ESSAYS ON THE PROVISION OF HEALTH AND EDUCATION

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

2015

vorgelegt von
Diplom-Mathematikerin Katharina Huesmann (geb. Busch)
aus Rheine

Referent:	Prof. Achim Wambach, Ph.D.
Korreferent:	Prof. Dr. Felix Bierbrauer
Tag der Promotion:	6. Juni 2016

To my family.

ACKNOWLEDGEMENTS

First of all, I would like to thank Achim Wambach for supervising this thesis. He provided me with strong support, helpful advice and encouragement whenever needed. I am also grateful to Felix Bierbrauer for numerous discussions and valuable feedback during my whole PhD and for refereeing this thesis. I thank Alexander Westkamp for chairing the defense board. Moreover, it was a pleasure to work with the co-authors of two chapters of this thesis, Achim Wambach and Wanda Mimra.

My research benefited a lot from various discussions and comments. On that account I thank Sandro Ambuehl, Sophie Bade, Pierre Fleckinger, Philipp Heller, Hartmut Kliemt, Fuhito Kojima, Liisa Laine, Albert Ma, Ulrike Malmendier, Timo Mennle, Matthew Mitchell, Yu Qin, and Al Roth. I especially appreciated valuable suggestions of Nicolas Fugger, Florian Gössl, Timo Mennle, and Wanda Mimra to improve the manuscript. I also like to thank all my colleagues and graduate fellows during the last years, in particular Anne, Alex, Christian, Christopher, David, Florian, Helene, Laura, Markus, Nicolas, Philippe, Vitali, and Wanda for many academic and non-academic discussions and a very pleasant work environment. I am thankful to Linda, Ute, Susanne, and all the student employees at our chair for the help. I am also grateful to Dagmar Weiler and the Cologne Graduate School in Management, Economics, and Social Sciences for support on various matters.

I owe big thanks to my family and friends. It especially goes to my parents, for their love, to Lioba, for repeatedly putting things into perspective, to Clara, for kicking me hard enough to come to an end, and to Martin, for so many things.

CONTENTS

INTRODUCTION

A key characteristic of markets is the interaction of demand and supply for goods or services. Rules for the interaction, however, differ significantly across markets. Conventional markets such as those for cars and garments are typically decentralized and price mechanisms are used to match demand and supply. In contrast, markets for the provision of health care and education are two salient examples of markets that are often partly centralized and subject to severe regulations that particularly concern prices. For instance, schooling is compulsory and free of charge in many countries. Typically, a significant number of schools are government-owned and the state manages the distribution of school places without using price mechanisms. Selling organs is prohibited almost everywhere in the world. In most countries of the European Union health insurance is compulsory.

What is special about health and education such that its provision requires intervention? The exceptional importance of the provision of health care and education is already emphasized in the Universal Declaration of Human Rights of 1948 as a *right to health* and a *right to education.*[1] The state is considered to be responsible for implementing these rights, which includes ensuring *availability, accessibility, acceptability, and quality.*[2] Therefore, as fairness and equality objectives may not be sufficiently reflected by free and decentralized markets, the state is obliged to take action. Intervention on health and education markets may furthermore be justified on efficiency grounds. Health and education entail positive externalities on society.[3] Regulations can also be a tool to improve welfare in view of information asymmetries which are due to the fact that it is difficult to judge the quality of health and education services.

In all three chapters of this thesis I deal with topics and questions that are particularly relevant in markets for the provision of health care and education. Each chapter is self-contained and contributes to the field of microeconomic theory.

Both Chapter 1 and Chapter 2 address the broad problem of public provision of scarce and indivisible goods. Therein, the role of wealth distribution and the

[1] Universal Declaration of Human Rights (1948), article 25 and article 26.

[2] General Comment No. 14 (2000) of the Committee on Economic, Social and Cultural Rights, article 12.

[3] Todaro and Smith (2003), for instance, describe the critical role of education and health for growth and development of a state.

impact wealth has on the assignment of the goods is of particular interest. In both chapters I consider settings in which wealth has an impact on what a consumer is willing to pay for a good and in which monetary transfers are allowed for the assignments of goods. Many economic models on the provision of indivisible goods do not reflect how wealth affects assignments since they either ban transfers or assume that wealth has no impact on somebody's willingness to pay. The burgeoning literature on *Matching Markets* considers the problem of assigning goods to consumers if monetary transfers must not be used for the assignment. Prominent applications are the distribution of school places and kidney exchange programs.[4] From an economic point of view, it is important to understand why banning monetary transfers might be desirable as it entails costs for society. Monetary incentives can improve on supply shortages like those faced in the market for organs. If consumers differ in their willingness to pay for a good, conditioning the admission for consuming the good on the payment of a price can increase revenues for funding or redistribution. On the downside, admitting monetary transfers for the assignment of goods might lead to inequality in access: if the assignment of goods conditions on the willingness to pay, inequalities in the distribution of wealth across the population gain importance on the question of "who gets what". This concern is of relevance whenever the willingness or ability to pay depends on wealth which is reflected by the settings of the first two chapters.

In contrast, Chapter 3 addresses quality concerns for the provision of health services that occur if quality cannot be observed precisely and cannot be contracted on. Therein, providers are in quality competition for patients and prices do not play a role, e.g, due to health insurance.

In the remainder of this introduction I provide an overview of the models and results of each chapter.

Chapter 1. In Chapter 1, I address the question of how to optimally assign an indivisible good of limited availability to a continuum of agents in a private information setting. A key assumption is that an agent's willingness to pay for the good increases with his wealth. While the benefit from consuming the good is the same for all consumers, the wealth level of some share of the agents is lower than the wealth level of the other agents. The model allows for randomization in the assignments, which implies that by assigning consumption probabilities to the agents each unit of the indivisible good can be treated as a divisible good with a supply of one.

[4] See, e.g. the work of Abdulkadiroğlu and Sönmez (2003) and Roth, Sönmez, and Ünver (2004).

I find that randomly assigning the good without transfers can be ex-post Pareto-efficient and can ex-ante Pareto-dominate selling the good for a market clearing price and redistributing the revenues. This explains the use of assignments without money rather than selling the good for a market clearing price in many real-world applications on grounds of efficiency. The main driver of the result is the non-linearity in preferences of the agents: the compensation an agent with low wealth is willing to accept for trading the good may exceed the willingness to pay for the good of an agent with high wealth.

I furthermore study welfare maximizing assignments. Independent of the curvature of the welfare function, neither the random assignment nor selling the good for a market clearing price and redistributing the revenues are optimal assignments. In general, optimal assignments involve wealth-dependent lotteries for the admission to consume the good and monetary transfers from agents with high wealth to agents with low wealth. The probability of being assigned to the good is lower for an agent with low wealth than it is for an agent with high wealth. The intuition is that if the good is randomly assigned without transfers, agents with low wealth have an incentive to sell at least a marginal share of their consumption probability to agents with high wealth. I find that resources of the good may be withheld by the social planner in order to increase redistribution from agents with high wealth to agents with low wealth. The welfare gains from redistribution then overcompensate the efficiency losses resulting from not assigning all available resources.

The wide-spread use of mechanisms without monetary transfers suggests that in certain markets social planners have other objectives than solely utilitarian ones. Particularly for health and education markets, it seems to be a worry that access to resources is linked to wealth if price mechanisms are used for provision. In the setting of Chapter 1, the only incentive compatible assignment that fully reflects the desire that consumption probabilities do not depend on endowments is the income-independent random assignment which is never optimal. Nevertheless, the welfare maximizing assignment might allow a greater number of low income agents to consume the good than the assignment implied by a market clearing price approach does.

Chapter 2. The model in Chapter 2 is joint work with Achim Wambach.[5] In contrast to Chapter 1, we impose an additional constraint on how to assign objects and transfers. We require that a social choice function assigning objects and

[5]Chapter 2 is a modified version of Huesmann and Wambach (2015).

transfers is discrimination-free, which is defined as the object somebody is assigned to must not depend on his wealth endowment. This reflects the concern that wealth inequalities lead to an unequal access to resources if money plays a role for the assignment of goods. We consider a finite set of agents and a set of heterogeneous objects. Agents are endowed with some wealth and, in analogy to Chapter 1, wealth impacts on the willingness to pay.

We find that mechanisms that do not use transfers and that are only based on ordinal preference information of the agents are already at the Pareto-frontier of discrimination-free social choice functions. In a private information setting, we furthermore find that requiring a social choice function to be discrimination-free implies that an agent's money assignment is independent of his preferences and his object assignment must not be sensitive to cardinal information about his preferences. In case the market designer is informed about wealth endowments, the only exception that allows him to condition somebody's object assignment on information beyond object rankings is to make ex-post wealth independent of endowments. Assigning objects without transfers can therefore be understood as a tool to satisfy a desire for wealth independent access to certain goods whenever the mechanism does not (or cannot) eliminate potential wealth differences in endowments. We furthermore find that even more restrictions than banning transfers are needed, if money can be used to improve access to a good outside the mechanisms. Examples include moving to a neighborhood of a popular school or bribing somebody to donate a kidney.

Chapter 3. The model of Chapter 3 is joint work with Wanda Mimra.[6] Therein, quality concerns for the provision of multi-dimensional health services are central. In contrast to Chapter 1 and Chapter 2, we do not consider scarce resources to assign but providers of health services are in competition for patients. It is characteristic of our model that providers solely compete in quality and not in prices. We analyze a setting where patients value quality differences in one attribute more than quality differences in the other attribute but the quality signal in the more important attribute is less precise. An example is the medical quality of the service a hospital provides that is presumably more important for a patient than the amenities a hospital offers, but also more difficult to measure. Providers' investments are stochastic and they can shift resources to raise expected quality in some attribute.

[6] Chapter 3 is based on Huesmann and Mimra (2015).

We find that rational patients focus on less important attributes if the precision of their quality signals is high compared to the more important attribute. We say that a patient focuses on an attribute if a high quality signal in this attribute drives his provider choice. This result explains the empirical finding of Goldman and Romley (2008) that various measures of treatment quality of hospitals have only a small effect on patient demand while improvements in amenities strongly raise demand. Focusing of patients has implications for the provision of quality and welfare. We find that if patients focus on less important attributes, any Perfect Bayesian Nash Equilibrium is inefficient. Furthermore, increasing signal precision can reduce welfare as the positive effect of better provider selection is overcompensated by a negative effect that a shift in patient focusing has on provider quality choice. Empirical findings of Feng Lu (2012) suggest that the shift in resources in dependence of signal precision is indeed relevant for real-world applications.[7] We finally discuss providers' strategic reporting incentives and reporting policies. In the case of optimal reporting, signals concerning the important attribute are always published. However, banning reporting on less important attributes might be necessary to enforce optimal reporting.

Our results are particularly relevant for applications since the availability of information on health services through feedback platforms as well as quality reports has significantly increased over the past years. Implications on overall quality provision are still little understood and controversially discussed. With the analysis in Chapter 3 we aim to contribute to a deeper understanding of the mechanisms of providing information that only imprecisely reflects quality.

[7]Feng Lu (2012) finds that after the introduction of public reporting scores of quality measures improve along the reported dimensions, but significantly deteriorate along the unreported ones.

CHAPTER 1

PUBLIC PROVISION OF SCARCE RESOURCES WHEN PREFERENCES ARE NON-LINEAR

Abstract

This paper considers the problem of assigning an indivisible good of limited availability to a continuum of agents who exhibit decreasing marginal utility of income. All agents equally benefit from consuming the good but may differ regarding their income. I find that randomly assigning the good to the agents without transfers can be ex-post efficient. It can also ex-ante Pareto-dominate selling the good for a market clearing price and redistributing the revenues. However, I also show that a random assignment without transfers is never optimal under the objective of welfare maximization. The unique second-best solution assigns income-dependent consumption probabilities and monetary transfers from agents with high income to agents with low income. Thereby, consumption probability of an agent with low income is distorted downwards compared to first-best, where consumption probabilities are independent of income. If income differences are large, the optimal assignment might not distribute all resources of the good in order to incentivize higher redistribution of income.

1.1 Introduction

Many goods and resources are not distributed to individuals via decentralized markets but by the state. Examples include the distribution of land, houses, education programs, and human organs. A common challenge is scarcity of resources while many people benefit from consumption. The state then has to decide whom to admit for consumption and what prices to charge. Real-world applications show a huge variety in assignment procedures. It ranges from market mechanisms where the good is sold for a market clearing price to non-market mechanisms like randomly assigning the goods without transfers.[1] Advocates of selling the goods for market prices often emphasize efficiency properties of this approach.

[1]The homepage of the US government lists various government-owned items that are sold via auctions in the US (https://www.usa.gov/auctions-and-sales, assessed on 16 December, 2015). Roth (2015) gives an extensive overview of markets where goods are distributed without price mechanisms.

In certain markets, however, price mechanisms are criticized based on moral and fairness concerns (see, e.g., Kahneman, Knetsch, and Thaler, 1986, Sandel, 2012, Satz, 2010).

Selling a good for a market clearing price assigns the good to those that have the highest willingness to pay for the good. If somebody's willingness to pay does not depend on his income, it fully reflects his benefit from consuming the good. However, if the willingness to pay increases with income, it is possible that selling the good for a market clearing price does not assign the good to the individual with the highest benefit from consumption but to the one with the highest income. This paper demonstrates that a use of random assignments without transfers instead of assignments via market prices can be explained on grounds of efficiency when preferences are non-linear in income. It is particularly relevant in markets where the benefit from consuming the good is high compared to income differences. I furthermore study assignments that maximize utilitarian welfare. I find that optimal assignments serve individuals with consumption probabilities that positively depend on income to gain revenues for redistribution. Therefore, utilitarian objectives alone cannot explain the wide-spread use of random assignments without transfers.

More specifically, I study the problem of assigning an indivisible good of limited availability to a continuum of agents when monetary transfers are possible. All agents equally benefit from consuming the good but may differ in income. Each indivisible good can be treated as a divisible good with a supply of one by assigning consumption probabilities. Each agent is either a low income type or a high income type. While the distribution of income is common knowledge, the realization of incomes is private information to the agents. Preferences of the agents are non-linear in income. In particular, utility of each agent with income e when consuming the good and paying a transfer t is described by $\theta + h(e - t)$. θ represents the benefit from consumption and $h(\cdot)$ is an increasing and strictly concave function that describes how the agent evaluates income. Therefore, the preferences exhibit decreasing marginal utility of income and the willingness to pay for the good increases in income. I restrict attention to admissible mechanisms. These are mechanisms that are feasible (i.e., resources availability is respected), incentive compatible, and individually rational.

I first discuss two classical approaches to assign the good from a welfare perspective, a market price approach and a matching market approach. In the market price approach the good is sold at a market clearing price and revenues generated are redistributed. The matching market approach assigns the indivisible good under the constraint that no transfers occur. In this setting where only one type of good exists and all agents benefit from the good in the same way, it corresponds to assigning the good randomly to the agents. I then describe optimal assignments if the objective is utilitarian welfare maximization. Non-linearities in preferences have important implications for market design, which quasilinear preferences, a typical assumption in the literature on the assignment of resources, cannot reflect.[2] Moreover, there are good reasons to assume that consumers' preferences are not quasilinear since consumers might be risk averse, budget constrained or exhibit income effects in their willingness to pay as our setting reflects.

The discussion in Section 1.4 about the market price approach and the matching market approach gives an intuition of how non-linearity in preferences impacts on the evaluation of assignment mechanisms. First, I find that random assignments without transfers are ex-post Pareto-efficient if income differences are not too large. This is because then, the compensation an agent with low income demands for selling the good is larger than the sum an agent with high income is willing to pay for the good.[3] Secondly, I find that the matching market approach might ex-ante Pareto-dominate the market price approach. The key to the problem is that an agent's willingness to pay for receiving the good with some probability π is larger than π times the willingness to pay for receiving the good with probability one. Thereby, an agent may prefer a random assignment to receiving revenue distribution from those buying the good for a market clearing price. I find that in general neither of the two classical approaches is ex-ante Pareto-efficient. The market approach can be Pareto-improved by selling admission to lotteries instead of offering the good for a fixed price. A random assignment without monetary transfers where all agents consume the good with the same probability can be Pareto-improved by shifting some marginal consumption probability from the low income agents to those with high income in return for positive monetary transfers.

In Section 1.5, I study a utilitarian mechanism design problem. Welfare is assessed via a strictly concave welfare function that evaluates the agents' ex-

[2] Baisa (2013) showed in a payoff environment similar to mine that mechanisms involving probabilistic assignments might yield higher revenues than standard auctions. Che, Gale, and Kim (2013) showed that from a welfare perspective, random assignments with resale might dominate competitive markets if consumers are budget constrained. Roughly speaking, I combine main aspects of both approaches by considering welfare maximizing assignments in a setting where agents' preferences are comparable to Baisa (2013).

[3] Knetsch and Sinden (1984), for instance, provide experimental evidence of a significant difference in the willingness to pay for a good and the compensation demanded.

pected utilities and thus reflects inequality aversion of the social planner.[4] In a first-best world, the social planner fully redistributes incomes and assigns income-independent consumption probabilities. In a second-best world, these two desires cannot be satisfied simultaneously. Generating revenues for redistribution requires that consumption probabilities depend on income in order to prevent agents with high income to mimic agents with low income. I find that there is a unique solution to the utilitarian problem. It involves transfers and income-dependent consumption probabilities. The consumption probabilities of low income agents increase in their income but are strictly smaller than the one for a high income type. If income differences are large, the optimal assignment may even not assign all resources. This is because the welfare loss of not assigning some of the resources is overcompensated by the welfare gain through increasing the revenues for redistribution. Furthermore, I show that the more inequality averse the social planner is, the more the consumption probabilities differ across agents with high and low income in order to increase monetary transfers. If the benefit from consumption becomes arbitrarily large, the difference in consumption probabilities goes to zero. This is because the higher the benefit, the easier it is to generate revenues.

Therefore, in a setting where differences in the willingness to pay are solely due to income differences, non-market mechanisms like a random assignment can be closer to welfare maximization than a market price approach is. However, the wide-spread use of non-market mechanisms cannot be solely explained by classical utilitarian objectives. Condorelli (2013) argues that a social planner might have objectives that are not directly linked to consumers' utilities. One potential desire that justifies a matching market approach in this setting is an income independent access to the resources. Compared to a market price approach, the welfare maximizing assignment allows a higher share of low income types to consume the good if income differences are small enough. For high income differences, however, the market price approach might lead to a higher share of low income types consuming the good. This is the case if the optimal assignment withholds resources in order to increase redistribution.

In the following section I discuss related literature. In Section 1.3, the setting and the general assignment problem are presented. Then, the market price approach, the matching market approach (Section 1.4) and welfare maximizing assignment (Section 1.5) are worked out in detail and discussed. In Section 1.6 I conclude.

[4]Thereby, I consider a setting where equal opportunities are desired rather than distributive justice.

1.2 Related Literature

Broadly speaking, this work relates to the economic literature on assigning indivisible goods and monetary transfers. In the following I first refer to the literature that shows analogies to how I treat the consumers' preferences. Then, I outline the literature on the objectives and constraints of a social planner that is especially relevant for my work.

Non-Linear Consumer Preferences. The central idea of my setting is the non-linearity of preferences expressed by a decreasing marginal utility for money. The literature on mechanism design with non-linear preferences is scarce although there are good reasons to assume that consumers exhibit non-linear preferences and that they are relevant for the design of certain markets.[5] There is some work on how non-linearity in preferences impacts on auction design. Maskin and Riley (1984) were among the first to consider auctions in a non-linear setting. In a recent work Baisa (2013) analyzed a canonical auction market when bidders are risk averse and their willingness to pay increases in income. He showed that probabilistic allocations might Pareto-dominate the second price auction. In contrast, I study welfare maximizing assignments in a large market and I am interested in how wealth endowments impact on the assignment. However, the attractiveness of randomizations when preferences are not linear in income is also reflected in his setting. Che et al. (2013) considered a problem of assigning scarce resources to a continuum of agents with quasilinear preferences but who might be budget constrained. They showed that random assignments with resale might dominate competitive markets. In their model, the value of income is the same for all agents. Therefore, in contrast to our work, there is no desire for redistribution. The work of Garratt and Pycia (2014) is an example of how dropping the assumption of quasi-linearity might turn impossibility results into possibility results. They showed that the efficient bilateral trade problem of Myerson and Satterthwaite (1983) is solvable in certain settings when preferences incorporate positive wealth effects.

Utilitarian Mechanism Design. In the main part of this paper I study welfare maximizing assignments of indivisible goods and monetary transfers in large

[5] The American Medical Association on distributing medical resources: "At present, though, the disparity among incomes across society distorts the accuracy of the market model as a fair tool for distributing scarce medical resources, for the amount an individual can spend to gain access to a needed treatment will often fall short of his or her actual valuation of it." (American Medical Association Council on Ethical and Judicial Affairs, 1995)

markets. Thereby, my research is related to the literature on optimal provision of excludable public goods in large markets. Hellwig (2005) showed that for a utilitarian social planner, excluding agents from consuming the good might be desirable for redistribution. In my setting, access is restricted exogenously through scarcity of resources. I also find that further restrictions in access might be desired to increase redistribution. However, in my setting the reason for which increasing redistribution through limiting access may be desired is not purely a matter of inequality aversion and heterogeneity in consumption benefits but is furthermore due to income differences that lead to different marginal utilities in income. Hellwig (2010) showed in setting similar to that of Hellwig (2005) that randomization in admissions can be desirable if inequality aversion is sufficiently high. The goals and constraints for a social planner to distribute the goods in this work are similar to the ones of Hellwig (2010) while the agents' preferences differ. In my setting, randomization in admission is attractive in general. This is driven by the non-linearity of the preferences.

Non-Utilitarian Objectives. Condorelli (2013) studied optimal allocations for quasi-linear preferences if a market designer has objectives other than utilitarian welfare maximization but the allocation can only be conditioned on the willingness to pay and observable characteristics. He showed that whether a market allocation is optimal depends on how much the social planner's goals and the agents' willingness to pay are linked. It relates to my work since it might not be optimal either to allocate the good to the one with the highest willingness to pay. In contrast to Condorelli (2013), it even holds if the social planner has utilitarian objectives but is rather driven by the fact that the willingness to pay does not reflect the benefit from consumption.

Mechanism Design under Constraints. The literature on Matching Markets deals with the problem of how to assign indivisible goods to consumers if money must not be used for the assignment. Popular applications are assignments of school places, housing, or kidneys (see, e.g., Abdulkadiroğlu and Sönmez, 1999, Abdulkadiroğlu and Sönmez, 2003, Roth et al., 2004). Several works discuss motivations behind prohibiting transfers (see, e.g., Roth, 2008, Sandel, 2012, Satz, 2010). However, concerns associated with prices are barely integrated in economic models. Huesmann and Wambach (2015) consider a setting in which the consumers' willingness to pay increases in wealth and in which it is desired to have income-independent access to resources. They show that then a price mechanism cannot be used to assign the good. In my setting, there is only one type of indivisible good. Therefore, a matching market approach is limited to assigning the good randomly with a uniform distribution across the agents. The matching market approach then reflects a desire of income independent access.

1.3 The General Assignment Problem

There is a continuum of agents with total mass equal to one. In the economy, there is an indivisible good that is available with capacity $Q < 1$. Therefore, not all agents can be provided with the indivisible good. Initially, nobody owns the good but it is owned by the state. Each agent can at most receive one unit of the indivisible good. Furthermore, each agent j initially owns an amount $e^j \in \mathbb{R}$ of money, his income. e^j is distributed across agents according to a distribution function $G(\cdot)$. I concentrate on the case that $G(\cdot)$ is a binary distribution. A share $G_L \in (0, 1)$ of the agents is endowed with $e_L \in \mathbb{R}$ (referred to as the low income type or L-type) and a share $G_H = 1 - G_L$ of agents is endowed with $e_H > e_L \in \mathbb{R}$ (referred to as the high income type or H-type).

An *assignment* $(\chi, t) = (\chi^j, t^j)_{j\in[0,1]}$ determines for each agent j a transfer t^j of money and whether he consumes one unit of the indivisible good, denoted by $\chi^j \in \{0, 1\}$. $\chi^j = 1$ corresponds to the agent being admitted to consuming the good, $\chi^j = 0$ corresponds to the agent not being admitted to consuming the good. The payoff of an agent with individual assignment (χ^j, t^j) and income e^j is

$$\theta\chi^j + h(e^j - t^j),$$

where $\theta \in \mathbb{R}_+$ is a parameter that determines the benefit of consuming the indivisible good. Since I am interested in the impact of wealth differences on optimal assignments rather than the impact of differences in benefits from consumption, I concentrate on the case where all agents have the same benefit θ when consuming the indivisible good. $h(\cdot)$ is three times differentiable with $h' > 0$, $h'' < 0$ and $h''' > 0$.[6] Therefore, agents exhibit decreasing marginal utility of income ($h'' < 0$). This implies positive wealth effects in the sense that their willingness to pay strictly increases for increasing income. Furthermore, since $h''' > 0$, agents exhibit decreasing absolute risk aversion.[7]

For any given benefit parameter θ, the assignment (χ^j, t^j) depends on agent j only through the agent's endowment e^j and the realization r^j of an exogenously

[6] I allow the agents to have negative levels of income. Restricting incomes level and, for instance, considering $h(e) = \ln e$ for $e \in \mathbb{R}_+$ neither changes the proceeding of my analysis nor the characteristics of the results.

