schindler, K. (1992). *Professionelles Optionsgeschäft*. Verlag Moderne Industrie, 3rd edition.

Index

Lebenslauf

Persönliche Angaben:

Name:	Marco Sahm
Geburtsdatum:	14. September 1973
Geburtsort:	Ottweiler (Saar)
Staatsangehörigkeit:	Deutsch

Schulischer Werdegang:

1980 – 1984:	Grundschule Furschweiler
1984 – 1993:	Arnold–Janssen–Gymnasium St.Wendel; Abitur (Note: 1,2)

Akademischer Werdegang:

1994 – 1996:	Grundstudium der Mathematik und Betriebswirtschaftslehre an der Universität des Saarlandes in Saarbrücken
1996 – 1997:	ERASMUS-Studium der Mathematik an der Universität 'La Sapienza' in Rom
1997 – 2001:	Hauptstudium der Mathematik und Betriebswirtschaftslehre an der Universität des Saarlandes in Saarbrücken; Diplom-Kaufmann (Abschlussnote: 1,6), Diplom-Mathematiker (Abschlussnote: 1,0)
seit 03/2001:	Wissenschaftlicher Mitarbeiter am Lehrstuhl für Finanzwissenschaft der Ludwig-Maximilians-Universität München

Sonstiges:

1993 – 1994:	Grundwehrdienst in Zweibrücken

München, 22. September 2005