[7] According to Kimball (1990), $h''' > 0$ can be interpreted as prudence in the sense that an agent undertakes precautionary savings when facing uncertainty.

given indicator variable $\tilde{r}^j$. $\tilde{r}^j$ takes values in the unit interval and has a uniform distribution denoted by ν that is the same for all agents. The introduction of the indicator variable allows for randomization in the assignments.

An assignment can then be described by

$$(\chi_L(r), \chi_H(r), t_L(r), t_H(r)) = (\chi(e_L, r), \chi(e_H, r), t(e_L, r), t(e_H, r)),$$

with $r \in [0, 1]$. Any agent j with realization r^j of the random variable $\tilde{r}^j$ is then assigned to $(\chi_L(r^j), t_L(r^j))$ if his income is e_L and to $(\chi_H(r^j), t_H(r^j))$ if his income is e_H. Furthermore, for $i \in \{L, H\}$,

$$\pi_i := \int \chi_i(r) d\nu(r)$$

denotes the probability that an agent with income e_i is admitted to consume the good.

$$v_i := \pi_i \theta + \int h(e_i - t_i(r)) d\nu(r) \tag{1.1}$$

denotes an agent's expected payoff if his endowment is e_i with $i \in \{L, H\}$.

Endowments $e^j \in \{e_L, e_H\}$ are assumed to be the agents' private information. In analogy to Hellwig (2010), from the perspective of other agents and the market designer, e^j is the realization of a random variable $\tilde{e}^j$ that takes the values e_L with probability G_L and e_H with probability G_H. Again like Hellwig (2010) I assume a large-number effect such that with probability one, the joint distribution $G \times \nu$ is the cross-section distribution of the pair (e^j, r^j) in any nonnegligible subset of the population. Based on the large economy effects, π_i then corresponds to the share of agents of type i that receive the indivisible good.

An assignment is *admissible* if it is *feasible*, *incentive compatible* and *individually rational*. Feasible assignments are those that respect resource availability. Formally,

$$G_L \int t_L(r) d\nu(r) + G_H \int t_H(r) d\nu(r) \geq 0, \tag{1.2}$$

$$G_L \pi_L + G_H \pi_H \leq Q. \tag{1.3}$$

An assignment is incentive compatible, if truthful reporting is a weakly dominant strategy for each agent. Therefore, for $i, j \in \{L, H\}$ it has to hold that

$$v_i \geq \theta \pi_j + \int h(e_i - t_j(r)) d\nu(r) \tag{1.4}$$

An assignment is individually rational if every agent weakly prefers his assignment to not participating. Therefore, for any $i \in \{L, H\}$

$$v_i \geq h(e_i). \tag{1.5}$$

The budget constraint for transfers reflects that the social planner does not need to extract any money that is needed for funding the provision of goods.[8]

Implication of the Preferences on the Willingness to Pay. The assumptions that preferences in income can be expressed by a strictly concave function $h(\cdot)$ and that consumption utility does not depend on income have some important implications for the following analysis. First, the willingness to pay t_p of an agent with income e for an object is lower than the compensation t_a he is willing to accept in order to give up the object (i.e., his willingness to accept). Formally, t_p and t_a are described by the equations

$$\theta + h(e - t_p) = h(e) \quad \text{and} \quad \theta + h(e) = h(e + t_a)$$

These equations imply that $h(e) = \frac{1}{2}(h(e + t_a) - h(e - t_p))$ which can only hold if $t_p < t_a$ since $h(\cdot)$ is strictly concave. A second implication of the specifications of preferences is that an agent's willingness to pay for attending a lottery that serves him the good with probability π is higher than π times his willingness to pay t_p for receiving the good with probability one. To see this, note that strict concavity of $h(\cdot)$ implies that $h(e) - \pi\theta < h(e - \pi t_p)$. This is equivalent to $\pi\theta + h(e - \pi t_p) > h(e)$ which implies the desired.

1.4 Two Classical Approaches

In this section I discuss a classical market and a classical non-market approach to assign resources and design transfers. Both are indeed widely used in real-world applications for assignments of goods. The first one is a market price approach where the indivisible good is offered for a market clearing price such that supply equals demand. Revenues are redistributed to society. The second one is a matching market approach that, in this setting, corresponds to a random assignment of resources to agents.

1.4.1 Market Price Approach

Suppose the social planner announces a price for the indivisible good such that supply equals demand. Revenues generated are redistributed to all agents. The

[8] If there is some cost K to cover, the budget constraint changes to total transfers being at least K instead of zero. If K is small enough, nothing significant changes in the analysis that follows. Then, revenues generated by the social planner are not completely redistributed but also used for financing the cost of provision. If K is large, the participation constraint may not be satisfied which implies that nothing is distributed.

prices and the share of each type consuming the indivisible good depends on the availability Q of the indivisible good.[9] Whether the H-types or the L-types determine the market clearing prices depends on the availability of the indivisible good.

For $Q \leq G_H$ market clearing prices are such that the H-type is just indifferent between buying the good or not. Then, a share $\frac{Q}{G_H}$ of the H-types consumes the good. Everybody consuming the good pays a transfer t_{MP} to those not consuming the good such that

$$\theta + h(e_H - t_{MP}) = h(e_H + \frac{Q}{1-Q} t_{MP}).$$

For $Q > G_H$ market clearing prices are such the L-type is just indifferent between buying the good or not. Then, all H-types consume the good and a share $\frac{Q-G_H}{G_L}$ of the L-types consumes the good. The ones consuming the good pay a transfer t_{MP} such that

$$\theta + h(e_L - t_{MP}) = h(e_L + \frac{Q}{1-Q} t_{MP}).$$

The final assignment is ex-post efficient. Agents do not have any incentives to trade the indivisible good after the assignment.

1.4.2 Matching Market Approach

Characteristic of a matching market approach is the exclusion of transfers. The matching market literature for one-sided markets deals with the question of how to assign indivisible goods to agents without using any transfers. For some real-world markets, transfers often are even forbidden by law in the respective countries. For instance, nearly everywhere in the world it is forbidden to sell organs.

This paper considers one indivisible good of limited availability and agents benefit from consumption in the same way. Differences in the willingness to pay are solely due to endowment differences. Then, requiring no transfers is equivalent to requiring income independent access to resources if endowments are private information to the agents. Therefore, an assignment of a matching market approach pools both income types: no transfers are made and all agents receive an indivisible good with probability Q, independent of their endowment. A share Q of L-types consumes the good and a share Q of H-types consumes the good.

[9] For $Q = G_H$ (i.e., availability corresponds exactly to the amount of H-types) the market clearing price is not unique but several prices such that supply equals demand exist. Then select the market clearing price that is the highest.

In the following, I discuss the agents' incentives to resell the good after the assignment (which is, in fact, forbidden in many markets). Because of the strict concavity of $h(\cdot)$, an L-type agent that received the good through a random assignment may not be willing to sell the good to an H-type agent that did not receive the good. Despite a lower endowment, the compensation that an L-type is willing to accept might exceed the willingness to pay for the good of an H-type. It occurs if income differences are small enough given any benefit θ from consumption. Therefore, there exist parameters such that the matching market approach is ex-post Pareto-efficient in the sense that no Pareto-improvements through trades after the assignment is performed can be realized. If resale is permitted, the share of L-types and H-types consuming the indivisible good then either corresponds to the one of the matching market approach and is the same for both types or it corresponds to the one that results by offering the good for a market clearing price. This discussion also shows that the Coase theorem does not hold in this setting since the initial allocation matters.

1.4.3 Market Price versus Matching Market Approach

An attractive feature of the market price approach is that no agent has an incentive to trade the good after the assignment is performed. For non-linear preferences, the matching market approach might be ex-post efficient as well as discussed above. Therefore, a classical argument by opponents of non-market approaches that random assignments lead to allocations where consumers wish to resell the good therefore does not necessarily hold if preferences exhibit wealth effects.

There is even an efficiency argument in favor of the matching market approach compared to the market price appraoch. The following proposition shows that a random assignment without transfers can ex-ante Pareto-dominate the market price approach.

Proposition 1.1. *Fix θ, Q, and e_H. There exists some $e_L^c < e_H$ such that for all $e_L \in (e_L^c, e_H)$ the assignment where each agent receives the indivisible good with probability Q and no transfers are made ex-ante Pareto-dominates the assignment where the good is sold for a market clearing price and revenues are redistributed.*

Proof. See appendix. □

The intuition for the result is once again based on the non-linearity in the preferences. An agent's willingness to pay for receiving a good with probability

Q is higher than Q times the willingness to pay for receiving the good with probability one. On the one hand, this implies that the type that defines the market clearing price (the H-type for $Q \leq G_H$ and the L-type for $Q > G_H$) prefers the random assignment over the good being sold for a market clearing price for any endowment values e_L and e_H. On the other hand, if e_L and e_H are close enough, the valuation for the L-type and the H-type served by each mechanisms are arbitrarily close. This implies that for low income differences, both types prefer the random assignment.

The analysis shows that for the L-type, the payment he receives when not buying the object for a market clearing price, does not compensate the benefit from attending a lottery with a chance of Q for receiving the good whenever income differences are low. If resources are such that more than only the H-types can be served, the L-types' preference for the random assignment is even independent of e_L and e_H. Therefore, a Rawlsian planner that aims to find an assignment that maximizes the valuation of the L-types prefers the random assignment over the market price approach independent of e_L and e_H.

1.5 Welfare Maximizing Assignment

In the following I consider a utilitarian social planner that aims to maximize welfare. The *utilitarian problem* is to find an assignment that maximizes welfare $\mathcal{W}$ within the set of admissible assignments. Welfare is assessed through

$$\mathcal{W} = G_L W(v_L) + G_H W(v_H) \tag{1.6}$$

where $W(\cdot)$ is a welfare function and $v_i = \theta\pi_i + \int h(e_i - t_i(r))d\nu(r)$ is the expected payoff of an agent with income e_i.[10] I consider welfare functions $W(\cdot)$ that are strictly increasing, strictly concave and twice continuously differentiable. This means that the social planner is inequality averse. For simplicity I assume that the share of L-types equals the share of H-types,[11] therefore

$$G_L = G_H = \frac{1}{2}.$$

Since $\chi_i(r)$ only enters the utilitarian problem through the valuation $v_i = \theta\pi_i + \int h(e_i - t_i(r))d\nu(r)$ with $\int \chi_i(r)d\nu(r) = \pi_i$, it implies that for solving the utilitarian problem it is sufficient to determine the optimal π_L and π_H. $\chi_i(r)$ is then

[10] Ex-post outcomes, i.e., outcomes after each lottery $\tilde{r}^j$ is conducted, do not enter the welfare function. Therefore, I consider a situation where equal opportunities are desired rather than distributive justice.

[11] Assuming different shares neither changes the procedure of the following analysis nor the characteristics of the results.

chosen such that $\int \chi_i(r)d\nu(r) = \pi_i$ for $i \in \{L, H\}$. Therefore, in the following any $(\pi_L^*, \pi_H^*, t_L^*(r), t_H^*(r))$ that maximizes welfare $\mathcal{W}$ within the set of admissible assignments is called optimal or a solution to the utilitarian problem.

1.5.1 First-best Solutions

In a first-best world the social planner is informed about income levels and is not restricted by incentive or participation constraints. Based on how agents evaluate assignments and the shape of the welfare function, the social planner aims at smoothing income levels as well at smoothing utility levels. In a first-best world, she can choose the assignment such that income levels and consumption probabilities of the good are equal for both types.

This is indeed the unique first-best solution. To see this, first note that it is not desirable that transfers depend on randomization. This is because removing an agent's uncertainty regarding his transfers by assigning a transfer that equals the expected transfer unambiguously increases welfare. Furthermore, utilities v_L and v_H need to be equal in the first-best assignment. If $v_L \neq v_H$, either a shift in consumption probabilities or a shift in income from the type that is better off to the other type is a welfare improvement. This is because the shift can be chosen such that inequality among the two types decreases while the sum of utilities is at least the same. Therefore, for the first-best assignment it has to hold that $v_L = v_H$. If this is achieved by different consumption probabilities for both types it cannot be optimal. This is because due to the strict concavity of $h(\cdot)$, money can be shifted from the type with the lower ex-post income to the other type in turn for some consumption probability of the indivisible good such that nobody is worse off but the sum of valuation increases. Therefore, the first-best assignment is the unique assignment that equalizes income and assigns the same consumption probabilities of the indivisible good to the consumers. The following proposition summarizes the first-best case. Notably, the solution is independent of the welfare function $W(\cdot)$.

Proposition 1.2 (First-best). *Fix θ, Q, and $e_L < e_H$. In the first-best assignment each agent is admitted to consume the indivisible good with probability Q, i.e., $\pi_L^* = \pi_H^* = Q$. Transfers satisfy $t_i^* = e_i - \mathbb{E}\tilde{e}$ with $\mathbb{E}\tilde{e} = \frac{1}{2}(e_L + e_H)$. Payoffs then are such that $v_L^* = v_H^* = \theta Q + h(\mathbb{E}\tilde{e})$*

Proof. The preceding discussion directly implies the proposition. □

It is straightforward to see that the first-best assignment is not incentive compatible and therefore cannot be implemented if income of the agents is unknown to the social planner. A high income type that imitates the low income type is better off compared to admitting being the high income type. Later in this section I will furthermore consider a setting where the social planner knows the agents' endowments but has to ensure that the assignment is individual rational such that all agents would like to participate.

1.5.2 Second-best Solutions

In the following, the incentive compatible constraints as well as the participation constraints are active. I first show that any second-best assignment incorporates transfers t_L and t_H that are independent of the randomization variable. The utilitarian problem can then be reduced to maximize welfare within the set of admissible consumption probabilities $\pi_L, \pi_H \in [0,1]$ and transfers $t_L, t_H \in \mathbb{R}$. This is an important insight when searching for solutions to the utilitarian problem. In a first-best world it is straightforward to see that any assignment where transfers depend on randomization is dominated by an assignment with constant transfers. However, in a second-best world randomization does not only impact on the objective function but also impacts on the incentive constraint.

Lemma 1.1. *Fix θ, Q, and $e_L < e_H$. Suppose $(\pi_L^*, \pi_H^*, t_L^*(r), t_H^*(r))$ is a solution to the utilitarian problem. Then, transfers are independent of the randomization variable, i.e., $t_L^*(r) = t_L^* \in \mathbb{R}$ and $t_H^*(r) = t_H^* \in \mathbb{R}$.*

Proof. See appendix. $\square$

The lemma relies on several drivers. On the one hand, preferences of the agents are such that they are risk neutral regarding their consumption of the good but are, due to the concavity of $h(\cdot)$, risk averse with respect to uncertainty in transfers. Therefore, uncertainties in transfers are worse for the agents than receiving the respective expected transfer. However, since the assignment choice is restricted by an incentive compatibility constraint, making the payments for the L-type dependent of the lottery might help to generate higher payments from the H-type without violating the incentive constraint of the H-type. To see the mechanism behind, suppose the incentive constraint of the H-type is binding. By introducing the lottery the assignment of the L-type is less worth for the H-type. Therefore, he is willing to pay more for his assignment. Since there is decreasing absolute risk aversion, the negative effect of the randomization for the L-type' valuation, however, outweighs the positive effect of higher payments (compare Hellwig, 2007). This implies that payments of the L-type are always independent of randomization. To prove that payments of the H-type are also independent of randomization, I construct admissible and constant transfers that strictly improve welfare.

Lemma 1.1 allows to limit attention to assignments with transfers that are independent of randomization when searching for solutions to the utilitarian problem. In the following, I present several characteristics of optimal assignments that help to reduce the utilitarian problem to a maximization problem in one dimension. In particular, it is useful to know which of the constraints are binding. A further simplification of searching for a solution of the utilitarian problem is that the participation constraint can be omitted. This is driven by the assumption that no cost of provision has to be extracted from the mechanism.

Lemma 1.2. *Fix θ, Q, and $e_L < e_H$. Any assignment that maximizes welfare within the set of feasible and incentive compatible assignments also satisfies individual rationality. Suppose $(\pi_L^*, \pi_H^*, t_L^*, t_H^*)$ is a solution to the utilitarian problem. It implies*

1. $\pi_L^* \leq \pi_H^*$, $t_L^* \leq t_H^*$, *and* $v_L^* < v_H^*$.
2. *The incentive compatibility constraint for the H-Type is binding.*
3. *The feasibility constraint with respect to transfers is binding.*
4. *If the feasibility constraint with respect to the indivisible good is not binding, it implies that $\pi_H^* = 1$.*

Proof. See appendix. □

By Lemma 1.2, for any solution to the utilitarian problem transfers satisfy $t_L^* = -t_H^*$. Denote by

$$t^* = t_H^* = -t_L^* \geq 0$$

the money each H-type agent pays and each L-type agent receives in optimum. The following discussion demonstrates that the utilitarian problem can be reduced to a one-dimensional maximization problem without constraints that is about finding the optimal transfer t^*. It will be summarized in Proposition 1.3. The main idea is that if t^* is the transfer of a solution to the utilitarian problem it uniquely implies the optimal consumption probabilities π_L^* of the L-type. π_L^*, in turn, uniquely implies the consumption probabilities π_H^* of the H-type. This allows to limit the search for optimal assignments to searching for optimal transfers.

Definition of $\pi_H(\pi_L^*)$. The insights of Lemma 1.2 about the feasibility constraint for the indivisible good imply that once $\pi_L^* \in [0,Q]$ is part of a solution to the utilitarian problem, it uniquely defines π_H^*. If $\pi_L^* \leq 2Q-1$, $\pi_H^* = 1$ cannot be part of an optimal solution. This implies that the constraint is binding, and therefore $\pi_H^* = 2Q - \pi_L^*$. This always holds for $Q < \frac{1}{2}$. If $\pi_L^* > 2Q-1$, which can only occur if $Q > \frac{1}{2}$, the constraint cannot be binding and, therefore, $\pi_H^* = 1$ holds. The relationship of π_H^* and π_L^* can be summarized by

$$\pi_H^* = \pi_H(\pi_L^*) = \min\{2Q - \pi_L^*, 1\} \text{ for all } \pi_L^* \in [0,Q]. \tag{1.7}$$

Definition of $\pi_L(t^*)$. The insights of Lemma 1.2 about the binding characteristics of the incentive compatibility constraint in combination with $\pi_H(\pi_L^*)$ implies a bijection between the optimal probability π_L^* that an L-type agent is admitted to consume the indivisible good and the optimal transfer t^* that is assigned. A binding incentive compatibility constraint of the H-type implies that

$$\pi_L^*\theta + h(e_H + t^*) = \pi_H^*(\pi_L^*)\theta + h(e_H - t^*) \tag{1.8}$$

$$\Leftrightarrow \quad h(e_H + t^*) - h(e_H - t^*) = (\min\{2Q - \pi_L^*, 1\} - \pi_L^*)\theta \tag{1.9}$$

Strict concavity implies that if the H-type is indifferent between the two assignments, the incentive constraint for the L-type is satisfied as well. Furthermore, the left hand side of the last equality is strictly increasing in t^* and the right hand side is strictly decreasing in π_L^*. Therefore, if for any $\pi_L^* \in [0,Q]$ there exists some t^* such that the equation is satisfied, it has to be unique. Such a transfer t^* exists, because the right hand side is larger or equal to zero while the left hand side is zero for $t^* = 0$ and converges to infinity for $t^* \to \infty$. In the following, let t_M the maximal transfer t^* that can be demanded by the H-types without violating the incentive constraint. $t^* = t_M$ is optimal if the difference in consumption share of the two types is maximal. This is the case for $\pi_L^* = 0$. Therefore, t_M is uniquely defined by

$$h(e_H + t_M) - h(e_H - t_M) = \min\{2Q, 1\}\theta.$$

Taken all together, there is a continuous bijection of optimal transfers $t^* \in [0, t_M]$ and optimal consumption probabilities $\pi_L^* \in [0,Q]$. The discussed insights directly lead to the following proposition that describes how to reduce the utilitarian problem to a one-dimensional maximization problem.

Proposition 1.3 (Reduced Utilitarian Problem). *Fix θ, Q, and $e_L < e_H$. Let t_M be defined by $h(e_H + t_M) - h(e_H - t_M) = \min\{2Q, 1\}\theta$. For any $t \in [0, t_M]$, let $\pi_L(t)$ be the unique $\pi_L \in [0,Q]$ with $h(e_H + t) - h(e_H - t) = (\min\{2Q - \pi_L, 1\} - \pi_L)\theta$.*

Any assignment $(\pi_L^*, \pi_H^*, t_L^*(r), t_H^*(r))$ *is a solution to the utilitarian problem if and only if* $t^* \in [0, t_M]$ *maximizes*

$$\mathcal{W}(t) = W\left(\pi_L(t)\theta + h(e_L + t)\right) + W\left(\pi_L(t)\theta + h(e_H + t)\right) \tag{1.10}$$

and $\pi_L^* = \pi_L(t^*)$*,* $\pi_H^* = \min\{2Q - \pi_L^*, 1\}$ *as well as* $t_H^*(r) = -t_L^*(r) = t^*$ *hold.*

Proof. The proof is implied by the analysis preceding the proposition. Note that for the definition of the objective function $\mathcal{W}(t)$ it is already exploited that the expected utility of the H-type, $\pi_H\theta + h(e_H - t)$, equals $\pi_L\theta + h(e_H + t)$ according to the binding incentive constraint for the H-type. □

By Proposition 1.3, the utilitarian problem is reduced to finding $t^* \in [0, t_M]$ that maximizes $\mathcal{W}(t)$ according to equation (1.10). This problem is a straight forward optimization problem without constraints.[12] The monotonicity properties of $\mathcal{W}(t)$ help to find a maximizer of expression (1.10). $\mathcal{W}(t)$ is continuous but it is not necessarily differentiable on the whole interval $[0, t_M]$. This is because the explicit functional form of $\pi_L(t)$ depends on whether the feasibility constraint with respect to the indivisible good is binding or not. For $Q \leq \frac{1}{2}$ the feasibility constraint is binding and the minimum function does not cause any kinks in $\mathcal{W}(t)$. Then, $\min\{2Q - \pi_L(t), 1\} = 2Q - \pi_L(t)$. This implies

$$\pi_L(t) = Q - \frac{1}{2\theta}[h(e_H + t) - h(e_H - t)] \text{ for } t \in [0, t_M]. \tag{1.11}$$

For $Q > \frac{1}{2}$, $\mathcal{W}(t)$ exhibits a kink at $t = t_1$ where t_1 is the unique transfer such that $\pi_L(t_1) = 2Q - 1$. Therefore, t_1 is uniquely defined by

$$h(e_H + t_1) - h(e_H - t_1) = (1 - Q)2\theta \tag{1.12}$$

The reason for the kink is that as long as $\pi_L^* < 2Q - 1$, the feasibility constraint with respect to the indivisible good is binding, and therefore π_H^* is strictly decreasing in π_L^*. For $\pi_L^* \geq 2Q - 1$, the constraint is not binding and $\pi_H^* = 1$ holds independent of π_L^*. Therefore, the function form of $\pi_L(t)$ changes at $t = t_1$. Formally,

$$\pi_L(t) = \underline{\pi}_L(t) = Q - \frac{1}{2\theta}[h(e_H + t) - h(e_H - t)] \text{ for } t \in [0, t_1], \text{ and} \tag{1.13}$$

$$\pi_L(t) = \overline{\pi}_L(t) = 1 - \frac{1}{\theta}[h(e_H + t) - h(e_H - t)] \text{ for } t \in (t_1, t_M] \tag{1.14}$$

[12] In the same manner equation (1.10) is expressed in dependence of t. One could write the equation in dependence of π_L based on the bijection developed above. However, the notation in dependence of t is more convenient for the following calculations.

$\mathcal{W}(t)$ is therefore twice differentiable on $[0, t_1]$ as well as on $(t_1, t_M]$. The first and second derivatives $\mathcal{W}'(t)$ and $\mathcal{W}''(t)$ can be meaningful defined on $[0, t_M]$ as the left derivatives. For all $t \in [0, t_M]$ except for t_1 it then holds that the left derivative equals the right derivative. $\mathcal{W}'(t)$ and $\mathcal{W}''(t)$ are then not continuous since they have a point of discontinuity at $t = t_1$.

The following proposition reveals that solutions to the utilitarian problem are unique and, also, how to find the optimal transfer by evaluating the first derivative of $\mathcal{W}(t)$.

Proposition 1.4 (**Optimal Assignments**). *Fix θ, Q, and $e_L < e_H$. There exists (up to modifications on a null-set) a unique solution to the utilitarian problem. The solution entails a monetary transfer $t^* > 0$ such that each H-type agent pays t^* and each L-type agent receives t^*. t^* is the unique $t^* \in [0, t_M]$ such that $\mathcal{W}'(t) \geq 0$ if and only if $t \leq t^*$. The share of L-type agents consuming the good is strictly lower than the share of H-type agents consuming the good.*

Proof. See appendix. □

The proposition is proved by showing that $\mathcal{W}'(t)$ is strictly decreasing and that $\mathcal{W}'(0) > 0$. An implication of the proposition is that if $\mathcal{W}'(t)$ attains zero at some $t \in [0, t_M]$, this transfer is optimal. For $Q \leq \frac{1}{2}$, $\mathcal{W}'(t)$ is continuous. Therefore, if no root exists on $[0, t_M]$, $t^* = t_M$ is optimal. For $Q > \frac{1}{2}$, $\mathcal{W}'(t)$ is discontinuous at $t = t_1$. Therefore, if no root of $\mathcal{W}'(t)$ on $[0, t_M]$ exists, $t^* = t_M$ is optimal if $\mathcal{W}'(t) > 0$ for all $t \in [0, t_M]$. $t^* = t_1$ is optimal if $\mathcal{W}'(t) \geq 0$ for all $t \in [0, t_1]$ and $\mathcal{W}'(t) < 0$ for all $t \in (t_1, t_M]$. Note that if $Q > \frac{1}{2}$ and $t^* = t_M$ resources are not exhausted since $\pi_L^* = 0$ and $\pi_H^* = 1$. In section 1.5.3 when studying the impact of e_L on optimal assignments I discuss in which settings a transfer t_M is optimal.

An important corollary of the proposition is that conducting a symmetric lottery for the goods among the two types is never optimal. This is a direct consequence of $\mathcal{W}'(0) > 0$.

Corollary 1.1. *Any assignment that assigns the same probabilities of consuming the indivisible good to both types is not a solution to the utilitarian problem.*

The corollary has an intuition that does not rely on the technical results of Proposition 1.4. Consider any assignment where all agents consume the indivisible good with the same probability. If it was part of a solution of the utilitarian problem, transfers are zero. Since the L-type's marginal utility of money is larger than the one of the H-type, the L-type has an incentive to sell a marginal probability share of consuming the good to the H-type. This holds despite the L-type may not be willing to sell the whole good to the H-type. The resulting assignment is admissible and dominates the original one such that assigning the same probabilities to both types cannot be optimal.

Comparison of second-best with first-best. The first-best assignment involves a symmetric lottery among both types and income equalization. The second-best assignment exhibits a downward distortion in the share of the L-types consuming the indivisible good and an upward distortion in the share of the H-types consuming the indivisible good. If more than the share of H-types can be served (i.e., $Q > \frac{1}{2}$), resources of the indivisible good might not be exhausted and withheld in order to generate higher transfers. The distortion in consumption shares holds despite there exists an admissible assignment with no distortion in consumption with respect to first-best.

1.5.3 Comparative Statics

In this section I discuss how the optimal assignments depend on the income e_L of the L-type, on the degree of inequality aversion, and on the benefit parameter θ.

Varying the low type income e_L. In the following fix the parameters θ and e_H. Focus of interest is how optimal consumption shares depend on the income e_L of the L-types.

Proposition 1.5 (**Dependence on e_L**)**.** *Fix θ, Q, and e_L. Let $\pi_L^*(e_L)$ denote the optimal share of L-types consuming the good in dependence of e_L. $\pi_L^*(e_L)$ is weakly increasing in e_L. Furthermore, $\pi_L^*(e_L) \to Q$ for $e_L \to e_H$.*

Proof. See appendix. □

To raise money for redistribution from the H-type, consumption shares of the L-type and the H-type need to differ. The larger the difference, the more money can be demanded from the H-type. The Proposition reflects the intuition that the lower the income of the L-type, the more he might be willing to sacrifice from his consumption utility in order to increase the monetary transfer. If his income is close to the H-type, marginal utilities in money are close for both types and the optimal transfer can already be reached by small differences in consumption probabilities. If, on the other hand, income differences are large, not all resources of the good might be distributed. It is discussed below for which settings this can occur.

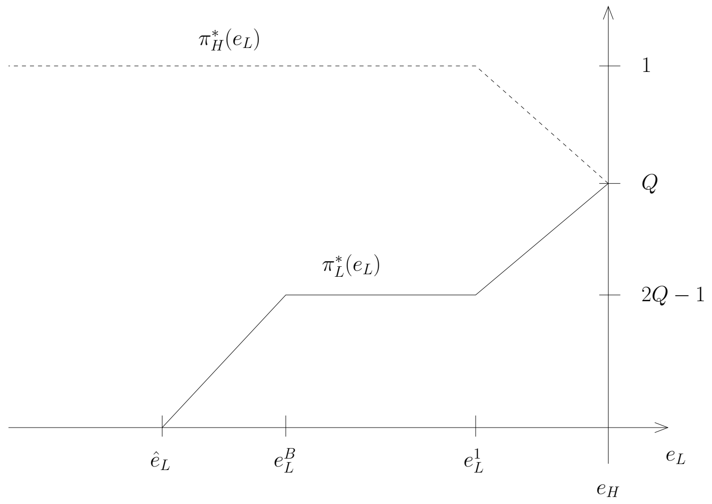

Figure 1.1: Sketch of Consumption Probabilities in Dependence of e_L for $Q > \frac{1}{2}$

Figure 1.1 illustrates the dependence of consumption probabilities on the income of the L-type for $Q > \frac{1}{2}$. The proof of the proposition reveals that if $\mathcal{W}'(t) \geq 0$ and e_L decreases, then $\mathcal{W}'(t)$ increases. When decreasing e_L and starting at $e_L = e_H$ there exist up to three critical values $\hat{e}_L < e_L^B < e_L^1$ of e_L that indicate a change in the shape of the function $\pi_L^*(e_L)$. For $e_L = e_H$ both types consume the good with probability Q. If e_L decreases up to some e_L^1, π_L^* strictly decreases while π_H^* strictly increases. e_L^1 is such that t_1 is the root of $\mathcal{W}'(t)$. Here, the optimal consumption share of the H-type is 1 and the one of the L-type is $2Q-1$. A further decrease in e_L to e_L^B does not impact on the consumption shares. This represents the interval of incomes on which $\mathcal{W}'(t)$ is positive for $t < t_1$ and negative for $t > t_1$. If e_L is lower than some e_L^B, the feasibility constraint for the good is not binding any more and resources are withheld to increase redistribution. If e_L falls below some $\hat{e}_L$, consumption probability of an L-type is zero and transfers are maximal. This is the case if e_L is such that $t = t_M$ is the unique root of $\mathcal{W}'(t)$. The critical value e_L^B exists independent of $h(\cdot)$. Whether e_L^1 and $\hat{e}_L$ exist depends on $h(\cdot)$. They particularly exist if $h'(e) \to \infty$ for $e \to -\infty$ because then $\mathcal{W}'(t_M) \geq 0$ holds if e_L is small enough.

For $Q \leq \frac{1}{2}$ the feasibility constraint with respect to the indivisible good is binding and the dependence of consumption shares on e_L is such that at $e_L = e_H$ all consume the good with equal probability Q. For decreasing e_L, π_L^* is strictly decreasing and π_H^* is strictly increasing. This holds up to a critical $\hat{e}_L$. For all $e_L < \hat{e}_L$ consumption probability of the L-type is zero and for the H-type it is $2Q$. Optimal transfers are then maximal.

Varying the Degree of Inequality Aversion. In the following I study the influence of the curvature of the welfare function $W(\cdot)$ on solutions of the utilitarian problem. I measure inequality aversion by the relative curvature $\rho_W(v) := -\frac{W''(v)}{W'(v)}$ of the welfare function (Atkinson, 1973). The higher $\rho_W(v)$, the larger the degree of inequality aversion. The following proposition shows that the higher the degree of inequality aversion, the (weakly) lower the share of L-types consuming the indivisible good and the (weakly) higher transfers are. Thereby, increasing inequality aversion results in increasing the gap in consumption probabilities to increase redistribution rather than decreasing the gap in consumption probabilities.

Proposition 1.6 (**Dependence on** ρ)**.** *Fix θ, Q, and $e_L < e_H$. Consider two welfare functions W_x and W_y with $\rho_{W_x}(v) > \rho_{W_y}(v)$ for all $v \in \mathbb{R}$. Let $\pi_L^*(W)$ denote the optimal share of L-types consuming the good in dependence of the welfare function W. Then, $\pi_L^*(W_x) \leq \pi_L^*(W_y)$.*

Proof. See appendix. □

The main idea of the proof is to show that if $\rho_{W_x}(v) > \rho_{W_y}(v)$ for all $v \in \mathbb{R}$ and $\mathcal{W}_y'(\cdot)$ is positive for some transfer t, then $\mathcal{W}_x'(\cdot)$ is positive at t as well. Since the domain $[0, t_M]$ of admissible transfers is independent of the welfare function, by Proposition 1.4, $t^*(W_x)$ is then at least as high as $t^*(W_y)$, and therefore $\pi_L^*(W_x) \leq \pi_L^*(W_y)$.

The monotonic impact of the degree of inequality aversion on optimal consumption shares makes it interesting to study the extreme ends of inequality aversion. This is a social planner that is not inequality averse with objective function $\mathcal{W}^0 = v_L + v_H$ and a Rawlsian planner with objective function $\mathcal{W}^R = v_L$. These objective functions are not included in the set of admissible objective functions yet. However, to find optimal solutions for the utilitarian problem they can be treated in the same way.

First, consider a Rawlsian planner with objective function $\mathcal{W}^R = v_L$. By the very same argumentation as seen for Propositions 1.3 and 1.4, $\pi_L^R = \pi_L(t^R)$ is part of the unique solution to the utilitarian problem if t^R maximizes $\mathcal{W}^R(t) = v_L(t)$. t^R is then the unique $t^R \in [0, t_M]$ such that $\mathcal{W}_R'(t) = v_L'(t) \geq 0$ for all $t \leq t^R$.

Second, consider a social planner that is not inequality averse with objective function $\mathcal{W} = v_L + v_H$. Again, by the same argumentation seen for Propositions 1.3 and 1.4, $\pi_L^0 = \pi_L(t^0)$ is part of a solution to the utilitarian problem if t^0 maximizes $\mathcal{W}^0(t) = v_L(t)$. t^R is then the unique $t^0 \in [0, t_M]$ such that $\mathcal{W}_0'(t) = v_L'(t) + v_H'(t) \geq 0$ for all $t \leq t^0$. However, there might exist further solutions to the utilitarian problem. This is because in contrast to the cases considered so far, the incentive constraint for the H-type is not necessarily binding. The reason is that a shift of consumption probability from the H-type to the L-type is welfare neutral. Therefore, if there is an incentive compatible assignment such that income levels are equalized and resources of the indivisible good are exhausted, the assignment maximizes welfare. This holds even if both incentive constraints are slack. Therefore, π_L^0 and t^0 represent the consumption share and transfer of the optimal solution to the utilitarian problem for which the incentive constraint of the H-type is binding.[13] In the following, when referring to the optimal solution of a social planner that is inequality averse, it is meant to be the unique optimal solution for which the incentive constraint for the H-type is binding.

By the same argumentation as seen in Proposition 1.6, the consumption probabilities for any welfare function $W(\cdot)$ are always between π_L^R and π_L^0.

Corollary 1.2. *Fix θ, Q, and $e_L < e_H$. Let π_L^0 and π_L^R be the optimal shares of L-types consuming the indivisible good for a planner that is not inequality averse and for a Rawlsian planner, respectively. For the optimal consumption share $\pi_L^*(W)$ where $W(\cdot)$ is any welfare function it holds that $\pi_L^R \leq \pi_L^*(W) \leq \pi_L^0$.*

In the following, the social planner that is not inequality averse is discussed in further detail since it gives some interesting insights about solutions for other objective functions. First, consider the case of $Q \leq \frac{1}{2}$. This implies that all resources of the good are assigned to the agents. Then, the optimal solution of the utilitarian problem is to smooth income levels as much as possible. This is because the social planner does not care about differences in consumption shares but only about differences in income. Therefore, if there is an admissible assignment with a transfer $t = \tilde{t}$ from the H-type to the L-type such that income levels are equalized, this assignment is a solution to the utilitarian problem.[14] If equalizing incomes is not admissible, the social planner chooses the transfers as large as possible, i.e. $t^* = t_M$. Particularly, if $\tilde{e_L}$ is the lowest income of the L-type such that there exists an admissible assignment that equalizes incomes, optimal transfers are maximal,

[13] Out of the set of all solutions it is the one that serves the L-type with the highest consumption probability. To see this, consider any second-best assignment such that the H-Type is not binding. Since the assignment has to be incentive compatible, $\pi_L \leq \pi_H$ and $t_L \leq t_H$ holds. Now shift the probability of consuming the indivisible good from the H-type to the L-type without influencing welfare just up to the point where the H-type is binding. This resulting assignment is then second-best as well. In particular, this also implies that whenever there is an optimal assignment such that the incentive constraint for the H-type is not binding, there exists also an optimal assignment such that it is binding.

[14] Technically, the social planner aims to maximize $\mathcal{W}_0(t) = v_L(t) + v_H(t) = 2\pi_L(t)\theta + h(e_L + t) + h(e_H + t)$ on $[0, t_M]$. Furthermore, $\mathcal{W}_0'(t) = h'(e_L + t) - h'(e_H - t)$. This term is zero if and only if $e_L + t = e_H - t$.

i.e. $t^* = t_M$, if and only if $e_L \leq \tilde{e_L}$. In combination with Proposition 1.6 the following corollary holds.

Corollary 1.3. *Fix θ and e_H. Suppose $Q \leq \frac{1}{2}$. Let $\tilde{e_L}$ be such that $\tilde{e_L} + t_M = e_H - t_M$. Then, for any welfare function $W(\cdot)$, $t^* = t_M$ and $\pi_L^*(e_L) = 0$ are optimal for all $e_L \leq \tilde{e_L}$.*

If the L-type's income is larger than $\tilde{e_L}$ it is admissible to equalize incomes. However, positive transfers from the H-type to the L-type come along with differences in consumption shares. Any social planner that is strictly inequality averse cares not only about equalizing incomes but also about inequalities in the utilities of the types. The corollary indicates that to further smooth utilities of the types an inequality averse planner rather increases transfers to $t > \tilde{t}$ than to reduce the difference in consumption probabilities. Then, the ex-post income of the L-type is larger than the one for the H-type. The H-type would prefer to sell some share of π_H to the L-type which implies that the second-best solution is not Pareto-efficient. However, the Pareto-improvement cannot be performed in an incentive compatible way since then the H-type prefers to mime the L-type.

For $Q > \frac{1}{2}$, even if it is admissible to equalize incomes, it may not be optimal for a social planner that is not inequality averse to do so. This is because on the one hand, the social planner aims to equalize incomes. On the other hand, to generate revenues for redistribution from the H-type to the L-type, the probability of consuming the indivisible good needs to be higher for the H-type than for the L-type. The larger the difference in consumption probability, the higher the realizable transfer. Depending on the setting, a transfer that equalizes incomes might only be admissible if resources Q are not exhausted. This, however, is on the cost of efficiency. Therefore, there are two conflicting desires: equalizing incomes and exhausting Q which leads to the optimal transfer being smaller than the one where incomes are equalized.[15]

Varying the Consumption Benefit θ. Varying the valuation θ for the indivisible good impacts on the analysis in two ways. On the one hand, it has an effect on the first derivative of $\mathcal{W}(t)$ and with it on the optimal transfer. On the other hand, it impacts on $\pi_L(t)$ that results from the binding incentive constraint for the H-type. An increase in θ, ceteris paribus, makes it easier to extract money from the H-type. For any fixed transfer t, an increase in θ therefore leads to an

[15]For details, let $\tilde{e}_L$ be such that for the maximal admissible transfer $t = t_M$ income levels are just equalized. It then holds that $\mathcal{W}'(t_M, \tilde{e}_L) < 0$ because $\mathcal{W}'(t_M) = v_L'(t) + v_H'(t) = -h'(e_H + t_M) - h'(e_H - t_M) < 0$ for $e_L = \tilde{e_L}$. Therefore, some $t^* < t_M$ is optimal.

increase in $\pi_L(t)$. Furthermore, the maximal transfer $t = t_M$ admissible increases when θ increases. This is in contrast to varying the L-type's income or the degree of inequality aversion since the incentive constraint for the H-type is independent of these two parameters.

The following proposition shows that if the consumption benefit θ increases, the optimal consumption share π_L^* of the L-type becomes arbitrarily close to Q. Therefore, the larger the benefit from consuming the good, the smaller the differences in consumption probabilities eventually are.

Proposition 1.7 (**Dependence on** θ)**.** *Fix any Q and $e_L < e_H$. Let $\pi_L^*(\theta)$ denote the optimal consumption probability of the L-type in dependence of the consumption benefit θ. Then, $\pi_L^*(\theta) \to Q$ for $\theta \to \infty$.*

Proof. See appendix. □

The proof is based on the idea to show the proposition for a Rawlsian planner. For a Rawlsian planner $\pi_L^*(\theta)$ is (weakly) increasing in θ and converges to Q for θ becoming arbitrarily large. Since the optimal $\pi_L^*(W)$ for any other welfare function is larger than $\pi_L^*(W_R)$ and lower than Q, the convergence result holds as well for arbitrary welfare functions. It is here more convenient to consider a Rawlsian planner than any other welfare function because for the Rawlsian planner the first derivative of $\mathcal{W}(t)$ is independent of θ. When searching for the optimal transfer t^*, θ therefore only impacts on the interval $[0, t_M]$ on which to search for the optimal transfers.

1.5.4 Comparison to Classical Approaches

Proposition 1.1 implies that the matching market approach Pareto-dominates the market price approach if the difference in income levels is small enough. The discussion of the comparative statics of the solution to the utilitarian problem reveals that the matching market approach is arbitrarily close to the utilitarian solution if the L-type's income approaches the H-type's income. The analysis of the solution to the utilitarian problem reveals that in general none of the classical approaches presented in Section 1.4 are optimal from a welfare maximizing perspective. The following intuition furthermore explains why both approaches are not ex-ante Pareto-efficient.

First, consider the market price approach where the good is sold for a market clearing price and revenues are redistributed. For $Q < \frac{1}{2}$ a share $2Q$ of the H-types buys the good and revenues are redistributed to all not buying the good.

Due to the concavity of $h(\cdot)$ this is, for instance, dominated by selling a lottery with winning probability $2Q$ to all H-types and redistributing the revenues to the L-type. For $Q > \frac{1}{2}$, by the same argument welfare increases when offering the participation in lotteries for the good instead of selling the good for a market clearing price. Only for $Q = \frac{1}{2}$ and income differences being large enough, a utilitarian planner might sell a lottery with winning probability of one to the H-types and redistribute the generated revenues to the L-types.

Second, consider the matching market approach that conducts a symmetric lottery for the goods and does not involve transfers. This is not optimal since the L-type has an incentive to sell at least a marginal consumption share to the H-type such that expected utilities of both types increase. This is because the compensation the L-type is willing to accept for a marginal loss in consumption probability is lower than the willingness to pay of the H-type for a marginal increase in consumption probability. It holds despite the L-type might not be willing to sell a (full) good that he received to the H-type.

The use of random assignment without monetary transfers in many real-world applications suggests that these assignments have some attractive features a classical welfare maximizing approach cannot reflect. In our setting conducting a symmetric lottery for the good is the only assignment that fully reflects the desire of income independent access. Once endowments are private information, it implies that no redistribution from the high income types to the low income types can be enforced. Notably, the first-best assignment of the utilitarian problem incorporates an income independent consumption of the good as well. The reason is that a social planner who is inequality averse cares about equalization of incomes as well as equality in consumption shares. Information asymmetries then lead to a distortion in the second-best assignment. A preference for income-independent access conflicts with a preference for redistribution. Although in a private information setting randomly assigning the good is admissible, solutions of the utilitarian problem sacrifice income independent consumption as faced in first-best in order to increase revenues for redistribution.

Nevertheless, comparing consumption shares of the optimal assignment with consumption shares of the market price approach, the optimal assignment may exhibit less distortion in consumption probabilities than the market price approach does. For $Q \leq \frac{1}{2}$, using the market price approach implies that none of the agents with low income consumes the indivisible good. For the L-type's income e_L or the benefit parameter θ being large enough, however, the welfare maximizing assignment involves a positive consumption share for the low income types. For

$Q > \frac{1}{2}$, using the market price approach implies that a share $2Q - 1$ of the L-types consumes the good. Again, by the same argument, there are parameter ranges such that the optimal assignment exhibits less distortion. However, since the optimal assignment may not exhaust resource if income differences are large enough, the optimal assignment might even assign less of the good to the L-types than the market price approach does in order to increase redistribution.

1.5.5 Availability of Income Information

For the preceding analysis of the utilitarian problem, the participation constraint could be neglected: any assignment that maximizes welfare within the set of all feasible and incentive compatible assignments is individual rational for both types as well. The participation constraint is of interest if the social planner is informed about endowments. Then, the incentive compatibility constraints are replaced by the participation constraints

$$\pi_L\theta + h(e_L - t_L) \geq h(e_L) \quad \text{and} \quad \pi_H\theta + h(e_H - t_H) \geq h(e_H) \tag{1.15}$$

In such a setting it is straight forward to argue that the feasibility constraints are both binding since the planner does not need to care about incentives to report the true type. Furthermore, transfers do not depend on randomization since dissolving any uncertainty in randomization has only a positive effect on welfare and no potentially negative effect on any constraint.

Remember that the first-best assignment is such that both types consume the indivisible good with probability Q and the optimal transfer $\tilde{t}$ is such that both have ex-post the same wealth level, i.e. $e_L + \tilde{t} = e_H - \tilde{t}$. If the first-best assignment does not violate the individual rationality constraint it is therefore optimal. If it violates the incentive constraint the social planner faces a trade-off between equalizing income and equalizing consumption shares. The following discussion shows that for a welfare maximizing social planner equalizing income has priority since the optimal assignment involves transfers that are as close as possible to equalizing incomes. The curvature of the welfare function will not play any role for the optimal assignment.

To describe the optimal assignment in this setting first note that the participation constraint for the H-type is binding as long as first-best cannot be implemented. This is because if first-best cannot be implemented, for any optimal assignment either the ex-post income of the H-type is larger than the one for the L-type or the consumption probability of the H-type is larger than the one for the L-type. Then either income or consumption probability can be shifted from the H-type to the L-type without violating individual rationality. Therefore, $\pi_H^*\theta + h(e_H - t^*) = h(e_H)$ in optimum. With $\pi_L^* = 2Q - \pi_H^*$ it implies that

$$\pi_L^* = \pi_L(t^*) = 2Q - \frac{1}{\theta}[h(e_H) - h(e_H - t^*)].$$

Let t_M be the maximum transfers implementable, i.e., t_M is such that $h(e_H) = \min\{2Q, 1\}\theta + h(e_H - t_M)$. The maximization problem then reduces to finding $t^* \in [0, t_M]$ such that it maximizes

$$\mathcal{W}(t) = W(h(e_H)) + W(\pi_L(t)\theta + h(e_L + t)).$$

The first derivative is

$$\mathcal{W}'(t) = W'(\pi_L(t)\theta + h(e_L + t))[h'(e_L + t) - h'(e_H - t)].$$

It is straight forward to see that $\mathcal{W}'(0) > 0$ and $\mathcal{W}'(t)$ is strictly decreasing in t. Furthermore, $\mathcal{W}'(t) = 0$ if and only if $e_L + t = e_H - t$. Let $\tilde{t}$ be such that $e_L + \tilde{t} = e_H - \tilde{t}$. For the optimal transfer t^* it therefore holds $t^* = \tilde{t}$ if $\tilde{t} \leq t_M$ and $t^* = t_M$ otherwise. Therefore, within the set of feasible and individually rational assignment, the welfare maximizing one is the one that includes a transfer t^* that is the closest to equalizing incomes. If transfers are such that incomes are equalized, the corresponding share π_L^* of L-types is Q, if first-best can be implemented. In all other cases π_L^* is such that the participation constraint of the H-type is binding. In particular, if the optimal transfer t^* is such that incomes are not equalized, the consumption share of the L-type is minimal and therefore either zero (if $Q \leq \frac{1}{2}$)) or $2Q - 1$ (if $Q > \frac{1}{2}$). The intuition behind for the priority of redistribution is that whenever the ex-post income of the L-type is lower than the H-type, it is a Pareto-improvement that the L-type sells a small part of p_L to the H-type. The following proposition summarizes the analysis above.

Proposition 1.8. *Fix any θ and $e_L < e_H$. Let $\tilde{t}$ be defined by $e_L + \tilde{t} = e_H - \tilde{t}$ and t_M be defined by $h(e_H - t_M) - h(e_H) = \min\{2Q, 1\}\theta$. Furthermore,*

$$\pi_L(t) = 2Q - \frac{1}{\theta}[h(e_H) - h(e_H - t)].$$

The assignment $(\pi_L^*, \pi_H^*, -t^*, t^*)$ that maximizes welfare within the set of feasible and individually rational assignments is such that

- *$\tilde{t} \leq t_M$ implies $t^* = \tilde{t}$, $\pi_L = \min\{\pi_L(t), Q\}$, and $\pi_H^* = 2Q - \pi_L^*$*
- *$\tilde{t} > t_M$ implies $t^* = t_M$, $\pi_L = \max\{2Q - 1, 0\}$, and $\pi_H^* = \min\{2Q, 1\}$.*

Proof. The proof is implied by the preceding discussion. □

1.6 Conclusion

In the literature dealing with the assignment of resources it is standard practice to assume that agents have preferences that are linear in income. An agent's willingness to pay then fully reflects his benefit of consuming the good. In many economic environments, however, preferences of agents exhibit wealth effects. In particular, increasing wealth may increase the willingness to pay. The willingness to pay then does not any more perfectly reflect the benefit the agents derive from consuming the good. Consequences for a social planner caring for welfare maximization might therefore crucially depend on assumptions about wealth effects.

This paper analyses the provision of an indivisible good of limited availability when agents' preferences regarding money are strictly concave. Therefore, not only the benefit of consumption but also an agent's income has an impact on the willingness to pay. I concentrate on a setting where agents differ in their income but not in the benefit they derive from consuming the good. Randomly assigning the good then might dominate selling the good for a market clearing price. However, none of them is optimal from a welfare maximizing perspective. Any solution to the utilitarian problem involves selling probability shares of the indivisible good and redistributing the revenues. Consumption probabilities of agents depend on their income and are distorted compared to first-best where both types receive the indivisible good with the same probability. The optimal assignment may not distribute all resources of the good in order to increase revenues for redistribution.

The model explains on grounds of efficiency that in certain settings, randomly assigning a good without using transfers should be preferred to selling the good for a market clearing price. It is particularly relevant for markets where the benefit from consuming the good is high which is presumably the case for scarce medical resources like organs or places at good schools and universities. However, any welfare maximizing solution in a private information setting implies that consumption probabilities of the indivisible good depend on income. Whenever income independent access to the good is desired, the welfare maximizing approach never fully reflects this desire. Nevertheless, the welfare maximizing assignment might admit a greater number of low income types for the consumption of the good than the assignment implied by a market clearing price approach.

1.7 Appendix

Proof of Proposition 1.1

I consider the cases $Q \leq G_H$ and $Q > G_H$ separately.

$Q \leq G_H$: The market clearing price t_{MP} is such that

$$\theta + h(e_H - t_{MP}) = h(e_H + \frac{Q}{1-Q} t_{MP}).$$

I first show that independent of e_L, the H-type prefers the random assignment over the assignment resulting from the setting the market clearing price. The equation defining t_{MP} implies that $(e_H + \frac{Q}{1-Q} t_{MP}) - (e_H - t_{MP}) = \frac{1}{1-Q} t_{MP}$ is the willingness to pay for receiving the good with probability one when facing a wealth level of $e_H + \frac{Q}{1-Q} t_{MP}$. Due to the strict concavity of $h(\cdot)$, the willingness to pay at this wealth level for receiving the indivisible good with probability Q is then larger than $Q \frac{1}{1-Q} t_{MP}$. Therefore,

$$Q\theta + h(e_H) > h(e_H + \frac{Q}{1-Q} t_{MP}).$$

This implies that the H-type prefers the random assignment over the good being sold for a market clearing price. Therefore, the H-type is always better off with the random assignment.

The L-type's valuation of the random assignment is $Q\theta + h(e_L)$, for the market clearing price it is $h(e_L + \frac{Q}{1-Q} t_{MP})$. If e_L converges to e_H, the valuations of both assignments converge to the valuations of the H-type. Since the H-type strictly prefers the random assignment over the market clearing price assignment, whenever e_L is large enough, the L-type prefers it as well.

$Q > G_H$: Here, the market clearing price t_{MP} is such that

$$\theta + h(e_L - t_{MP}) = h(e_L + \frac{Q}{1-Q} t_{MP}).$$

With exactly the same argument as for $Q \leq G$ and the H-type, the L-type always prefers the lottery to paying the market clearing price for the good. It remains to show that for e_L large enough, the H-type prefers the lottery as well. Analogously to above, e_L converging to e_H results in the valuations of the H-type of the two assignment procedures approaching the valuations of the L-type. Since the L-type always strictly prefers the random assignment over the market clearing price, the H-type prefers it as well whenever e_L is close enough to e_H.

Proof of Lemma 1.1

It is to show that if $(\pi_L^*, \pi_H^*, t_L^*(r), t_H^*(r))$ solves the utilitarian problem, then $t_L^*(r)$ and $t_H^*(r)$ are both independent of r. This is shown in two steps. First, I show that $t_L^*(r)$ is independent of r, then I show that $t_H^*(r)$ is independent of r. Since it should be clear from the context, in the following I omit the asterisk * for denoting a solution of the utilitarian problem.

1. $t_L(r) = t_L \in \mathbb{R}$**:** In the following I show that if the optimal transfer $t_L(r)$ depends on r it can be replaced by a constant transfer without violating the feasibility and incentive compatibility constraint such that the L-type is strictly better off under the new assignment while the H-type is indifferent. The new assignment is then individually rational as well because the original one was. This contradicts the assumption that $t_L(r)$ is part of the optimal solution. Therefore, $t_L(r)$ cannot depend on r.

Suppose that $t_L(r)$ depends on r. It therefore can be expressed as $t_L(r) = t_L + \epsilon(r)$ with $t_L \in \mathbb{R}$ and $\int \epsilon(r)d\nu(r) = 0$. Let $k_H > 0$ be the unique solution of the equation

$$\theta\pi_L + \int h(e_H - t_L - \epsilon(r))d\nu(r) = \theta\pi_L + h(e_H - t_L - k_H).$$

k_H is thus chosen such that the H-type is indifferent between the assignments $(\pi_L, t_L(r))$ and $(\pi_L, t_L - k_H)$ and can be interpreted as the willingness to pay of the H-type to avoid the lottery in the assignment of the L-type. It remains to show that replacing the assignment $(\pi_L, t_L(r))$ of the L-type by $(\pi_L, t_L - k_H)$ is feasible, incentive compatible and a strict welfare improvement.

First, it is feasible since the budget of the new assignment is even lower than the budget of the old assignment. Second, it is incentive compatible: For the H-type the value of both assignments is the same. Therefore, if he preferred his assignment to the one of the L-type before, he still does so. For the L-type, the value of the new assignment increases. This is because $h''' > 0$ implies decreasing absolute risk aversion and therefore, the L-type is willing to pay even more than k_H to avoid the lottery in $t_L(r)$. Therefore, the L-type strictly prefers $(\pi_L, t_L - k_H)$ to $(\pi_L, t_L(r))$ and thereby also strictly prefers it to (π_H, t_H). This furthermore implies that the L-type is strictly better off under the new assignment while the H-type is indifferent. Therefore, replacing $(\pi_L, t_L(r))$ by $(\pi_L, t_L - k_H)$ is an admissible welfare improvement.

2. $t_H(r) = t_H \in \mathbb{R}$**:** In analogy to the first part of the proof, assume that $t_H(r)$ depends on r. The proof is then completed if this implies that $t_H(r)$ cannot be optimal since there exists an admissible assignment that is welfare improving.

If $t_H(r)$ depends on r it can be expressed by $t_H(r) = t_H + \epsilon(r)$ with $t_H \in \mathbb{R}$ and $\int \epsilon(r) d\nu(r) = 0$. Replacing $t_H(r)$ in the assignment of the H-type by t_H strictly improves the value of this assignment to him. Therefore, he strictly prefers his assignment to the one of the L-type. This allows to define $x > 0$ as the unique solution of the equation

$$\theta\pi_L + h(e_H - t_L + x) = \theta\pi_H + h(e_H - t_H - x) \tag{1.16}$$

Consider now the two assignments $(\pi_L, t_L - x)$ and $(\pi_H, t_H + x)$. These assignments are feasible since the budget is the same as it was for the initial assignment. In the following, I furthermore argue that assignments are incentive compatible and increase total welfare. I distinguish two cases.

Case 1: $\pi_L \leq \pi_H$**.** Rearranging equation (1.16) yields

$$h(e_H - t_L + x) - h(e_H - t_H - x) = \theta\pi_H - \theta\pi_L > 0.$$

Due to the strict concavity of $h(\cdot)$, holds that

$$h(e_L - t_L + x) - h(e_L - t_H - x) > \theta\pi_H - \theta\pi_L > 0.$$

Therefore, the H-type is indifferent among the assignments $(\pi_L, t_L - x)$ and $(\pi_H, t_H + x)$, while the L-type strictly prefers the $(\pi_L, t_L - x)$. This implies incentive compatibility. In the following I argue that both types are better off with the new assignment such that it increases welfare and furthermore satisfies the incentive compatibility constraint because the original assignment did. For the L-type it is directly implied by $x > 0$ that he is strictly better off by the assignment $(\pi_L, t_L - x)$ compared to (π_L, t_L). To see that the H-type is better off as well, note that the incentive compatibility constraint of the H-type is binding in this case. This is because if it is not binding, a marginal shift in consumption probability from the H-type to the L-type keeps the sum of valuations constant but reduces the gap of the L-type's and the H-type's valuations.[16] Therefore, the original assignment has a value of $\pi_L\theta + h(e_H - t_L)$ to the H-type. The new assignment serves him a value of at least $\pi_L\theta + h(e_H - t_L + x)$ which implies that he is strictly better off.

[16] $\pi_L \leq \pi_H$ implies that $\pi_H^* > 0$ because otherwise no resources are distributed which cannot be optimal. The detailed argument for why the incentive constraint of the H-type is binding is the same as used in the proof of Lemma 1.2.

Case 2: $\pi_L > \pi_H$. Rearranging equation (1.16) yields

$$h(e_H - t_H - x) - h(e_H - t_L + x) = \theta\pi_L - \theta\pi_H > 0.$$

Due to the strict concavity of $h(\cdot)$, it holds that

$$h(e_L - t_H - x) - h(e_L - t_L + x) > \theta\pi_H - \theta\pi_L > 0.$$

Therefore, a new assignment where the H-type receives $(\pi_L, t_L - x)$ and the L-type receives $(\pi_H, t_H + x)$ is incentive compatible. To show that this is also a welfare improvement, first note that the L-type originally faces a value of $\pi_L\theta + h(e_L - t_L)$. His new assignment has a value to him that is strictly higher than $\pi_L\theta + h(e_L - t_L + x)$. Therefore, he is better off and the utility gain is at least $h(e_L - t_L + x) - h(e_L - t_L)$. The H-type originally faced a value of $\pi_H\theta + \int h(e_H - t_H - \epsilon(r))d\nu(r)$. The value of the new assignment is as high as $\pi_H\theta + h(e_H - t_H - x)$. Therefore, the (potential) loss he faces is at most $h(e_H - t_H) - h(e_H - t_H - x)$. It is to show that the minimum gain exceeds the maximum loss, i.e., $h(e_L - t_L + x) - h(e_L - t_L) > h(e_H - t_H) - h(e_H - t_H - x)$. By concavity of $h(\cdot)$ and the differences in incomes being x on both sides of the inequality, it is sufficient to show that $e_H - t_H > e_L - t_L + x$. Equation (1.16) which defines x implies $e_H - t_H > e_H - t_L + x$ and this is larger than $e_L - t_L + x$. Therefore, the new assignment makes the L-type strictly better off and the sum of utilities for both types increases which implies that it is a welfare improvement. It remains to show that the new assignment is also individually rational. It is individually rational for the L-type since he is better off by the new assignment than he was before. If $\pi_L + h(e_L - t_L) \geq h(e_L)$ holds, it also holds that $\pi_L + h(e_L - t_L) \geq h(e_H)$ due to the strict concavity of $h(\cdot)$. Therefore, the new assignment is individually rational for the H-type as well.

Note that this analysis also implies that $\pi_L > \pi_H$ is never optimal.

Proof of Lemma 1.2

I first proof the characteristics of optimal assignments and then show that the participation constraint can be neglected.

Part 1. The second part proof of Lemma 1.1 implies that for any optimal solution it has to hold that $\pi_L^* \leq \pi_H^*$ because otherwise there is an admissible assignment where the H-type receives a consumption probability of π_L^* and the L-type a consumption probability of π_H^* that yields higher welfare. Furthermore, if $\pi_L^* \leq \pi_H^*$, incentive compatibility implies $t_L^* \leq t_H^*$. Incentive compatibility combined with $e_L < e_H$ furthermore implies that $v_L^* < v_H^*$.

Part 2. Let $(\pi_L^*, \pi_H^*, t_L^*, t_H^*)$ be a solution to the utilitarian problem. Suppose the incentive compatibility constraint for the H-type is not binding. By Part 1 of the lemma, $\pi_L^* \leq \pi_H^*$ holds. $\pi_H^* > 0$ holds because otherwise all transfers have to be zero and consumption probabilities have to be zero which contradicts the assumption that the H-type strictly prefers her assignment to the one of the L-type. Furthermore, $\pi_L^* \leq \pi_H^*$ and $Q < 1$ implies that $\pi_L^* < 1$. Now shift some consumption probability of π_H^* from the H-type to the L-type. Since the incentive compatibility constraint of the H-type is slack, this can be performed in an incentive compatible way. The L-type's utility increases. Therefore, the new assignment is individually rational for the L-type as well. It is therefore also individually rational for the H since the H-type likes his assignment at least as much as the one of the L-type: if $\pi_L + h(e_L - t_L) \geq h(e_L)$ holds, it also holds that $\pi_L + h(e_L - t_L) \geq h(e_H)$ due to the strict concavity of $h(\cdot)$. The value $v_L + v_H$ is not affected by the shift but the inequality in valuation decreases. Since the welfare function is strictly concave, total gains are strictly positive which contradicts the assumption that $(\pi_L^*, \pi_H^*, t_L^*, t_H^*)$ is an optimal assignment within the set of admissible assignments.

Part 3. Let $(\pi_L^*, \pi_H^*, t_L^*, t_H^*)$ be a solution to the utilitarian problem. Assume Part 3 of the lemma does not hold such that $t_L^* + t_H^* > 0$. The Lemma is proved if there exists an admissible assignment that yields higher welfare. By Part 2, the incentive compatibility constraint for the H-type is binding. If it is binding for the L-type as well, it has to hold that

$$h(e_H - t_L^*) - h(e_H - t_H^*) = h(e_L - t_L^*) - h(e_L - t_H^*) = \theta(\pi_H^* - \pi_L^*).$$

Since $h(\cdot)$ is strictly concave it implies that $t_L^* = t_H^*$. Then it is admissible to reduce both payments to zero which is a Pareto-improvement and contradicts the assumption of t_L^* and t_H^* being part of an optimal assignment. If the constraint for the L-type is not binding, t_H^* can be decreased without violating any constraint. This is a Pareto-improvement which again contradicts the assumption that t_L^* and t_H^* are part of an optimal assignment.

Part 4. Assume that for any $Q < 1$ the feasibility constraint with respect to the indivisible good is not binding. Now assume that for any solution of the utilitarian problem $\pi_H^* < 1$ holds. Since $\pi_L^* \leq \pi_H^*$ the consumption utility of both types can be raised by some $\Delta > 0$ such that $\pi_H^* + \Delta \leq 1$ and $\pi_L^* + \pi_H^* + 2\Delta \leq Q$. The resulting assignment is still admissible since both types value the gain of both assignments in the same way. This contradicts the assumption of $\pi_H^* < 1$

and therefore $\pi_H^* = 1$ has to hold if the feasibility constraint with respect to the indivisible good is not binding.

Participation Constraint. Consider any assignment that maximizes welfare within the set of feasible and incentive compatible assignments. All results on optimal assignments derived so far hold as well if there is no participation constraint. By the other results of this Lemma that are proved above, the assignment of the L-type is such that transfers from the H-type are larger or equal to zero and his consumption probability of the good is larger or equal to zero as well. Therefore, his participation constraint is satisfied. Since H-types like their assignments at least as much as the one of the L-types, his participation constrained is satisfied as well.

Proof of Proposition 1.4

To show Proposition 1.3, I consider the first and second derivatives $\mathcal{W}'(t)$ and $\mathcal{W}''(t)$. I will show that $\mathcal{W}'(t)$ is strictly decreasing for all $t^* \in [0, t_M]$ and that $\mathcal{W}'(0) > 0$. This then implies that $\mathcal{W}(t)$ has a unique maximizer $t^* \in [0, t_M]$ and that t^* is such that $\mathcal{W}'(t) \geq 0$ if and only if $t \leq t^*$.

First consider the case that more than only H-types can be served, i.e., $Q > G_H$. Based on this analysis the case for $Q \leq G_H$ will be straight forward. According to equation (1.10),

$$\mathcal{W}(t) = W(\pi_L(t)\theta + h(e_L + t)) + W(\pi_L(t)\theta + h(e_H + t)).$$

The explicit expression for $\pi_L(t)$ depends on whether $t \in [0, t_1]$ or $t \in (t_1, t_M]$ (see discussion preceding Proposition 1.4 and particularly equations (1.13) and (1.14)). At $t = t_1$ there is a kink in $\mathcal{W}(t)$ and a point of discontinuity in $\mathcal{W}'(t)$. To show that $\mathcal{W}'(t)$ is strictly decreasing on $[0, t_M]$ I show that on each domain it holds that $\mathcal{W}'(t)$ is strictly decreasing in t, and that at $t = t_1$, $\mathcal{W}'(t)$ decreases as well.

$\mathcal{W}'(\mathbf{t})$ is strictly decreasing on $[\mathbf{0}, \mathbf{t_1}]$ and on $(\mathbf{t_1}, \mathbf{t_M}]$. Using $v_L(t) = \pi_L(t)\theta + h(e_L + t)$ and $v_H(t) = \pi_L(t)\theta + h(e_H + t)$, the first and second derivatives of $\mathcal{W}(t)$ on each domain are

$$\begin{aligned}
\mathcal{W}'(t) &= W'(v_L(t))v_L'(t) + W'(v_H(t))v_H'(t) \\
\mathcal{W}''(t) &= W''(v_L(t))(v_L'(t))^2 + W'(v_L(t))v_L''(t) + W''(v_H(t))(v_H'(t))^2 + W'(v_H(t))v_H''(t)
\end{aligned}$$

In the following I show that $\mathcal{W}''(t) < 0$ for all $t \in [0, t_1]$ and $\mathcal{W}''(t) < 0$ for all $t \in (t_1, t_M]$. For this, it is sufficient to show that $v_H''(t) \leq 0$ on each domain. This is because, first, it implies that $v_L''(t) \leq 0$ as well. This is a consequence of the assumption that $h''' > 0$: for $h''' > 0$, it holds that $h''(e_L + t) < h''(e_H + t)$, and therefore, $v_L''(t) = \pi_L''(t)\theta + h''(e_L + t) < \pi_L''(t)\theta + h''(e_H + t) = v_H(t)$. Second, it holds that $W''(v_L(t))(v_L'(t))^2 < 0$, $W''(v_H(t))(v_H'(t))^2 < 0$, $W'(v_L(t)) > 0$, and $W'(v_H(t)) > 0$. Therefore, if $v_H''(t) \leq 0$ on both domains, each summand of the equation for $\mathcal{W}''(t)$ is smaller than zero. This implies that $\mathcal{W}''(t) < 0$ on both domains.

$t \in [0, t_1]$: For $t \in [0, t_1]$ it holds that

$$\pi_L(t) = \underline{\pi}_L(t) = Q - \frac{1}{2\theta}(h(e_H + t) - h(e_H - t))$$

Therefore,

$$\begin{aligned} v_H''(t) &= \pi_L''(t)\theta + h''(e_H + t) && (1.17)\\ &= -\frac{1}{2\theta}\theta[h''(e_H + t) - h''(e_H - t)] + h''(e_H + t) && (1.18)\\ &= \frac{1}{2}[h''(e_H + t) + h''(e_H - t)] && (1.19) \end{aligned}$$

Since $h'' < 0$, for all $t \in [0, t_1]$ it holds that $v_H''(t) < 0$.

$t \in (t_1, t_M]$: For $t \in (t_1, t_M]$ it holds that

$$\pi_L(t) = \overline{\pi}_L(t) = 1 - \frac{1}{\theta}(h(e_H + t) - h(e_H - t))$$

It implies

$$\begin{aligned} v_H''(t) &= \pi_L''(t)\theta + h''(e_H + t) && (1.20)\\ &= -\frac{1}{\theta}\theta[h''(e_H + t) - h''(e_H - t)] + h''(e_H + t) && (1.21)\\ &= h''(e_H - t). && (1.22) \end{aligned}$$

Therefore, $v_H''(t) < 0$ by the assumption $h'' < 0$.

$\mathcal{W}'(\mathbf{t})$ is strictly decreasing on $[\mathbf{0}, \mathbf{t_M}]$. To evaluate the information received from considering the restricted domains, I now go back to considering the characteristics of $\mathcal{W}(t)$ on the whole domain of $[0, t_M]$. Since $\pi_L(t)$ is continuous on $[0, t_M]$, $\mathcal{W}(t)$ is continuous $[0, t_M]$ as well. Furthermore, $\mathcal{W}'(t)$ is continuous and strictly decreasing on $t \in [0, t_1]$ and it is continuous and strictly decreasing on $t \in (t_1, t_M]$. To show that $\mathcal{W}'(t)$ is strictly monotonically decreasing on the whole domain $[0, t_M]$ it is to show that $\lim_{t \nearrow t_1} \mathcal{W}'(t) > \lim_{t \searrow t_1} \mathcal{W}'(t)$.

$$
\begin{aligned}
\lim_{t\nearrow t_1}\mathcal{W}'(t)-\lim_{t\searrow t_1}\mathcal{W}'(t) &= W'(v_L(t_1))[\underline{\pi}'_L(t_1)+h'(e_L+t_1)]\\
&+ W'(v_H(t_1))[\underline{\pi}'_L(t_1)+h'(e_H+t_1)]\\
&- W'(v_L(t_1))[\overline{\pi}'_L(t_1)+h'(e_L+t_1)]\\
&- W'(v_H(t_1))[\overline{\pi}'_L(t_1)+h'(e_H+t_1)]\\
&= [W'(v_L(t_1))+W'(v_H(t_1))][\underline{\pi}'_L(t_1)-\overline{\pi}'_L(t_1)]
\end{aligned}
$$

The equations use the continuity of $\pi_L(t)$ at $t = t_1$. The sign of $W'(v_L(t_1)) + W'(v_H(t_1))$ is positive, so it remains to show that the sign of $\underline{\pi}'_L(t_1) - \overline{\pi}'_L(t_1)$ is positive as well. This directly follows by considering the equations for $\underline{\pi}'_L(t_1)$ and $\overline{\pi}'_L(t_1)$:

$$
\begin{aligned}
\underline{\pi}'_L(t_1) &= -\frac{1}{2\theta}[h'(e_H+t)+h'(e_H-t)]\\
\overline{\pi}'_L(t_1) &= -\frac{1}{\theta}[h'(e_H+t)+h'(e_H-t)]
\end{aligned}
$$

Therefore, $\underline{\pi}'_L(t_1) > \overline{\pi}'_L(t_1)$ which shows that $\lim_{t\nearrow t_1}\mathcal{W}'(t) > \lim_{t\searrow t_1}\mathcal{W}'(t)$.

$\mathcal{W}'(\mathbf{0}) > \mathbf{0}$ holds. For the first derivative of $v_H(t)$ it holds that $v'_H(0) = \pi'_L(0)\theta + h'(e_H) = h'(e_H) > 0$. Furthermore

$$v'_L(t) = \pi'_L(t)\theta + h'(e_L+t) > v'_H(t) = \pi'_L(t)\theta + h'(e_H+t).$$

Then, $\mathcal{W}'(0) = W'(v_L(0))v'_L(0) + W'(v_H(0))v'_H(0) > 0$ since each summand is strictly larger than zero.

Case $\mathbf{Q \leq G_H}$. Considering $Q \leq G_H$ simplifies the analysis since $\pi_L(t) = Q - \frac{1}{2\theta}(h(e_H+t) - h(e_H-t))$ for all $t \in [0, t_M]$. This corresponds to the equation for $\pi_L(t)$ on $[0, t_1]$ if $Q > G_H$. Therefore it holds that $\mathcal{W}'(0) > 0$ and $\mathcal{W}'$ is strictly monotonically decreasing on $[0, t_M]$.

Proof of Proposition 1.5

To show that $\pi^*_L(e_L)$ is increasing in e_L it is sufficient to show that the corresponding transfers $t^*(e_L)$ are decreasing in e_L. The maximal transfer t_M that can be demanded from each H-type does not depend on e_L. Therefore, the domain of transfers on which to search the optimal transfers is constant in e_L. Let $\mathcal{W}_{e_L}(t)$ indicate the objective function to maximize if the L-type's income is e_L. By Proposition 1.4, for any e_L, $t^*(e_L)$ is the unique $t^* \in [0, t_M]$ such that $\mathcal{W}'_{e_L}(t) \geq 0$

if and only if $t \leq t^*$. I first show that $\mathcal{W}'_{e_L}(t) \geq 0$ implies that for $e'_L < e_L$, $\mathcal{W}'_{e'_L}(t) > 0$ holds.

Suppose that for any $t \in [0, t_M]$, $\mathcal{W}'_{e_L}(t) \geq 0$. It holds that

$$\mathcal{W}'_{e_L}(t) = W'_{e_L}(v_L(t))v'_L(t) + W'_{e_L}(v_H(t))v'_H(t)$$

with $v_L(t) = \theta\pi_L(t) + h(e_L + t)$ and $v_H(t) = \theta\pi_L(t) + h(e_H + t)$. Since $W'_{e_L} > 0$ and $v'_L(t) > v'_H(t)$ it has to hold that $v'_L(t) > 0$ whenever $\mathcal{W}'_{e_L}(t) \geq 0$. Furthermore note that $\pi_L(t)$ is independent of e_L since it is defined by the binding incentive constraint of the H-type. Now consider how $\mathcal{W}'_{e_L}(t)$ changes when lowering the income of the L-type from e_L to e'_L. It implies that $v_L(t)$ decreases and therefore $W'(v_L(t))$ increases. Furthermore, $v'_L(t)$ increases since $h'(e_L + t)$ increases. $W'(v_H(t))v'_H(t)$ does not depend on e_L. $v'_L(t) > 0$ implies that $\mathcal{W}'_{e_L}(t)$ increases for decreasing e_L and $\mathcal{W}'_{e'_L}(t) > 0$ holds for $e'_L < e_L$. This completes the first part of the proof: for e_L, the optimal transfer satisfies $\mathcal{W}'_{e_L}(t^*(e_L)) \geq 0$. For some e'_L it then holds as well that $\mathcal{W}'_{e'_L}(t^*(e_L)) \geq 0$. Therefore, if $t^*(e_L)$ is optimal for e_L, the transfer $t^*(e'_L)$ that is optimal for $e'_L < e_L$ is at least as large as $t^*(e_L)$.

For $e_L \to e_H$, $\mathcal{W}'_{e_L}(t)$ converges to $\mathcal{W}'_{e_H}(t) = 2W'(v_H(t))v'_H(t)$. It holds that $W'(v_H(t)) \geq 0$ for any transfers and furthermore $v'_H(t) = \frac{1}{2}(h'(e_H + t) - h'(e_H - t)) \leq 0$ for all $t \in [0, t_M]$ with $v'_H(t) = 0$ if and only if $t = 0$. Therefore, $t^* = 0$ is optimal which implies $\pi^*_L = \pi^*_H = Q$.

Note that optimal transfers $t^*(e_L)$ are not necessarily strictly increasing in e_L. However, the proposition shows that they are strictly increasing whenever $t^*(e_L) \notin \{t_1, t_M\}$. This is because for $Q > \frac{1}{2}$, $\mathcal{W}'(t)$ is not continuous at $t = t_1$. Thereby, the conversion of signs might occur at $t = t_1$ and small variations of e_L do not have an impact on it. Furthermore, once $t^* = t_M$ is optimal for some e_L it is also optimal for all $e'_L < e_L$.

Proof of Proposition 1.6

Consider W_x and W_y such that $\rho_{W_x}(v) > \rho_{W_y}(v)$ for all $v \in \mathbb{R}$. This particularly implies that $\frac{W'_x(v)}{W'_y(v)}$ is strictly monotonically decreasing in v. To see this, consider the first derivative with respect to v of $\frac{W'_x(v)}{W'_y(v)}$:

$$\frac{d}{dv}\frac{W'_x(v)}{W'_y(v)} = \frac{W''_x(v)W'_y(v) - W''_y(v)W'_x(v)}{(W'_y(v))^2}$$

Since $\rho_{W_x}(v) > \rho_{W_y}(v)$ is equivalent to $W''_x(v)W'_y(v) < W''_y(v)W'_x(v)$ the expression above is smaller than zero and therefore $\frac{W'_x(v)}{W'_y(v)}$ is strictly monotonically decreasing in v. This will be used to show that $\pi^*_L(W_x) \leq \pi^*_L(W_y)$.

Showing $\pi_L^*(W_x) \leq \pi_L^*(W_y)$ is equivalent to showing that for the corresponding optimal transfers $t^*(W_x) \geq t^*(W_y)$ holds. Since it is known that $\mathcal{W}'(t)$ is strictly monotonically decreasing in t it is sufficient to show that whenever $\mathcal{W}_y'(t) \geq 0$, it holds that $\mathcal{W}_x'(t) > 0$. Then, if t^* is optimal for W_y the optimal transfer is at least as large as t^* for W_x. It holds that

$$\mathcal{W}_y'(t) \geq 0 \;\Leftrightarrow\; W_y'(v_L(t))v_L'(t) + W_y'(v_H(t))v_H'(t) \geq 0 \tag{1.23}$$

$$\Leftrightarrow\; v_H'(t) \geq -\frac{W_y'(v_L(t))v_L'(t)}{W_y'(v_H(t))} \tag{1.24}$$

Now have a closer look at $\mathcal{W}_x'(t)$. Using the above boundary for $v_H'(t)$ it holds that

$$\mathcal{W}_x'(t) = W_x'(v_L(t))v_L'(t) + W_x'(v_H(t))v_H'(t) \tag{1.25}$$

$$\geq W_x'(v_L(t))v_L'(t) - \frac{W_x'(v_H(t))W_y'(v_L(t))v_L'(t)}{W_y'(v_H(t))} \tag{1.26}$$

Therefore, $\mathcal{W}_x'(t) > 0$ can be shown by showing that the last term of the inequality above is lager than zero. Indeed,

$$W_x'(v_L(t))v_L'(t) - \frac{W_x'(v_H(t))W_y'(v_L(t))v_L'(t)}{W_y'(v_H(t))} > 0 \tag{1.27}$$

$$\Leftrightarrow\; W_x'(v_L(t))W_y'(v_H(t)) > W_x'(v_H(t))W_y'(v_L(t)) \tag{1.28}$$

$$\Leftrightarrow\; \frac{W_x'(v_L(t))}{W_y'(v_L(t))} > \frac{W_x'(v_H(t))}{W_y'(v_H(t))} \tag{1.29}$$

The last part directly follows by $\frac{W_x'(v)}{W_y'(v)}$ being strictly monotonically decreasing in v. This proves the proposition.

Proof of Proposition 1.7

I show that the proposition holds if the social planner is a Rawlsian planner that aims to maximize the utility of the low income types. Since for any welfare function π_L^R is a lower bound for the L-types' optimal consumption probability π_L^* and Q is an upper bound, considering a Rawlsian planner is sufficient.

For a Rawlsian planner, it holds that $\mathcal{W}_R'(t) = v_L'(t) = \theta\pi_L'(t) + h'(e_L + t)$. $\pi_L(t)$ is implied by the incentive constraint of the H-type according to equations (1.11), (1.13), or (1.14) that are discussed in the preceding Proposition 1.4. Which equation is to be used depends on Q and t. In any case it can be easily seen that $\theta\pi_L'(t)$ is independent of θ which implies that $\mathcal{W}_R'(t)$ is independent of θ. Therefore, varying θ does not impact on $\mathcal{W}_R'(t)$ but only impacts on the domain on which to find the maximizer of $\mathcal{W}_R'(t)$ since t_M depends on θ. The proof now proceeds as follows. I first show that $t_M(\theta) \to \infty$ for $\theta \to \infty$ and that, for $Q > G_H$, $t_1(\theta) \to \infty$ where t_1 is defined by equation (1.12) as the kink in $\mathcal{W}(t)$. Second, I show that $\mathcal{W}_R'(t) < 0$ for some $t \in \mathbb{R}_+$. This implies that there exists some θ such that $t_R^* < t_M(\theta)$ is the maximizer of $\mathcal{W}_R(t)$. Since $\mathcal{W}_R'(t)$ is independent of θ, t_R^* is then the optimal transfer for all $\theta' > \theta$. Since $t_1(\theta)$ also converges to infinity for θ becoming arbitrarily large, it holds that $\pi_L^*(t^*) = Q - \frac{1}{2\theta}(h(e_H + t_R^*) - h(e_H - t_R^*))$ whenever θ is large enough. Since t_R^* is constant for $\theta' > \theta$, it therefore holds that $\pi_L^*(t^*) \to Q$ for $\theta \to \infty$.

1. It is to show that $t_M(\theta) \to \infty$ for $\theta \to \infty$ and $\mathcal{W}'(t) < 0$ for some $t \in \mathbb{R}_+$. t_M is uniquely defined by

$$\min\{2Q, 1\}\theta = h(e_H + t_M) - h(e_H - t_M).$$

Since the right hand side is strictly increasing in t, $t_M(\theta)$ strictly increases in θ. $t_M(\theta)$ also becomes arbitrarily large, since for any t_M there exists some θ such that the equation above is satisfied. For $Q > G_H$, t_1 is uniquely defined by

$$(1 - Q)2\theta = h(e_H + t_1) - h(e_H - t_1).$$

Therefore, the same argument as used for t_M shows that $t_1(\theta) \to \infty$ for $\theta \to \infty$.

2. To show that $\mathcal{W}_R'(t) < 0$ for some $t \in \mathbb{R}_+$, explicitly consider the first derivative for the case that either $Q \leq G_H$ or $t \in [0, t_1]$ and show that is negative for some t (it is sufficient to consider this case since t_1 becomes arbitrarily large). It holds that

$$\mathcal{W}_R'(t) = -\frac{1}{2}[h'(e_H + t) + h'(e_H - t)] + h'(e_L + t).$$

Since $h'(\cdot)$ is strictly decreasing and bounded below by zero, there exist $K > k > 0$ such that for t large enough $h'(e_L + t) - \frac{1}{2}h'(e_H + t) < \frac{1}{2}k$ and $h'(e_H - t) > K$. This implies that $\mathcal{W}_R'(t) < \frac{1}{2}k - \frac{1}{2}K < 0$ for t being large enough. This completes the proof.

CHAPTER 2

CONSTRAINTS ON MATCHING MARKETS BASED ON MORAL CONCERNS

Abstract

Monetary transfers are banned or heavily restricted in many markets. These restrictions are often motivated by moral concerns. However, it is not obvious whether the observed restrictions on monetary transfers are the appropriate market design answer to these concerns. Instead of exogenously restricting monetary transfers on a market for indivisible objects, we introduce a desideratum based on egalitarian objectives and study its market design implications. The desideratum we consider is *discrimination-freeness*, which requires that one's access to certain resources is independent of one's wealth endowment. A key assumption in our model is that wealth impacts on the agents' willingness to pay. We show that if discrimination-freeness is desired monetary transfers cannot be used to Pareto-improve ordinal assignment mechanisms that do not involve monetary transfers. Moreover, we find that implementable social choice functions are discrimination-free if and only if an agent's object assignment depends on his ordinal object ranking only and his money assignment is independent of his preferences. In situations where money can be used outside a market designer's control, we show that externality-freeness is needed: an agent's object assignment has to be independent of other agents' preferences. We discuss applications of our results in the context of discrimination-freeness including compensation for kidney donors.

2.1 Introduction

Why worry that we are moving toward a society in which everything is up for sale? ... One [reason] is about inequality ... Where all good things are bought and sold, having money makes all the difference in the world.

Michael Sandel in "What Money Can't Buy"[1]

Various markets ban monetary transfers or heavily regulate them by law. Selling

[1]Compare Sandel (2012), p.8

organs or financially compensating organ donors is prohibited almost everywhere in the world. School and university places are free of charge and must not be traded for money in many countries. A classical utilitarian welfare perspective cannot explain the prohibition of transactions when all involved parties would give their consent. However, anxiety and repugnance towards transactions involving transfers clearly exist in several markets (Frey and Pommerehne, 1993, Kahneman et al., 1986, Roth, 2007). As Satz (2010) puts it, *"From the egalitarian's angle of vision, what underlies noxious markets ... is a prior and unjust distribution of resources, ... the fairness of the underlying distribution of wealth and income is extremely relevant to our assessment of markets."*[2] Inequality concerns are also considered as one of the main sources for market disapproval by Sandel (2012). Intense public debates demonstrate the ambivalent character of using money for allocating certain types of resources. Price mechanisms allow to promote the efficiency of an allocation, but since somebody's willingness or ability to pay might depend on wealth, it also implies that who gets what depends on wealth.

In this paper, we study market design implications if wealth-independent access to goods is a desideratum.[3] We develop a formal model for the assignment of objects and money to agents who are characterized by preferences that are not linear in money and a wealth level. The assignment is required to be *discrimination-free* in the sense that the object an agent is assigned to does not depend on his wealth endowment. Consider any social choice function that assigns objects only based on information about rankings and that is at the Pareto-frontier of social choice functions that do not use transfers. We find that discrimination-free social choice functions *with* monetary transfers cannot realize Pareto-improvements compared to this one. In a private information setting, we find that requiring a social choice function to be discrimination-free already implies that only an agent's object ranking can be used for his object assignment and that his money assignment needs to be independent of his preferences. The only way to incorporate any information beyond ordinal preferences is to ensure that an agent's ex-post wealth is independent of his wealth endowment. This, however, requires that the market designer can condition the mechanism on the agents' wealth endowments. Assigning objects without using transfers can, therefore, be understood as a tool to satisfy a desire for wealth-independent access to certain goods whenever the mechanism does not (or cannot) eliminate potential wealth differences in endow-

[2] Compare Satz (2010), p.5

[3] Inequality is clearly not the only argument used by opponents of transfers in certain markets. However, other arguments are not in our focus here. We furthermore do not aim to answer the question on which markets inequality is desired.

ments. We show that even stricter restrictions than banning transfers are needed, if money can be used to improve access to a good outside the market designer's control.

Intense discourses about the role of money in various markets reveal the importance of studying motivations behind the desire to ban transfers and their implications for market design. In the US, there is an ongoing debate about compensations for kidney donors. Proposals range from free markets to regulated markets to strictly prohibiting any transfers. In Germany, a back and forth in charging tuition fees at universities was accompanied by intense debates.[4] By introducing discrimination-freeness as a constraint on market design, we formally capture a desire that underlies the wide-spread reluctance towards transfers in certain markets. The content of the Universal Declaration of Human Rights supports our conjecture that discrimination-free access is a deeper desire than restricting transfers: it incorporates both a *right to education* as well as a *right to health* and highlights the importance of discrimination-free access to both.[5] Furthermore, empirical findings suggest that whether a third party considers it unethical to receive monetary incentives in return for participating in a transaction, depends on whether, from his financial perspective, he would accept the incentives and take part in the transaction (Ambuehl, Niederle, and Roth, 2015). To the best of our knowledge, our's is the first paper on the provision of indivisible resources that explicitly models a wealth independent access to goods as a fairness criterion.

Our Analysis. We consider the problem of assigning indivisible objects to agents. Each agent is characterized by a *type* containing information about his initial wealth endowment and a utility function that describes how he evaluates bundles of objects and wealth. A *social choice function* assigns one object to each agent and determines monetary transfers.[6] It is called *discrimination-free*, if the object assignment of an agent does not depend on his wealth endowment. A key assumption we impose on the agents' utility functions is that they are not linear in money. While an agent's ranking of objects is assumed to be wealth-independent, his marginal utility of money and his willingness to pay for preferred objects de-

[4] After a period of having (basically) not charged any fees, from 2006 on universities were allowed to charge up to 1000 EUR per year. Protests were huge and finally, in 2014, there is no university left charging fees (see, e.g., The Conversation, 2014).

[5] See, e.g., articles 25 and 26 the Universal Declaration of Human Rights (1948). By General Comment No. 14 (2000): "Health facilities, goods and services have to be accessible to everyone without discrimination [...]".

[6] Thereby, we concentrate on deterministic social choice functions and therefore do not incorporate potential ex-ante improvements via allowing for probabilistic outcomes. We discuss this later in more detail.

pend on wealth.[7] In particular, if his wealth increases, the amount that is required to compensate him for a less preferred object increases as well. A high willingness to pay for an object can thus be due both to a high utility benefit associated with the object and to high wealth. The assumption of non-linear preferences is in contrast to many standard mechanism design models and is crucial for our analysis. This is because two agents wish to trade an object if and only if the price the owner is willing to accept for giving up the object is lower than the price the potential buyer is willing to pay for the object. For preferences that are linear in money, the desire to trade does not depend on wealth and therefore no discrimination concerns occur.

In a world without money and private information about preferences, a market designer is restricted to mechanisms such that an agent's object assignment is not sensitive to cardinal information about his preferences.[8] In our analysis, we are interested in the implications on the design of social choice functions that are allowed to use transfers but that are required to be discrimination-free. In particular, we explore what information about preferences can be exploited for the assignment of objects and money. Can money be used to Pareto-improve money-free mechanisms based on ordinal information by trading-off differences in preference intensities? Can payments be used to elicit private information about preference intensities? What are necessary and sufficient conditions on social choice functions to meet discrimination-freeness?

First, we find that social choice functions with wealth-independent transfers cannot be Pareto-improved on by using transfers without violating discrimination-freeness. Therefore, by allocating the objects without using transfers and exploiting only information about the agents' rank order lists, we can already reach the Pareto-frontier of discrimination-free social choice functions.[9] The main driver of the result is that on the one hand, the amount of money compensating an agent for a worse object becomes larger if the agent gains wealth, and on the other hand, the

[7]Therefore, for which objects agents compete, is independent of their wealth. Otherwise, moral concerns occurring might rather belong to segregation concerns that are not further considered here. We briefly discuss dropping the assumption of non-constant rankings as an extension.

[8]To see this, assume that there are two preference profiles of an agent that both represent the same ordinal ranking but that serve him different objects. Then, whenever the agent has preferences according to the profile that serves him the less preferred object, he has an incentive to misreport.

[9]This is relevant independent of the information setting. In particular, allocating objects via a *Serial Dictatorship* mechanism where one agent after each other selects an object is implementable in a private information setting and at the Pareto-frontier of discrimination-free social choice functions. Our results then imply that even if we had full information about preferences, the mechanism cannot be Pareto-improved in a discrimination-free way.

money all other agents are willing to pay for an object improvement is bounded. However, any discrimination-free social choice function with wealth-independent transfers is not Pareto-efficient within the set of all social choice functions. We show that discrimination-freeness and Pareto-efficiency are not exclusive, but to satisfy both, wealth-dependent transfers are needed. To condition transfers on wealth, however, it is relevant which information the market designer has about the agents' wealth.

Second, we consider a setting where types are agents' private information and implementability of the social choice function is required.[10] We show that a social choice function is discrimination-free if and only if an agent's money assignment is independent of his type *and* his object assignment depends on his object ranking only. Money, therefore, can not be used to elicit and exploit information beyond rankings. Inefficiencies in markets without transfers are obtained as second best outcomes. Again, wealth effects are crucial: if transfers depend on the types, we show that there exist preferences such that for high wealth levels, the agent benefits by focusing on the preferred object, while for low wealth levels, he benefits by focusing on the monetary difference. This induces incentives to misreport, since a discrimination-free mechanism must not condition the object assignment on wealth. The toolkit to allocate objects if transfers are banned, therefore, corresponds to the one available if discrimination-freeness is desired. A simple mechanism that is implementable, discrimination-free, and at the Pareto-frontier of discrimination-free social choice function is the *Serial Dictatorship* mechanism, where one agent after the other selects an object.[11] We find that if wealth levels are public information, only a social choice function that fully eliminates an agent's potential wealth differences can exploit information beyond his object ranking for his object and money assignment. Otherwise, to meet discrimination-freeness and implementability, only information about an agent's object ranking can be used for his object assignment and his money assignment must be independent of his preferences (but might depend on wealth). Examples for preference-independent but wealth-dependent transfers are goods that are financed via taxes and the consumption of which does not require any additional fee.

[10]We call a social choice function implementable if reporting the truth type is a dominant strategy.

[11]Serial Dictatorship is not Pareto-efficient within the set of all social choice function when transfers are allowed. However, this does not contradict the classical *Gibbard-Satterthwaite Theorem* (Gibbard, 1973, Satterthwaite, 1975) which implies that a strategy-proof social choice function that reaches all outcomes is Pareto-efficient. We restrict our attention to a discrimination-free social choice function and, therefore, not all outcomes can be reached.

Finally, we discuss implications of discrimination-freeness in situations where money can be used to improve access to goods outside a mechanism. Technically, we extend our model by taking into account that somebody bribes somebody else to misreport preferences in line with Schummer (2000b). Externality-freeness (i.e., an agent's outcome must not depend on other agents' preferences) is then sufficient to ensure the preservation of discrimination-freeness under bribes. For nonbossy social choice functions (i.e., an agent cannot change another agent's outcome without changing his own), externality-freeness is also necessary to ensure the preservation of discrimination-freeness under bribes. This is, because bribing incentives appear as soon as other agents' preferences play a role for an agent's outcome.[12] However, if the wealth endowment of an agent who is bribed increases, the incentive to bribe him eventually disappears such that the object he receives depends on wealth. Externality-freeness is a severe restriction for the design of social choice functions. If the number of objects equals the number of agents, externality-freeness implies that the allocation of objects must not depend on anyone's preferences. If more objects than agents are available, externality-freeness implies wastefulness (i.e., an agent may prefer an unassigned object over the object he is assigned to). The analysis of bribes can be interpreted more generally as using money outside a centralized mechanism to influence one's access to a good. Applications include co-existing private markets or priority parameters like living in a school's neighborhood where paying more for a house can help to improve one's priority at the school.

Overall, our results explain the wide-spread use of a matching market approach to assign objects, whenever discrimination-freeness is desired and differences in wealth are not fully eliminated. The analysis is relevant for several real-world applications. In particular, for the question whether or not two persons should be allowed to trade a good like a kidney, discrimination-freeness requires that the transaction takes place independent of the wealth of anyone involved.

Related Work. Our work relates to the literature on repugnance on markets, in particular on the desire of third parties to restrict transfers (e.g., Ambuehl et al., 2015, Frey and Pommerehne, 1993, Kahneman et al., 1986, Roth, 2007). In contrast to that literature, we explicitly integrate a concern underlying the desire to ban transfers into an economic model. Our definition of discrimination-freeness appears to be in line with what people judge as immoral according to Ambuehl

[12]This is closely related to Schummer (2000a) and Schummer (2000b). His results imply that bribe-proofness is equivalent to externality-freeness for a very general class of quasilinear preferences.

et al. (2015).[13] While we concentrate on the concern of inequality, Ambuehl (2015) studies the concern of coercion in the context of financial incentives. In contrast to our work, he studies how incentives affect those whom they target.[14]

Our research complements the literature on the implications of fairness concerns on allocating resources. Thomson (2011) provides a comprehensive overview on fair allocation rules. Popular fairness criteria typically refer to how an agent evaluates his bundle in comparison to another agent's bundle. For instance, *no envy* requires that no agent prefers any other agent's bundle, *equal treatment of equals* requires that no agent prefers any other agent's bundle whenever the other agent has the same preferences over bundles. In contrast, discrimination-freeness is grounded in the analysis of a single individual and refers to the object an agent is assigned to if his wealth level changes.

Key for our analysis is the non-linearity of preferences. Here, our model differs from the standard assumption in many economic models where consumers have quasilinear preferences. There are some works that deal with the impact of non-linearities in preferences such as budget constraints, risk aversion or wealth effects, to the provision of indivisible goods (see, e.g., Baisa, 2013, Che et al., 2013, Garratt and Pycia, 2014, Maskin and Riley, 1984).

Outlook. The remainder of the paper is organized as follows: in Section 2.2 we describe the basic model. In Section 2.3 we introduce discrimination-free social choice functions. We also discuss implications of discrimination-freeness on efficiency and characterize discrimination-free social choice functions. In Section 2.4 we consider consequences for discrimination-free social choice functions if money is used outside the mechanism designer's control. Then we discuss several extensions (Section 2.5) and applications (Section 2.6). We conclude with Section 2.7.

[13] In that work, they present a basic model based on survey results assuming that people judge a transaction as immoral if, from their financial perspective, they would not take part in the transaction. In their context, our definition of discrimination-freeness then generally speaking translates to requiring moral approval from anyone's financial perspective.

[14] There is a large literature dealing with how incentives impact on the moral behavior of individuals (Frey and Oberholzer-Gee, 1997, Gneezy and Rustichini, 2000, Mellström and Johannesson, 2008, Richard, 1970). In contrast, we are interested in how monetary incentives impact on who receives what.

2.2 Model

We consider the problem of assigning a set Ω of $k \geq n$ distinct and indivisible objects to a set N of $n \geq 2$ agents. Each agent receives exactly one object.[15]

Payoff Environment. Preferences of each agent i are described by a utility function $u_i : \Omega \times \mathbb{R} \to \mathbb{R}$. $u_i(\omega, A)$ denotes the utility that agent i derives from owning object $\omega \in \Omega$ and having a total wealth of $A \in \mathbb{R}$. We assume that the agents' preferences are twice differentiable in wealth. Furthermore, we make the following assumptions on how wealth affects preferences.

1. *Strict and wealth independent object ranking:* u_i implies a strict and unique rank order of objects denoted by r_i. Formally,

$$\begin{aligned} u_i(\omega, A) \neq u_i(\omega', A) \quad &\Leftrightarrow \quad \omega \neq \omega' \quad \text{and,} \\ u_i(\omega, A) > u_i(\omega', A) \quad &\Rightarrow \quad u_i(\omega, A') > u_i(\omega', A') \ \forall \ A' \in \mathbb{R}. \end{aligned}$$

2. *Monotonicity and strict concavity in wealth:*

$$\frac{\partial}{\partial A} u_i(\omega, A) > 0 \quad \text{and} \quad \frac{\partial^2}{\partial A^2} u_i(\omega, A) < 0 \ \forall \omega \in \Omega$$

3. *Unbounded willingness to accept:* Let u_i be such that object ω is preferred over ω'. For any $m > 0$ there exists $\overline{A}_i \in \mathbb{R}$ such that

$$u_i(\omega, A) > u_i(\omega', A + m) \ \forall A > \overline{A}_i$$

$\mathcal{U}$ denotes the set of all utility functions that satisfy the above assumptions. According to the first assumption, wealth does not influence how an agent ranks the objects.[16] The second assumption ensures that each agent has a finite willingness to pay for any object improvement: for any $u_i \in \mathcal{U}$, $A \in \mathbb{R}$, $a, b \in \Omega$ where a is preferred to b there exists a unique $M > 0$ such that $u_i(a, A - M) = u_i(b, A)$.[17] The third assumption stated in word means the following: suppose an agent prefers object ω to object ω' and he is offered any amount $m > 0$ for taking ω' instead of ω. Then, whenever his wealth is large enough, he refuses this offer and rather

[15] There is only one copy of each object; however, it is straight forward to include objects with more copies in this setting and the analysis.

[16] Technically, the assumption of wealth independent object rankings already follows by the assumptions of continuity in wealth and strict ranking. However, due to the importance of unique order rankings, we explicitly state this as an assumption.

[17] For a formal proof see the proof of Proposition 2.1.

takes ω. Therefore, the compensation to accept for an object impairment becomes arbitrarily large for increasing wealth.

Examples for utility functions in $\mathcal{U}$ are those that can be described by $u_i(\omega, A) = v_i(\omega) + h_i(A)$ where $v_i : \Omega \to \mathbb{R}$ and $h_i : \mathbb{R} \to \mathbb{R}$ is a twice continuously differentiable function with $h_i' > 0$, $\lim_{A\to\infty} h_i'(A) = 0$, and $h_i'' < 0$. All results we develop continue to hold if the set of admissible utility functions is restricted to utility functions that are of this shape. Note that we allow agents to have negative wealth and no budget constraints exist. In Section 2.5 we discuss why this assumption is not critical for our analysis, and explain how many of our results even do not rely on the assumption that the possible wealth endowments of agents are not bounded from above.

Each agent is endowed with an initial wealth level $e_i \in \mathbb{R}$. $t_i = (u_i, e_i) \in T = \mathcal{U} \times \mathbb{R}$ denotes the *type* of each agent that implies how agent i evaluates bundles of objects and wealth. T^n is the space of all type profiles $t = (t_i)_{i\in N}$ and $t_{-i} \in T^{n-1}$ is the type profile of all agents except agent i.

In our analysis we are interested in shared characteristics of different types. $T(r_i)$ denotes the set of all types that describe the same ordinal ranking $r_i \in R$ of objects where R denotes the set of all possible rankings over objects. $T(u_i)$ denotes the set of all types that describe the same utility function u_i. $T(e_i)$ denotes the set of all types with equal wealth endowment e_i. While all types in $T(r_i)$ agree on the ranking of objects, they might disagree on what any object improvement is worth. This heterogeneity can have two sources. First, even if t_i and t_i' describe the same endowment and the same object ranking, the cardinal appreciation for the objects might differ according to different utility functions u_i and u_i'. Second, even if t_i and t_i' describe the same utility function $u_i = u_i'$, the willingness to pay for object improvements might differ due to endowment differences. Furthermore, if two types t_i and t_i' both belong to $T(u_i)$ they also both belong to $T(r_i)$ for some object ranking r_i. Therefore, $T(u_i) \subset T(r_i)$. If two types t_i and t_i' in $T(u_i)$ disagree on what an object improvement is worth this can only be due to heterogeneity in wealth levels.

Social Choice Functions. An *outcome* $x = (\sigma, m) \in \Omega^n \times \mathbb{R}^n$ assigns exactly one object to each agent expressed by $\sigma \in \Omega^n$ and defines monetary transfers by $m \in \mathbb{R}^n$.[18] $\sigma_i = \omega$ means that object ω is assigned to agent i. $m_i \in \mathbb{R}$ is the money agent i receives.

[18] In particular, no agent remains unassigned. The setting can be easily extended by adding an object $\emptyset$ with n copies to Ω where $\emptyset$ corresponds to remaining unassigned.

Each type $t_i = (u_i, e_i)$ uniquely defines preferences over outcomes. In particular, agent i of type $t_i = (u_i, e_i)$ evaluates his individual outcome (σ_i, m_i) according to $u_i(\sigma_i, A_i)$ where $A_i = e_i + m_i$ is agent i's *ex-post wealth*. In contrast to quasilinear preferences, knowing u_i is not sufficient to evaluate outcomes but we also need to know an agent's wealth endowment because two agents with the same utility function u_i might evaluate outcomes differently due to differences in wealth. On the other hand, two agents might evaluate outcomes in the same way but their types differ.

$\varphi = (\sigma, m)$ denotes a *social choice function* (or direct *mechanism*, if types are private information) that selects for each type profile $t \in T^n$ an outcome $\varphi(t) = (\sigma(t), m(t))$. $\varphi_i = (\sigma_i, m_i)$ is agent i's assignment. We call $\sigma : T^n \to \Omega^n$ the *object assignment* and $m : T^n \to \mathbb{R}^n$ the *money assignment*.[19] φ might use tie-breaking rules like priorities (e.g., based on districts in school choice) or lotteries. We assume that those tie-breakers are determined *before* the mechanism is conducted and are fixed for each agent independent of the realization of types. We concentrate on deterministic outcomes instead of lotteries over deterministic outcomes. This corresponds to taking an ex-post perspective. Therefore, we do not restrict our attention to anonymous mechanisms since agents might differ according to priorities or a lottery number. This perspective is more suited to our analysis because we are interested in whether money can be used to increase efficiency and not on whether ex-ante efficiency gains can be achieved via lotteries.

Definitions. A social choice function $\varphi' = (\sigma', m')$ (or an object assignment σ') *Pareto-dominates* $\varphi = (\sigma, m)$ (or σ) if for all type profiles $t \in T^n$ all agents are weakly better off and at least for one $t \in T^n$ there is one agent who is strictly better off. σ is a *Pareto-efficient object assignment* if there is no object assignment σ' that Pareto-dominates σ. $\varphi = (\sigma, m)$ is a *Pareto-efficient social choice function* if there is no social choice function $\varphi' = (\sigma', m')$ with the same budget $\sum m'_i = \sum m_i$ that Pareto-dominates φ. Thereby, we allow social choice functions not to be budget-balanced. For instance, money might be extracted to fund the provision of resources. For the definition of Pareto-efficiency, we restrict our attention to Pareto-improvement without extending the budget. If extending the budget was allowed, a social choice function could be Pareto-improved on just by increasing each agent's wealth.

A social choice function $\varphi = (\sigma, m)$ is *implementable* if it can be implemented as a dominant strategy equilibrium of a direct mechanism. By the revelation

[19] With a slight abuse of notation we denote by σ the assignment that maps profiles to a an object allocation as well as the allocation itself; the same holds for m.

principle, for implementability we limit our attention to social choice functions where truthtelling is a dominant strategy. Truthtelling is a dominant strategy if and only if $u_i(\sigma_i(t_i, t_{-i}), e_i + m_i(t_i, t_{-i})) \geq u_i(\sigma_i(t'_i, t_{-i}), e_i + m_i(t'_i, t_{-i}))$ for each agent i and all $t_i, t'_i \in T$ and $t_{-i} \in T^{n-1}$. A social choice function $\varphi = (\sigma, m)$ is *ordinal* if it is not sensitive to cardinal information. Formally, $\varphi(t) = \varphi(t')$ if for all i it holds that $t_i, t'_i \in T(r_i)$ for some rank order $r_i \in R$. An ordinal object assignment is defined analogously. In line with Satterthwaite and Sonnenschein (1981) we call a social choice function φ *nonbossy* if for any agent i, $\varphi_i(t_i, t_{-i}) = \varphi_i(t'_i, t_{-i})$ implies $\varphi(t_i, t_{-i}) = \varphi(t'_i, t_{-i})$. Therefore, an agent cannot change another agent's outcome without changing his own.[20]

2.3 Discrimination-Free Social Choice Functions

In our model we deliberately omit the typical restriction of a matching market that monetary transfers are not allowed. Instead, we introduce a desideratum that is used in many discourses as an argument for restricting transfers: discrimination-freeness with respect to wealth. We call a social choice function discrimination-free if an agent's object assignment does not depend on his wealth endowment. Hence, discrimination-freeness refers to what determines how objects are allocated but does not a priori impose restrictions on transfers.

Definition 2.1 (Discrimination-Free). *A social choice function $\varphi = (\sigma, m)$ is discrimination-free (with respect to wealth) if for any agent i, utility function $u_i \in \mathcal{U}$, and type profile $t_{-i} \in T^{n-1}$ from the other agents*

$$\sigma_i(t_i, t_{-i}) = \sigma_i(t'_i, t_{-i}) \text{ for all } t_i, t'_i \in T(u_i).$$

φ discriminates if it is not discrimination-free.

An appealing feature of our definition of discrimination-freeness is that to judge whether or not this fairness criteria is satisfied it is sufficient to consider one agent. This allows us to concentrate on deterministic outcomes.[21]

[20] Whether or not nonbossiness is a desirable characteristic of a social choice function appears to be disputable. Thomson (2014), for instance, discusses several interpretations of nonbossiness and questions their validity. This paper remains agnostic to whether or not nonbossiness should be required. In the context of Proposition 2.4 we rather discuss implications of imposing it.

[21] In contrast, classical fairness criteria like *envy-freeness* or *equal treatment of equals* make restrictions on how an agent evaluates another agent's outcome.

For quasilinear utilities, preferences over outcomes do not depend on wealth and therefore discrimination-freeness does not impose restrictions on how a social choice function depends on preferences. However, since we impose income effects, discrimination becomes a valid concern. For illustration consider two agents and two objects and assume that both agents prefer object a to object b. One agent is willing to pay more to receive object a instead of object b than the other one. The willingness to pay is driven by preferences over bundles of objects and wealth as well as by endowments. A discrimination-free social choice function must not take account of the wealth effect but might regard utility effects. A central question in our following analysis is to what extent discrimination-free social choice function can use information about preferences to assign the objects.

2.3.1 Pareto-Efficiency

Free markets allow a transfer of utility via money and therefore offer the opportunity to realize Pareto-improvements via trades of objects and money. Markets without transfers provide less opportunities for Pareto-improvements since there is no divisible good available to transfer utility. In any environment without transfers, mechanisms that assign objects to agents by only exploiting information about the agents' object rankings are at the Pareto-frontier of all mechanisms that do not use transfers. However, agents that agree on the object ranking might disagree on what an object improvement is worth. This is why ordinal mechanisms without transfers are, in general, not Pareto-efficient. The central question for the following analysis is whether money can be used to realize Pareto-improvements compared to a classical money-free matching without violating discrimination-freeness.

For a simple example, consider two agents i and j and two objects a and b. Both agents prefer object a over object b. Then, assigning a to i and b to j, for instance, is not Pareto-dominated by swapping the objects. However, both agents might be better off when exchanging objects in return for a money transfer. This is the case whenever agent j's willingness to pay for a is higher then agent i's willingness to accept for giving up a. When increasing agent i's wealth level, agent i might not any more be willing to give up the preferred object in turn for a transfer that agent j is willing to make. This is why we cannot Pareto-improve an object assignment by admitting transfers, such that the object allocation is independent of wealth. The following proposition formalizes and generalizes this insight. It is in contrast to a setting with quasilinear preferences where what somebody is willing to pay or to accept in turn for an object exchange does not depend on wealth.

Proposition 2.1. *Consider a discrimination-free social choice function $\varphi = (\sigma, m)$ such that its money assignment m does not depend on wealth endowments $(e_i)_{i\in N}$. Suppose that σ is a Pareto-efficient object assignment. Then, φ is not a Pareto-efficient social choice function. Any social choice function $\varphi' = (\sigma', m')$ with $\sum_i m'_i = \sum_i m_i$ that Pareto-dominates φ does discriminate.*

Proof. See Appendix. □

The main driver of the proposition is in analogy to the simple example above that the compensation an agent is willing to accept becomes arbitrarily large for increasing wealth. Then, the amount of money compensating somebody for receiving a less preferred object under φ' than under φ becomes arbitrarily large when wealth of this agent increases (since transfers do not depend on wealth). On the other hand, there is a maximal amount of money that each agent is willing to pay for improving the object assignment of φ. This maximal willingness to pay can be determined independently of the other agent's wealth (again since transfers do not depend on wealth). Then, any agent that receives a less preferred object under φ' compared to φ cannot be compensated any more for this object impairment when being wealthy enough. Therefore, if any φ' Pareto-improves φ, there is at least one agent for whom the object assignment of φ' depends on his wealth level. This implies that φ' is not discrimination-free. To get an intuition on why φ is not Pareto-efficient consider a type profile where all agents agree on the ranking. By repeatedly changing the wealth levels of two agents we can induce a situation where a trade of objects in return for monetary transfers is a Pareto-improvement.[22]

Note that the proposition only considers potential Pareto-improvements of social choice functions that have the same budget. When extending the budget is allowed, Pareto-improvements without discrimination are straight-forward by just increasing each agent's wealth level. Therefore, the case where the budget is constrained is the interesting one when searching for Pareto-improvements.

Proposition 2.1 implies that by using object assignments that do not use trans-

[22]Proposition 2.1 requires that transfers of φ are independent of wealth. This implies that some delta in the endowment of an agent implies the same delta in the ex-post wealth level of this agent. It would be sufficient to require that φ preserves the wealth status of an agent in the sense that an agent's ex-post wealth is unbounded in dependence of his own wealth but bounded in dependence of other agents' endowments. However, we state it in the simplified way to not distract from the main point.

fers and that only exploit information on object rankings, the Pareto-frontier of discrimination-free social choice functions can be reached.

Corollary 2.1. *Consider any ordinal object assignment σ. Suppose σ is not Pareto-dominated by any other ordinal object assignment σ'. Then $\varphi(t) = (\sigma(t), 0)$ is at the Pareto-frontier of all budget-balanced and discrimination-free social choice functions.*

A second and direct implication of Proposition 2.1 is that if any social choice function is Pareto-efficient and discrimination-free the transfers of the social choice function necessarily depend on wealth. With the following corollary we furthermore show that efficiency and discrimination-freeness are not exclusive.

Corollary 2.2. *There is a discrimination-free social choice function $\varphi = (\sigma, m)$ that is Pareto-efficient. Any social choice function that is discrimination-free and Pareto-efficient assigns transfers that depend on wealth.*

As an example for a discrimination-free and Pareto-efficient social choice function consider the following one that performs the assignment in two steps. First, wealth of each agent i is adjusted according to some wealth level which is independent of his initial endowment e_i. Second, given this new wealth distribution, the mechanism assigns objects such that the sum of utilities is maximized. This allocation is Pareto-efficient. Furthermore, an agent's object assignment is independent of his wealth endowment. The social choice function described is not necessarily budget balanced. However, if wealth endowments are drawn from a distribution such that expected total endowment is $\overline{e}$, the mechanism above is budget balanced in expectation if each agent's wealth is adjusted to $\frac{1}{N}\overline{e}$.[23]

On the Information Structure. For Proposition 2.1 we did not impose any specific information structure about the types of the agents. Social choice functions that are discrimination-free with Pareto-efficient object assignments and wealth-independent transfers can be implemented in a setting where the mechanism designer has no information about types. The *Serial Dictatorship* mechanism where one agent after the other selects an object and no transfers are made is an example for such a social choice function that is implementable in dominant strategies. Therefore, the Pareto-frontier of discrimination-free social choice

[23]It is furthermore also feasible to construct a mechanism that is Pareto-efficient and ex-post budget balanced. Such a mechanism can be constructed by allocating objects such that for some specific wealth endowment no Pareto-improvements are feasible via transfers. Then, for any other wealth endowment it is possible to redistribute wealth such that this allocation of objects cannot be Pareto-improved within the budget.

functions can be reached in a setting of incomplete information. Proposition 2.1 can then be interpreted in the sense that *even if* we have full information about types, no Pareto-improvement can be realized. Furthermore, any admission of ex-post trades leads to discrimination. The Pareto-efficient mechanism presented in the context of Corollary 2.2, however, depends on the information structure and requires that the mechanism designer is informed about the agents' types. In the following section, we deal with the implications of incomplete information on the set of implementable social choice function.

2.3.2 Implementability

The previous section focused on the question whether money can be used to achieve Pareto-improvements compared to a classical matching market without transfers in a discrimination-free way. In the following we consider a setting of incomplete information and are interested in whether money can be used to exploit more information than ordinal rankings to assign the objects to the agents. Furthermore, we analyze the restrictions that arise for the money assignment. This is of interest since payments might be used for funding resources or to redistribute wealth.

We first assume that no information about the type is available (i.e., both the utilities profile $(u_i)_{i\in N}$ and the wealth profile $(e_i)_{i\in N}$ are unknown). Later we assume that the wealth profile $(e_i)_{i\in N}$ is known.

Proposition 2.2. *Let $\varphi = (\sigma, m)$ be an implementable social choice function. φ is discrimination-free if and only if for each agent i and $t_{-i} \in T^{n-1}$ fixed,*

- *$\sigma_i(t_i, t_{-i}) = \sigma_i(t'_i, t_{-i})$ for all $t_i, t'_i \in T(r_i)$, and all rankings $r_i \in R$, and*
- *$m_i(t_i, t_{-i}) = m_i(t'_i, t_{-i})$ for all $t_i, t'_i \in T$.*

Proof. See appendix.[24] □

The proposition formalizes that if implementability and discrimination-freeness are required, an agent's object assignment cannot be sensitive to cardinal information about his preferences and each agent's money assignment must not depend on his type. To get an intuition for the proof first note that discrimination-freeness

[24] The proof presented is more complex than needed for the domain $\mathcal{U}$ of utility functions. However, it reveals that Proposition 2.2 even holds if the domain of utility function is modified such that every i's utility function can be described by $u_i(\omega, A) = v_i(\omega) + h(A)$ for some $h : \mathbb{R} \to \mathbb{R}$ with $h' > 0$, $\lim_{A\to\infty} h'(A) = 0$, and $h'' < 0$. It furthermore also allows for a restriction of admissible endowments to some $E \subset \mathbb{R}$ such that E contains at least two elements. For this, see also Section 2.5.

and implementability of a social choice function imply neither the object assignment nor the money assignment can be conditioned on endowments. Assume that monetary payments are not type independent such that for two types t_i and t_i' it holds that $m_i(t_i) < m_i(t_i')$. Implementability of the social choice function implies that there are at most $|\Omega| = k$ outcomes that are available to agent i via varying his report. We can then construct a utility function such that the outcome of t_i is the most preferred one for one wealth level and the outcome of t_i' is the most preferred one for another wealth level. This contradicts discrimination-freeness of φ and therefore agent i's payments cannot depend on his type. The restriction on σ that σ_i must only depend on agent i's ordinal ranking is a direct implication of the restrictions on m: since m_i is independent of agent i's type, considering more information than rank order lists for the object allocation contradicts implementability.

In Proposition 2.1 we saw that object assignments without transfers that are based on ordinal rankings are already at the Pareto-frontier of discrimination-free social choice functions. By Proposition 2.2, the toolset to distribute objects is even restricted to the one that can be used if no transfers are admitted since only ordinal information about an agent's preferences can be exploited for his object assignment. With the *Serial Dictatorship* mechanism where one agent after the other selects an object we can then implement a social choice function at the Pareto-frontier.[25] Inefficiencies of such a mechanism without transfers are obtained as second-best outcomes when requiring discrimination-freeness. In the context of school choice problems, where students are often ordered according to a priority structure, two popular ordinal and implementable matchings are the *Deferred-Acceptance-Algorithm* proposed by Gale and Shapley (1962) or the *Top Trading Cycles Mechanisms* (see, e.g., Abdulkadiroğlu and Sönmez, 2003). The latter mentioned is at the Pareto-frontier of transfer-free assignments while the former is not.

Availability of Wealth Information. Proposition 2.2 deals with a setting where neither information about utilities nor about wealth endowments is known. If implementability and discrimination-freeness are desired, only information about an agent's object ranking can be exploited by the social choice function for his outcome.

[25]Since we concentrate on deterministic matchings, any lotteries that might be needed for serial dictatorship (or other mechanisms) are assumed to be conducted before the matching takes place.

We now assess a setting where wealth information is available while preferences in the form of utility function over objects and wealth still are unknown. Intuitively, this increases the scope for a mechanism designer to use information about preferences. For instance, she now might, in a first step of the mechanism, adjust the agents' wealth levels such that they do not depend on their initial endowment any more. By the following proposition it turns out, that *only if* ex-post wealth of an agent is independent of his initial wealth, an agent's object assignment can be based on more information than only the agent's object ranking. Otherwise, in accordance to Proposition 2.2, an agent's object assignment is not sensitive to cardinal information about his preferences, and each agent's money assignment is independent of his preferences (but might depend on wealth).

Proposition 2.3. *Let $\varphi = (\sigma, m)$ be an implementable social choice function. Wealth endowments $(e_i)_{i\in N}$ are public information. Assume that for every agent i, $u_i \in \mathcal{U}$ and $t_{-i} \in T^{n-1}$ fixed, agent i's ex-post wealth $A_i = e_i + m_i(u_i, e_i, t_{-i})$ is not constant in his wealth endowment e_i. φ is discrimination-free if and only if for every agent i and t_{-i} fixed*

- $\sigma_i(t_i, t_{-i}) = \sigma_i(t_i', t_{-i})$ *for all $t_i, t_i' \in T(r_i)$ and all rankings $r_i \in R$, and*
- $m_i(t_i, t_{-i}) = m_i(t_i', t_{-i})$ *for all $t_i, t_i' \in T(e_i)$.*

Proof. See Appendix.[26] □

The proof of the proposition uses similar characteristics of preferences as the proof of Proposition 2.2 does. However, it is more complex compared to Proposition 2.2 since varying wealth might vary transfers. Suppose that ex-post wealth of an agent i is not independent of e_i and transfers are not constant. Then there are two wealth levels e_i and e_i' such that the ex-post wealth that is associated with some object a he can reach by varying his report differs for the two wealth levels. We can then construct a utility function such that for one of the wealth levels, the agent prefers object a with the associated transfers, and for the other wealth level he prefers another object b with the associated transfers that he can reach by varying his report. This contradicts discrimination-freeness. The construction of such an utility function works by exploiting that the evaluation of gaining additional money depends on the reference level of wealth .

[26] In analogy to the proof of Proposition 2.2, the proof reveals that the proposition holds as well for a modification of the domain $\mathcal{U}$ of utility functions. Here, instead of $\mathcal{U}$ we can consider the domain of utility functions that can be expressed via $u_i(\omega, A) = v_i(\omega) + h_i(A)$ with $h_i' > 0$, $\lim_{A\to\infty} h_i'(A) = 0$, and $h_i'' < 0$. For this, see also Section 2.5.

If wealth information is known to the market designer, the assignment of objects does not depend on wealth if discrimination-freeness is desired, but transfers might depend on wealth. An example for this is the collection of income-dependent taxes to fund the provision of a good that are independent of the actual consumption of the good.

The only exception that allows a discrimination-free mechanism to exploit information about preferences beyond rankings is if ex-post wealth is made constant with respect to wealth endowments. Proposition 2.3 implies that if φ depends on more information about an agent's preferences than only rank order lists, ex-post wealth of this agent has to be independent of his initial wealth. As an example consider the mechanism presented in the context of Corollary 2.2. First, the mechanism adjusts each agent's wealth level to any predefined wealth level that is independent of his initial wealth (possibly the same for all agents). Then, the mechanism assigns objects and money. In this second step of the mechanisms, utilities $u = (u_i)_{i \in N}$ can play a role. In particular, if we consider the wealth levels that result after the first step and the preferences of the agents as the new types, the mechanism in the second step has all the flexibility that mechanisms have where we do not impose discrimination-freeness.

2.4 Preserving Discrimination-Freeness under Bribes

Even if objects are assigned to consumers without using transfers, there still might be ways how wealth influences the assignment if it is possible to influence one's outcome outside a market designer's control. An example are neighborhood priorities in school choice: Moving houses to an area of a preferred school raises the chances to receive a place at this school. Those being able to afford high house prices have the choice where to live. This in turn influences the access to schools. Black (1999) analyzed housing prices and showed that house prices are correlated with school quality. In the context of organ donations, there are also ways to use money to gain priority. Steve Jobs, for instance, reportedly obtained his liver transplantation because he was advised to raise his chances by subscribing to waiting lists in other states than his home state California.[27] This approach required to be rich enough to be able to quickly move to any location. Co-existing private markets are also examples where being wealthy improves the access to certain markets. Examples for private markets are private schools or private insurances.

[27] See, e.g., CNN (2009).

In what follows we consider a setting similar to the one of Proposition 2.2 where the mechanism designer is not informed about the agents' types. Agents can use money outside the mechanism designer's control by bribing another agent. With bribing we mean that one agent offers money to another agent for reporting false preferences. Bribing therefore provides agents the opportunity to use their money to change parameters of the game. A real-world example of bribes is that somebody bribes somebody else to agree with donating a kidney. Also, the examples at the beginning of this section can be interpreted as a special case of bribing: instead of using money to influence other agents' reports, the money is used to influence other parameters that influence the outcome.

We first define bribing in the spirit of Schummer (2000b).

Definition 2.2 (**Bribing**). *Let $\varphi = (\sigma, m)$ be a social choice function. Agent i has an incentive to bribe agent j if there is a profile $t \in T^n$, a corrupted type $t'_j \neq t_j \in T$, and a bribe amount $\tau \geq 0$ such that*

- $u_i(\sigma_i(t'_j, t_{-j}), e_i + m_i(t'_j, t_{-j}) - \tau) > u_i(\sigma_i(t), e_i + m_i(t))$ *and*
- $u_j(\sigma_j(t'_j, t_{-j}), e_j + m_j(t'_j, t_{-j}) + \tau) > u_j(\sigma_j(t), e_j + m_j(t))$.

φ is bribe-proof if no incentives to bribe exist.

For any agent i and any type profile $t \in T$ define $\sigma_i^B(t) \subset \Omega$ such that $\omega \in \sigma_i^B(t)$ if and only if $\omega = \sigma_i(t)$ or $\omega = \sigma_i(t')$ where $t' = (t'_i, t_{-i}) \in T$ is a corrupted report of types if agent i is bribed.

An agent therefore has an incentive to bribe another agent if paying another agent to state false preferences makes both agents better off. $\sigma_i^B(t)$ contains all object assignment of i that might result if agent i is bribed including the object assignment if no bribes occur. If φ is bribe-proof, then $\sigma_i^B(t)$ contains only $\sigma_i(t)$. We now extend the definition of discrimination-freeness to account for potential bribes.

Definition 2.3 (**Preserving Discrimination-Freeness Under Bribes**). *Let $\varphi = (\sigma, m)$ be a discrimination-free social choice function. φ preserves discrimination-freeness under bribes if and only if for any agent i, $u_i \in \mathcal{U}$, and $t_{-i} \in T^{n-1}$,*

$$\sigma_i^B(t_i, t_{-i}) = \sigma_i^B(t'_i, t_{-i}) \text{ for all } t_i, t'_i \in T(u_i).$$

With preserving discrimination-freeness under bribes we therefore require that agent i's set of available object assignments when being potentially bribed does not depend on his wealth. Note that we here focus on the object assignments of an agent if he himself is bribed. This reflects a desire to avoid that a change in an agent's wealth influences his decision to accept a bribe that assigns him a worse object.

In what follows, we are primarily interested in necessary and sufficient conditions such that an implementable and discrimination-free social choice function preserves discrimination-freeness under bribes. Obviously, a sufficient condition for preserving discrimination-freeness under bribes is bribe-proofness. With the following proposition we show that for discrimination-free and implementable choice functions bribe-proofness is equivalent to externality-freeness. Nonbossiness of the social choice function makes bribe-proofness a necessary condition for preserving discrimination-freeness under bribes. By externality-freeness we mean that an agent's outcome is independent of other agents' types.

Definition 2.4 (Externality-freeness). *A social choice function is externality-free if for any agent i and any $t \in T^n$ and $t'_{-i} \in T^{n-1}$,*

$$\varphi_i(t_i, t_{-i}) = \varphi_i(t_i, t'_{-i}).$$

Proposition 2.4. *Consider an implementable and discrimination-free social choice function φ. φ is bribe-proof if and only if φ is externality-free. Suppose φ is nonbossy. Then, φ preserves discrimination-freeness under bribes if and only if φ is externality-free.*

Proof. See Appendix. □

The result of the equivalence of bribe-proofness and externatility-freeness is closely related to Schummer (2000a) and Schummer (2000b). His results imply that for very general class of quasilinear preferences over bundles of objects and transfers, bribe-proofness implies that an agent's payoff is independent of other agents' reports. His general idea can be transferred straight forward to the utility domain with non-linear preferences that we consider. Main intuition for the equivalence result is that once an agent can influence another agent's outcome by his report, there exist type profiles such that there is an agent that is willing to pay a certain amount of money to profit from a misreport of another agent. On the other hand, there is an agent that would be willing to accept this amount to misreport in favor of the first agent. To construct those types described we exploit that the social choice function is implementable and discrimination-free and that therefore one's payments are independent of one's type (see Proposition 2.2).[28] Bribe-proofness becomes a necessary condition for preserving discrimination-freeness

[28] With requiring discrimination-freeness we even further restrict the domain of social choice functions considered compared to Schummer (2000a) and Schummer (2000b). Since his argu-

under bribes if whether bribing incentives exist depends on the wealth of the agents. This is the case if the social choice function is nonbossy. Once a bribing incentive exists, the bribing incentive vanishes whenever the agent is rich enough such that the other agents cannot afford any more to bribe this person. Nonbossiness here ensures that the bribe amount that is necessary to bribe is not arbitrarily small. For social choice functions that are not nonbossy, bribes might be quasi-free because there might be an agent who is indifferent between two reports, but his report influences the outcome of another agent.

The examples presented in the beginning of this section on how money might be used outside a system can be interpreted as a special case of bribes. Parameters that influence an outcome (like neighborhood-priority, paying a fee for a private school or subscribing on multiple waiting lists) can be treated as substitutes for preferences of a second side of the market that can be bribed. Preserving discrimination-freeness then requires that whether or not there is an incentive to use money to influence the outcome must not depend on wealth. This can be ensured by making the outcome independent of corruptible parameters.

On Externality-Free Mechanisms Externality-freeness heavily restricts the information about preferences a mechanism designer can use to assign objects. This goes on the cost of efficiency.

To get an intuition for the restrictions consider the problem of assigning goods without transfers. An agent's choice set is the set of objects he can achieve given any report of the other agents. Externality-freeness then is equivalent to the choice set of each agent being constant.[29] Therefore, the choice sets of the agents need to be disjoint such that each object only appears in one choice set. Designing externality-free mechanisms is therefore about designing the distinct choice sets of agents. Independently of the types, n disjoint subsets of Ω need to be build (e.g. via a lottery) and being assigned to the agents. Then for each agent an object is chosen out of the subset that was assigned to this particular agent. This is the only step where type-dependence is allowed. In particular, if there are exactly as many objects as agents, the allocation is constant, i.e., the allocation is type-independent. A simple lottery satisfies this condition. If more objects than agents exist externality-freeness implies wastefulness, i.e., there is a type profile

ments are transferable to our utility domain, the equivalence of bribe-proofness and externality-freeness can be even shown for implementable social choice function. However, we are primarily interested in discrimination-freeness and therefore do not further elaborate on this.

[29] See Schummer (2000a) for a formal description.

such that an object remains unassigned that is preferred by at least one agent to his assigned object.

Corollary 2.3. *Let φ be an implementable and nonbossy social choice function that is discrimination-free. Assume that more objects than agents are available, i.e. $n < k$. If φ preserves discrimination-freeness under bribes then φ is wasteful.*

To see why the corollary holds, first note that $n < k$ implies that for any type profile t there exists an object ω that remains unassigned. If ω belongs to nobody's choice set it remains unassigned for all type profiles, even if somebody ranks ω first. Therefore, φ is wasteful. If ω belongs to the choice set of some agent i it does not belong to any other agent's choice set. In particular, if any agent j other than agent i ranks ω first and all other types remain unchanged, object ω is still unassigned. This implies that φ is wasteful.

Tools to improve the assignment of objects are limited if externality-freeness is desired. One way to improve an externality-free social choice function is increasing the choice sets of the agents by increasing the number of objects or the number of copies. For instance, if each object has at least n copies, each agent can be provided with a choice set containing all objects and therefore can always receive his first choice. Another lever of improvement is how to build the choice sets. The following corollary shows that expected utility of each agent is a concave function in the number of objects available. Hence, given a uniform distribution of types with regard to the valuation of objects, highest expected total welfare is obtained if choice sets of preferably equal size are build. This is because an agent that is facing a choice set of size j randomly chosen out of a set Ω is facing decreasing utility gains when increasing the size j of the set.

Corollary 2.4. *Assume that the agents are homogeneous in the sense that their types are drawn from the same distribution. If every agent receives a random choice set of Ω such that all choice sets are disjoint, the highest expected utility is achieved in case that the differences in size of the choice sets are minimal. Furthermore, total expected utility gains are decreasing with an increase in the number of objects.*

Proof. See appendix. □

2.5 Discussion and Extensions

In the following we discuss some assumptions of the model and illustrate how the basic model presented might be extended to address several settings relevant for real-world applications. In particular, we will highlight the role of the domain $\mathcal{U}$ for the admissible utility functions and the role of the domain of admissible wealth levels.

Budget Constraints. Adding budget constraints to our setup implies that the willingness to pay might exceed the ability to pay. The results derived above then still hold, except that further restrictions on the admissible social choice function might be necessary because a social choice function must not assign payments to an agent that are larger than his wealth. A slight change is needed in Proposition 2.1 because budget constraints do no longer imply that any social choice function with wealth independent transfers is inefficient, but only implies it for social choice functions without transfers.[30]

Type Domain $\mathcal{U} \times \mathbb{R}$. First, consider potential restrictions of $\mathcal{U}$. Whether enlarging or further restricting $\mathcal{U}$ weakens or strengthens the derived results depends on the character of the analysis. For the results on the Pareto-frontier of discrimination-free mechanisms in Proposition 2.1, further restrictions of the domain of admissible utility functions $\mathcal{U}$ only weaken the results. However, when considering implementable social choice functions, the larger the domain $\mathcal{U}$ the more freedom to construct implementable and discrimination-free social choice functions. A further restriction of $\mathcal{U}$ then strengthens the results. It turns out that the proofs of Propositions 2.2, 2.3 and 2.4 do not need the universal character of $\mathcal{U}$. Therein, the domain $\mathcal{U}$ can be restricted to the domain of all utility functions that can be expressed as $u_i(\omega, A) = v_i(\omega) + h_i(A)$ where $v_i : \Omega \to \mathbb{R}$ and $h_i : \mathbb{R} \to \mathbb{R}$ is any function being twice continuously differentiable with $h_i' > 0$, $\lim_{A\to\infty} h_i'(A) = 0$ and $h_i'' < 0$.[31] For Proposition 2.2 and Proposition 2.4 the domain $\mathcal{U}$ can be even further restricted such that all admissible utility function of all agents entail the same fixed $h(\cdot)$. $h(\cdot)$ can be arbitrarily chosen in line with the requirements above. Then, all agents value money in the same way but differ only according to the benefit $v_i(\cdot)$ they attach to each object.

Second, consider the domain of wealth types $\mathbb{R}$. Based on the above discussion about budget constraints, assuming some minimum endowment $\underline{e} \in \mathbb{R}$ does not

[30]Note that considering a model where budget constraints occur but agents have quasilinear preferences does not imply the same results we conducted. In the presence of budget constraints, the willingness to pay is independent of wealth while the ability to pay becomes arbitrarily low if wealth decreases. However, the willingness to accept is independent of wealth. Therefore, compensations agents might receive do not have any consequences for discrimination-freeness which is in contrast to the implications of assuming non-linear preferences.

[31]The condition $\lim_{A\to\infty} h_i'(A) = 0$ is only needed for Proposition 2.1 and can be dropped for the others.

impact on the general analysis. Assuming a maximum endowment $\overline{e} \in \mathbb{R}$, impacts on Proposition 2.1 while it does not impact on the other propositions. The main step of the proofs for the propositions 2.2, 2.3, and 2.4 was to construct utility functions that satisfy certain criteria. In all cases, the construction works whenever the domain of the agents' endowments contains at least two elements. Only if the wealth domain is restricted to one element, requiring discrimination-freeness does not restrict the design of social choice functions. Consequences are different for Proposition 2.1. The result depends on the assumption that for increasing wealth, the willingness to accept becomes arbitrarily large. Restricting wealth endowments restricts the willingness to pay and willingness to accept as well. Then, there are potentially settings such that an agent might be compensated for a worse object by the other agent independent of his wealth level. In particular, in a simple setting with two agents and two goods of which both agents prefer the same, a Pareto-improvement can be performed without violating discrimination-freeness if independent of the wealth distribution, one agent is always willing to pay more for the preferred object than the other agent is willing to accept to give up the preferred object.

Outside Options. In many real-world applications outside options are available. An example is a co-existing private market like private schools or private health insurances. We can integrate an outside option into our model via adding an outcome (ω_o, m_o) with n copies to Ω. Each agent is free to choose the outcome (ω_o, m_o) instead of any other outcome. Hence, we concentrate on social choice functions that assign for any type profile nothing worse than (ω_o, m_o) to each agent.

Adding an outside option mainly implies some further restrictions on implementable and discrimination-free social choice functions compared to those seen in Proposition 2.2. First, if φ is implementable and discrimination-free each agent needs to be assigned to an object that is at least as good as the object of the outside option ω_o. Otherwise agents that are rich enough choose the outside option, independent of how large m_o is. Second, any money assignment of φ has to be greater or equal m_o. Otherwise there exists some agent that prefers his outcome for being rich and the outside option for being poor (see arguments in the proof of Proposition 2.2). In the context of Proposition 2.3 where wealth information is available, adjusting wealth to a constant level is not realizable any more if agents can avoid this redistribution by choosing the outside option.

Two-sided Market. We consider a one-sided market where only the agents that receive the objects have preferences and might act strategically. Whenever providers of the objects are strategic players our notion of discrimination-freeness can be applied for the other side of the market as well.

Seller and Buyer Model. Suppose that a seller owning an object is willing to sell the object for a certain price and some buyer is willing to buy it for a certain price. If the seller's will-sell-price is lower than the buyer's will-buy-price the transaction takes place. Then, discrimination-freeness can be transferred to whether or not the trade takes place must not depend on wealth of both the seller and the buyer. Considering a predetermined and fixed price, our specification of preferences implies that to ensure discrimination-freeness the price has to be zero. For any fixed price that is not zero there are wealth levels such that the transaction takes place and for others not.

Non-constant Ranking. A main assumption on the agents' preferences is that the ranking of objects is wealth independent. Technically, the assumptions of continuity and strict preferences over objects imply constant rankings. Relaxing the assumption of continuity and requiring only continuity from below, ranking of objects might differ with wealth. For instance, wealthier agents might have another first choice than poorer agents. When rankings depend on wealth, it is not straight forward how to define discrimination-freeness. Sticking to our definition implies that even rankings must not play a role for the object distribution. An alternative is to treat agents' preferences as if the ranking was wealth-independent. This might be a valid approach if payments in the mechanism are small enough such that constant rankings are a reasonable approximation. However, then concerns for segregation rather than concerns for discrimination might become relevant.

Assigning Probability Shares. Proposition 2.1 implies that exploiting any information about preferences beyond object rankings does not yield Pareto-improvements compared to a money-free social choice function. Since we are primarily interested in whether money can be used to trade-off cardinalities, in our analysis we concentrate on deterministic outcomes and therefore take an ex-post perspective. When allocating objects, assigning probability shares of objects to the agents might improve ex-ante efficiency since lotteries allow to exploit cardinal information about preferences. Our model can be extended to probabilistic outcomes (with some further specifications on how lotteries are evaluated by the agents). Discrimination-freeness then can be defined as the assignment of probability shares for receiving an object beeing independent of wealth. In analogy to Proposition 2.1, if σ is not ex-ante Pareto-dominated by any discrimination-free matching σ' there is no discrimination-free social choice function with $\sum_{i \in N} m_i = 0$ that ex-ante Pareto-dominates $\varphi = (\sigma, 0)$.[32] However, the resulting social choice function is still not Pareto-efficient within the set of social choice functions with transfers. Furthermore, only if preferences over lotteries are wealth-independent, it is assured that using lotteries to perform ex-ante Pareto-improvements is not in conflict with discrimination-freeness. To elicit cardinal information about preferences for the design of probabilistic assignments, virtual money might be used (compare, for instance, the Pseudomarket described in Hylland and Zeckhauser (1979)). Each agent receives a fixed amount of virtual money that he can split among several objects. Based on this, probability shares are assigned. Here again, only if preferences over lotteries are wealth-independent, it is assured that using probabilistic assignments is not in conflict with discrimination-freeness.

2.6 Applications

The results we obtained in the previous chapters have some interesting applications in markets where wealth-independent access to resources appears to be desirable. Our results provide an explanation why in certain markets preference intensities of agents are not exploited by using transfers. Within real-world applications that distribute object without transfers, there are indeed examples that use externality-free mechanisms by not taking preferences into account but simply using a lottery for distribution. Furthermore, even if money cannot be used to trade-off preference intensities, there might be other ways to account for preference intensities. Sandel (2012), for instance, argues that queuing for a good can be a tool for screening according to preference intensities without using any transfers.

School Choice. Many cities distribute school places via a centralized assignment procedure without using monetary transfers. The probably most popular examples, since extensively discussed in the literature on matching markets, are the school choice procedures in Boston (Abdulkadiroglu, Pathak, Roth, and Sonmez, 2006) and New York (Abdulkadiroğlu, Pathak, and Roth, 2005). School places (at least at public schools) are often fully funded by taxes and parents

[32] Details are available upon request.

do not have to pay additional fees. Furthermore, schooling up to a certain age is compulsory in most countries. Such school assignment procedures are then discrimination-free under the assumption that money cannot be used to influence any parameters of the procedure. Once, for instance, a private sector co-exists that charges fees, discrimination occurs. Or, if living in the neighborhood of a good school is more expensive than living in the neighborhood of a bad school (Black, 1999), the wealthier might have better access to better schools as well. Therefore, if discrimination-freeness is a desire, current assignment procedures may need some revision about whether or not they sufficiently meet this desire.

Kidney Donations. To increase donations from living donors, several models of incentivizing donors are currently discussed intensively. Our model implies that a free market for kidneys leads to discrimination (see also discussion on the seller and buyer model in Section 2.5). Any monetary lump-sum as a compensation for the donor leads to discrimination. Non-cash incentives, on the other hand, do not conflict with discrimination-freeness as long as they incentivize a donation independent on the wealth level of a person. However, not reimbursing cost of donation might lead to discrimination as well as wealthier people are rather willing to bear the costs. Gill, Dong, and Gill (2014) show that in the US the wealthier donate at a higher rate. Another potential source for discrimination in the context of kidney donations are bribes.

In the case of deceased donations it is current policy in many countries that the allocation of kidneys out of the cadaver queue to patients does not depend on subjective preferences intensities of the patients. Kidneys available are distributed based on exogenous factors such as urgency, region, blood type etc. (priority based matching) that makes manipulation very difficult. However, as soon as the report of those exogenous factors can be manipulated, misreports could be incentivized by bribes. This is not just a theoretical case as a 2013 uncovered scandal in Germany regarding transplant corruption shows: doctors manipulated factors that determine priorities for receiving a kidney.[33]

Health Insurance. Health insurance systems of several countries are examples for markets with regulated fees and regulated access. Fees often are mainly based on income characteristics. In several countries, health insurance is compulsory. Furthermore, the assignment of insurees to insurers is often regulated to avoid selection by the insurers. In the US, with Medicare there is made an effort to

[33]See, e.g., BBC (2013).

ensure discrimination-free insurance for people age 65 or older by making everybody eligible for Medicare. In Germany, health insurance is compulsory for all ages, payments are (roughly speaking) a certain percentage of income and people can choose the health insurer of their choice since health insurers must not refuse insurees.[34] Fees are therefore independent of preferences and ensure funding, difference in fees have solely distributive reasons. The assignment of insurees to insurers is even externality-free since everybody receives his first choice. The health insurance system in Germany that was in place until 1996 is another example for an externality-free mechanism. Until then, insurees were automatically assigned to an insurer depending on their occuption and therefore their preferences did not play a role either.

With a co-existing private market discrimination occurs. In Germany, there are indeed ongoing complaints about a two-tier health care system as people above a certain income threshold are free to choose a private insurer. If private insurance does not only mean more comfort but even better health treatment, it leads to discrimination. An example of a country where basically no private health insurance market co-exists is Austria.

Childcare. The assignment of childcare places in Germany is an example for a system that is partly tax-funded but also charges additional income-dependent fees. Local authorities decide on the concrete market design. Most German cities installed a system where parents pay an income-dependent fee which is independent on the specific childcare center chosen. All costs exceeding this fee are funded by the local authority (i.e. via taxes). The specific assignment to the childcare centers then is executed separately from the transfers. Some cities use a decentralized system where parents directly apply at the childcare centers, others use a centralized assignment where parents can submit preferences. Childcare centers do not have an incentive to select parents by income as they receive a lump-sum per child from the local authority. Participation is not mandatory and therefore parents might decide whether to apply for a childcare place. If poorer parents send their child to childcare centers while wealthier parents do not, discrimination might be an issue.

[34] According to the German social security statutes, the health insurance must not decline membership (SGB V § 175: "Ausübung des Wahlrechts: (1) Die Ausübung des Wahlrechts ist gegenüber der gewählten Krankenkasse zu erklären. Diese darf die Mitgliedschaft nicht ablehnen [...]").

2.7 Conclusion

In this paper, we study the problem of assigning indivisible goods to consumers under the constraint that access to goods should not depend on wealth. We find that no information beyond an agent's object ranking can be used for his object assignment, whenever the mechanism cannot (or does not) fully eliminate potential wealth differences in his endowments. Furthermore, ordinal mechanisms that do not use transfers and that are efficient within the set of mechanisms without transfers, are already at the Pareto-frontier of discrimination-free social choice functions. To ensure wealth-independent access to the goods also in cases where money might be used outside a market designer's control even further restrictions are needed such that the object an agent is assigned to must not depend on other agents' preferences.

We, therefore, find that a violation of moral concerns is not equivalent to the presence of money. However, requiring discrimination-freeness restricts to what extent a mechanism can exploit preference information and with it the use of transfers. Thereby, our model explains the very restricted use of transfers in certain markets based on inequality concerns. If there is a use of money outside the mechanisms to improve the access to resources, even further restrictions are required to ensure discrimination-freeness. Some currently used mechanisms are apparently not aligned with discrimination-freeness. Within school choice applications, for instance, if better schools are rather in more expensive neighborhoods, living in a rather expensive neighborhood already implies better access to schools. There are indeed claims for rethinking the current system. The chairman of the Black Alliance for Educational Option wrote: *"If access to high-performing schools has to come down to a number, better it be a lottery number than a ZIP code."*[35] Even if we cannot (and do not want to) deduce any advice as to whether or not to ban transfers, our work is a step into understanding the implications of concerns that underlie the desire to restrict markets. Before deciding to put specific restrictions on markets, a market designer should be aware of grounded desires and take implications of meeting them into account.

This paper contributes to an understanding of the implications of moral concerns behind a desire to ban monetary transfers. There is a branch of questions for further research. For instance, we deferred the question on which markets discrimination-freeness is desired and why. Furthermore, we did not yet consider any trade-offs between discrimination-freeness and efficiency. Knowing more about how preferences depend on wealth in real world applications, can facilitate

[35] See New York Times (2011).

a further differentiation of our results. Even if discrimination is a major concern, there also might be further moral concerns beyond discrimination-freeness. Slippery-slope effects are often feared in the context of an introduction of monetary transfers, even if they are small and regulated. Another concern mentioned, is the exploitation of people in a sense that financial distress might make people unable to decide in their best interest and they might thus regret a decision later. Zargooshi (2001) surveyed people in Iran who sold their kidney after some years. A striking 85% percent of the questioned people indicated that they regret the donation.

2.8 Appendix

Proof of Proposition 2.1

We prove Proposition 2.1 in several steps. First, we argue that there is a maximum amount that each agent is willing to pay for any improvement in the object he is assigned to via φ. This maximum amount can be chosen independently of the wealth endowments of other agents. Second, we show that if φ' Pareto-dominates φ and does not exceed the budget of φ it discriminates. Finally, we show that φ is not Pareto-efficient.

Maximal Willingness to Pay. Fix any utility profile $(u_i)_{i\in N}$ and wealth profile $(e_i)_{i\in N}$. By assumption, each agent i's ex-post wealth level $A_i(t) = m_i(t)+e_i$ does not depend on the other agents' wealth levels. We aim to find some $\overline{M} > 0$ such that for every agent i and any two objects a and b with a being preferred to b by agent i, it holds that

$$u_i(a, A_i - \overline{M}) \leq u_i(b, A_i). \tag{2.1}$$

Then, $\overline{M}$ is such that agent i is not willing to pay more than $\overline{M}$ for an improvement from b to a. Since the set of agents and the set of objects is finite, it is sufficient to show that for any agent i preferring object a over object b we can find $\overline{M}$ such that the inequality above holds. $\overline{M}$ might then depend on i, a and b. We can then take the maximum over all objects and over all agents to define $\overline{M}$ independent of these parameters.

$\overline{M} > 0$ such that 2.1 holds can be defined as the willingness to pay of agent i with wealth A_i for an object improvement from b to a. Formally, define $\overline{M}$ as the solution of the equation $u_i(a, A_i - \overline{M}) = u_i(b, A_i)$. It remains to show that $\overline{M}$ exists and that it is well defined. First note, that if such an $\overline{M}$ exists, it has to be unique since $u_i(a, m)$ is strictly increasing in m. To show the existence, we use that $u_i(a, A_i) > u_i(b, A_i)$. Since $u_i(a, m)$ is strictly increasing in m and strictly concave in m, it has to hold that $u_i(a, m) \to -\infty$ for $m \to -\infty$. Therefore, for some $\overline{M}$ it holds that $u_i(a, A_i - \overline{M}) = u_i(b, A_i)$.

φ' discriminates. Consider any social choice function $\varphi = (\sigma, m)$ such that m does not depend on wealth and assume that σ Pareto-efficient. It is to show that if φ' Pareto-dominates φ and has the same budget as φ has, it discriminates. To prove this, assume that φ' is discrimination-free. We show that this assumption leads to a contradiction. Select some agent i that received a less preferred object under φ' than under φ for some type profile $t = (t_i)_{i \in N}$. Such an agent exists because if for all type profiles nobody faced an object impairment under φ' compared to φ and furthermore φ' Pareto-dominates φ and has the same budget, then σ' needs to Pareto-dominate σ. However, σ was selected such that it is not Pareto-dominated by any σ'.

Now assume that agent i is assigned to a by φ and to b by φ'. Due to discrimination-freeness of φ and φ' agent i is assigned to those objects for any wealth endowments e_i. Pareto-dominance of φ' implies that for every wealth endowment e_i, agent i has to be compensated for receiving object b instead of a by a monetary transfer $M(e_i)$.

The amount $M(e_i)$ that compensates agent i for receiving b instead of a becomes arbitrarily large for increasing wealth: if e_i increases, his ex-post wealth $A_i = m_i(t) + e_i$ becomes arbitrarily large as well since $m_i(t)$ does not depend on e_i. Therefore, the willingness to accept for receiving b instead of a becomes arbitrarily large for increasing wealth.

At the same time, the amount of money that is available to compensate agent i is bounded above by $(n-1)\overline{M}$ when varying agent i's wealth level. Therefore, there exists some wealth endowment e_i of agent i such that agent i cannot be compensated any more by the other agents for the object impairment. Then, φ' is not a Pareto-improvement of φ which is a contradiction.

φ is not Pareto-efficient. To show that $\varphi = (\sigma, m)$ is not Pareto-efficient, we have to find a type profile $t = (t_i)_{i \in N}$ for which $\varphi(t)$ can be Pareto-improved without exceeding the budget of $\varphi(t)$. Consider a type profile $t = (t_i)_{i \in N}$ such that all agents have the same ordinal ranking over objects. Furthermore, choose the endowments e_i of each agent small enough such that for some M each agent is willing to accept at least M in return for an object impairment based on the outcome of φ. This construction can be performed, for instance, by using a utility function $u_i(\omega, A_i) = v_i(\omega) + h(A_i)$ with $h' > 0$, $h'' < 0$ and $\lim_{A_i \to \infty} h'(A_i) \to 0$.

Now consider the assignment of objects $\sigma(t)$. Then, select an agent that did not receive the most preferred object a. Since the transfers of φ do not depend on endowments, increasing the wealth level of agent i does not impact on wealth levels of the other agents. If agent i's level is high enough, he is willing to pay at least M for any object improvement. All other agents are still willing to accept M for any object impairment. Therefore, there are two agents that are both better off if they trade objects in turn for money. Since this is a Pareto-improvement φ cannot be Pareto-efficient.

Proof of Proposition 2.2

Throughout the proof we concentrate on the outcome of an agent i and fix the type of the other agents t_{-i}. Therefore we omit t_{-i} in the notation. First, we show that an agent i's monetary transfer is independent of his type. Second, we use this to show that his object assignment only depends on his preferences through his ordinal ranking.

Dependence of m_i on t_i. Suppose φ is discrimination-free and implementable and agent i's payment is *not* type-independent. Then there exist two types $t_i = (u_i, e_i)$ and $t_i' = (u_i', e_i')$ with $m_i(t_i) < m_i(t_i')$. Implementability of φ requires that for the two types t_i and t_i' the objects they are assigned to differ. Implementability furthermore implies that $|\varphi(T)| \leq k$ where $\varphi(T)$ is the set of all outcomes that agent i can reach by varying his report. This is because any two outcomes in $\varphi(T)$ need to differ regarding the object they contain. By assumption, $\varphi(T)$ contains at least two elements that differ in their money assignment. Let (b, m) be the assignment in $\varphi(T)$ with the highest monetary assignment and (a, m') any other outcome in $\varphi(T)$ with $m' < m$.

We now aim to construct a utility function u_i^* and find two wealth levels e_i^1 and e_i^2 such that agent i's object assignment differs for reporting $t_i^1 = (u_i^*, e_i^1)$ and $t_i^2 = (u_i^*, e_i^2)$. This then contradicts discrimination-freeness and therefore completes the proof. We choose $u_i^* \in \mathcal{U}$, e_i^1, and e_i^2 such that

- Object a is the most, object b the second most preferred object
- For e_i^1, (b, m) is preferred over (a, m')
- For e_i^2, (a, m') is preferred to (b, m).

For any $e_i^1 < e_i^2 \in \mathbb{R}$, we can construct u_i^*, for instance, by $u_i^*(\omega, A) = v_i(\omega) + h(A)$ with any $h : \mathbb{R} \to \mathbb{R}$ and $h' > 0$ and $h'' < 0$. $v_i(a)$ and $v_i(b)$ are chosen such that $v_i(a) - v_i(b) < h(m+e_i^1) - h(m'+e_i^1)$ and $v_i(a) - v_i(b) > h(m+e_i^2) - h(m'+e_i^2)$. Therefore, for a wealth level of e_i^1 and utility according to u_i^* agent i prefers receiving object b in combination with a transfers of m to all other bundles that can be reached. An increase in agent i's wealth level from e_i^1 to some e_i^2 results in agent i not preferring (b, m) anymore to all other bundles in $\varphi(T)$. Implementability then implies that the object assignment of agent i depends on his wealth. This is a contradiction to discrimination-freeness.

Dependence of σ_i on t_i. Consider two types t_i and t_i' that represent the same object ranking r_i, i.e. $t_i, t_i' \in T(r_i)$. From the first part of the proof we know that $m_i(t_i) = m_i(t_i')$. Implementability of φ implies that $\sigma_i(t_i) = \sigma_i(t_i')$ because otherwise either t_i or t_i' would have an incentive to deviate. Therefore, agent i's object assignment only depends on his rank order list of objects.

Proof of Proposition 2.3

Throughout the proof we concentrate on the outcome of an agent i and fix the type of the other agents t_{-i}. Therefore we omit t_{-i} in the notation. It is sufficient to show that m_i is independent of u_i. Then it follows in analogy to the proof of Proposition 2.2 that σ_i is not sensitive to cardinal information of u_i.

Assume that φ is discrimination-free and implementable and that ex-post wealth is not constant (see Proposition). We show that assuming that m_i is *not* independent of u_i results in a contradiction. For this, we construct a preference profile u_i^* such that there are two types $t_i, t_i' \in T(u_i^*)$ that only differ in their wealth level but receive different objects. This then contradicts discrimination-freeness and therefore, m_i has to be independent of u_i.

Construction of u_i^*. If m_i is not independent of u_i there exists e_i, u_i and u_i' such that $m_i(u_i, e_i) < m_i(u_i', e_i)$. Choose e_i' such that $A_i = e_i + m_i(u_i, e_i) \neq e_i' + m_i(u_i, e_i') = A_i'$. Such an e_i' exists because ex-post wealth is not constant.

In the following it is convenient to consider choice sets of agents given their wealth endowment. A choice set $C_{e_i}(\mathcal{U})$ is the set of all bundles of objects and ex-post wealth available to an agent i with wealth endowment e_i by varying his report (t_{-i} is still fixed). Formally,

$$C_{e_i}(\mathcal{U}) = \{(\sigma_i(u_i, e_i), m_i(u_i, e_i) + e_i) | u_i \in \mathcal{U}\}.$$

Implementability of φ implies that two different bundles in $C_{e_i}(\mathcal{U})$ need to differ in their object (otherwise φ cannot be implementable) and therefore $C_{e_i}(\mathcal{U})$ contains at most k bundles. Furthermore, define $a = \sigma_i(u_i, e_i)$ and $b = \sigma_i(u'_i, e_i)$. $a \neq b$ holds because φ is implementable and $m_i(u_i, e_i) < m_i(u'_i, e_i)$. Then, for the wealth endowment e_i the bundles (a, A_i) and $(b, A_i + x)$ with $x > 0$ are in the choice set $C_{e_i}(\mathcal{U})$ of agent i. On the other hand, for e'_i the bundles (a, A'_i) and $(b, A'_i + x')$ with some $x' \in \mathbb{R}$ are in the choice set $C_{e'_i}(\mathcal{U})$. This is because if only agent i's wealth varies, the objects that can be reached by varying the preferences need to be the same due to discrimination-freeness.

We now aim to construct a utility function u^*_i such that the object of the most preferred bundle in $C_{e_i}(\mathcal{U})$ differs from the object of the most preferred bundle in $C_{e'_i}(\mathcal{U})$ given preferences u^*_i. Implementability then implies that φ needs to assign different objects to an agent with preferences u^*_i for wealth e_i and e'_i.

To construct u^*_i, we first consider $x' \leq 0$. Then consider any u^*_i such that a is the most preferred object and b the second most preferred object, and $(b, A_i + x)$ is the most preferred bundle in $C_{e_i}(\mathcal{U})$. This is feasible with any utility function of the shape $u^*_i(\omega, A) = v_i(\omega) + h_i(A)$ with $h'_i > 0$, $h''_i < 0$. Since a is preferred over b and $x' \leq 0$, it holds that $u^*_i(a, A'_i) > u^*_i(b, A'_i + x')$. Therefore, the most preferred bundle in $C_{e'_i}(\mathcal{U})$ does not entail object b. This contradicts discrimination-freeness.

Second, consider $x' > 0$. Again, consider a utility function of the shape $u^*_i(\omega, A_i) = v_i(\omega) + h_i(A_i)$ with $h'_i > 0$, $h''_i < 0$. Here, let $h_i(\cdot)$ be such that $h_i(A_i + x) - h_i(A_i) \neq h_i(A'_i + x') - h_i(A'_i)$. This is feasible since $A_i \neq A'_i$. Choose $v_i(\omega)$ such that object a is the most preferred object and object b the second most preferred one.

Furthermore, for $h_i(A_i + x) - h_i(A_i) < h_i(A'_i + x') - h_i(A'_i)$ let $v_i(a)$ and $v_i(b)$ be such that

$$h_i(A_i + x) - h_i(A_i) < v_i(a) - v_i(b) < h_i(A'_i + x') - h_i(A'_i).$$

For all other objects that might be entailed in bundles of $C_{e_i}(\mathcal{U})$ assume that the distance in valuation to objects a and b are large enough, such that those bundles are never preferred bundles in $C_{e_i}(\mathcal{U})$ for u^*_i. Then, $(a, A_i + x)$ is the most preferred bundle in $C_{e_i}(\mathcal{U})$ but the most preferred bundle in $C_{e'_i}(\mathcal{U})$ does not entail a. This contradicts discrimination-freeness.

For $h_i(A_i + x) - h_i(A_i) > h_i(A'_i + x') - h_i(A'_i)$ choose

$$h_i(A_i + x) - h_i(A_i) > v_i(a) - v_i(b) > h_i(A'_i + x') - h_i(A'_i)$$

Again, for all other objects that might be entailed in bundles of $C_{e_i}(\mathcal{U})$ assume that the distance in valuation to objects a and b are large enough, such that those bundles are never preferred bundles in $C_{e_i}(\mathcal{U})$ for u_i^*. Then, $(b, A_i + x)$ is the most preferred bundle in $C_{e_i}(\mathcal{U})$ but the most preferred bundle in $C_{e_i'}(\mathcal{U})$ does not entail b. This contradicts discrimination-freeness.

Proof of Proposition 2.4

In the following we assume that φ is an implementable and discrimination-free social choice function.

Bribe-proofness $\Leftrightarrow$ Externality-freeness: It is straight forward to show that externality-freeness implies bribe-proofness: If no agent can influence another agent's outcome it never pays off to pay somebody else to state other preferences. Since φ is implementable, no agent has an incentive to misreport. This implies that no bribing incentives exist such that an agent i is bribing himself with $\tau = 0$. Therefore, φ is bribe-proof.

We now show that bribe-proofness implies externality-freeness. To ease notation we denote for an agent of type t_i the strict preferences over outcomes by P_i, the weak preferences by R_i, and indifferences by I_i. The proof proceeds in two steps. First, we show that if φ is bribe-proof, then for any agent i another agent j's report does not influence his utility, i.e., $\varphi_i(t_j, t_{-j}) I_i \varphi_i(t_j', t_{-j})$. Then, we show that it implies $\varphi_i(t_j, t_{-j}) = \varphi_i(t_j', t_{-j})$. Externality-freeness, i.e., $\varphi_i(t_i, t_{-j}) = \varphi_i(t_i, t_{-j}')$ then follows by induction. Whenever reports of other agents are fixed in the following, it is omitted in the notation for better readability.

Bribe-proofness Implies $\boldsymbol{\varphi_i(t_j, t_{-j}) I_i \varphi_i(t_j', t_{-j})}$. Assume the contrary holds such that there is some $t_{-j} \in T^{n-1}$ fixed and $t_j, t_j' \in T$ with $\varphi_i(t_j') P_i \varphi_i(t_j)$. We show that this assumption produces a contradiction because we can find a type profile such that agent i has an incentive to bribe another agent.

Continuity of the preferences in money implies the existence of $\delta > 0$ such that $(\sigma_i(t_j'), m_i(t_j') - \delta) P_i \varphi_i(t_j)$ (i would pay δ to change type t_j's report from t_j to t_j').

We now consider a utility function u_j^* that represents the same ordinal ranking as u_j does and a wealth level e_j^* such that

$$(\sigma_j(t_j'), m_j(t_j') + \delta)) P_j^* \varphi_j(t_j^*) \text{ with } t_j^* = (u_j^*, e_j^*)$$

This construction is feasible since whenever u_j^* represents the same ordinal ranking as u_j does, the outcomes for the two utility functions are the same - for any wealth levels: since φ is implementable, reporting type t_j' instead of t_j^* needs to yield a weakly worse outcome for agent j if agent j has a type t_j^*. Since m_j must not depend on the report (an implication of discrimination-freeness), the object assignment needs to be weakly worse than the one for reporting t_j^*. Then, for instance, for any u_j^* such that $u_j^*(\omega, A_j) = v_j(\omega) + h(A_j)$ with $h' > 0$ and $h'' < 0$ it is feasible to choose $v_j(\cdot)$ such that the equation above is satisfied.

By the discussion above, $(\sigma_i(t_j'), m_i(t_j') - \delta)P_i\varphi_i(t_j)$ holds. While the outcome for agent j is independent of whether reporting t_j or t_j^*, the outcome for agent i might be different. Whenever agent i prefers the outcome for a report t_j^* compared to t_j, he has an incentive to bribe an agent j that has type t_j with any amount $\tau < \delta$ (since j is anyway indifferent between reporting t_j or t_j^*). So assume that the outcome for a report t_j^* is weakly worse for agent i compared to a report t_j. Then agent i has an incentive to bribe agent j that has type t_j^* with an amount $\tau = \delta$ in order to report t_j'. Therefore there exists an incentive to bribe which completes the proof.

$\varphi_\mathbf{i}(\mathbf{t_j}, \mathbf{t_{-j}})\mathbf{I_i}\varphi_\mathbf{i}(\mathbf{t_j'}, \mathbf{t_{-j}})$ Implies $\varphi_\mathbf{i}(\mathbf{t_j}, \mathbf{t_{-j}}) = \varphi_\mathbf{i}(\mathbf{t_j'}, \mathbf{t_{-j}})$. Suppose the contrary: For any agent i, $t_{-ij} \in T^{n-2}$ fixed, and $t_i, t_j, t_j' \in T$ it holds that $\varphi_i(t_j, t_{-j})I_i\varphi_i(t_j', t_{-j})$, but

$$(a, m_1) = \varphi_i(t_i, t_j) \neq \varphi_i(t_i, t_j') = (b, m_2).$$

It implies that $a \neq b$ and $m_1 \neq m_2$ because otherwise, agent i cannot be indifferent. Without loss of generality assume that $m_1 > m_2$. Now consider any agent i with a type t_i^* such that t_i^* represents the same ordinal ranking as t_i does but it holds that $(a, m_1)P_i^*(b, m_2)$. Since φ is implementable and discrimination-free, reporting t_i and reporting t_i^* need to yield the same outcome for agent i. Therefore,

$$\varphi_i(t_i^*, t_j) = (a, m_1) \text{ and } \varphi_i(t_i^*, t_j') = (b, m_2).$$

Furthermore, the first part of the proof implies that $\varphi_i(t_i^*, t_j)I_i^*\varphi_i(t_i^*, t_j')$ holds which is a contradiction to the construction of t_i^* such that (a, m_1) is strictly preferred over (b, m_2).

Nonbossy Social Choice Functions: By the first part of the proposition externality-freeness is equivalent to bribe-proofness. Furthermore, bribe-proofnees implies that discrimination-freeness under bribes is preserved. Therefore, it remains to show that if φ is nonbossy and preserves discrimination-freeness under bribes, then φ has to be bribe-proof.

Assume that φ is implementable and preserves discrimination-freeness under bribes but is not bribe-proof. Then, there exists $t = (t_i)_{i\in N}$ such that an agent j has an incentive to bribe $i \neq j$. Since φ is nonbossy, the outcome for agent

i needs to differ when being bribed in order to report t'_i instead of t_i. Due to implementability, the object agent i receives for t'_i is worse than it is for t_i (since the money assignment is independent of the type). Therefore, $\sigma_i^B(t_i, t_{-i})$ contains an object assignment that is worse than the one for a report t_i. Furthermore note that the choice set of agent i, i.e., the set of bundles that agent i can reach by varying his report, has at most $|\Omega| = k$ elements and is therefore finite. Since φ is nonbossy, the number of different outcomes for each agent that can be reached by a variation of a report of agent i is therefore also finite. Therefore, there is some $\overline{M} > 0$ such that any agent is not willing to pay more than $\overline{M}$ in order to bribe agent i independent of agent i's type.

Now consider a utility function u_i^* such that u_i^* represents the same ordinal ranking as u_i does and two wealth levels e_i^1 and e_i^2 such that agent i with type $t_i^1 = (u_i^*, e_i^1)$ is willing to accept a bribe of agent j but agent i with type $t_i^1 = (u_i^*, e_i^2)$ is not willing to accept the bribe and is even not willing to accept anything less than $\overline{M}$ to change his report. This construction is feasible since t_i, t_i^1 and t_i^2 yield the same outcome for agent i. Furthermore, outcomes for the other agents are also independent of whether agent i reports t_i, t_i^1, or t_i^2 (due to nonbossiness). Therefore, no agent has an incentive to bribe agent i. $\sigma_i^B(t_i^1, t_{-i})$ with $t_i^1 = (u_i^*, e_i^1)$ contains at least one element that is worse than the object assignment for a report t_i. $\sigma_i^B(t_i^2, t_{-i})$ with $t_i^2 = (u_i^*, e_i^2)$ contains only the object that is assigned for a report t_i. This contradicts preserving discrimination-freeness under bribes which proves the desired.

Proof of Corollary 2.4

We show the corollary by showing that an agent facing a choice set of size j randomly chosen out of a set Ω is facing decreasing expected utility gains. For any agent i let Z_j denote the random variable that describes the element with maximal utility of a randomly chosen subset of Ω of size j. Let $\mathbb{E}(Z_j)$ denote the expected utility of Z_j for agent i. We have to show that the marginal utility gain of raising j is decreasing meaning that

$$\mathbb{E}(Z_{j+1}) - \mathbb{E}(Z_j) \leq \mathbb{E}(Z_j) - \mathbb{E}(Z_{j-1}).$$

We order the objects with respect to the valuation of the objects, a_1 denotes the object with the lowest valuation, a_k the object with the highest valuation. We consecutively draw objects out of the set $\{a_1, ..., a_n\}$, Y_j denotes the random variable representing the j-th draw. Then we can write the random variable Z_j as $Z_1 = Y_1$ and $Z_j = \max\{Y_j, Z_{j-1}\}$ for $j > 1$. By using conditional expectation it is then sufficient to show that

$$\mathbb{E}[(Z_j - Z_{j-1})1_{Y_1=a_{i_1},\ldots,Y_{j-1}=a_{i_{j-1}}}] \geq \mathbb{E}[(Z_{j+1} - Z_j)1_{Y_1=a_{i_1},\ldots,Y_{j-1}=a_{i_{j-1}}}]$$

for any possible sequence of draws $a_{i_1}, \ldots, a_{i_{j-1}}$. However, this just depends on the value of Z_{j-1} and therefore it is sufficient to prove this for $j = 2$. This can be done by explicit calculation.

CHAPTER 3

QUALITY PROVISION AND REPORTING WHEN HEALTH CARE SERVICES ARE MULTI-DIMENSIONAL AND QUALITY SIGNALS IMPERFECT

Abstract

We model competition for a multi-attribute health service where patients observe attribute quality imprecisely before deciding on a provider. High quality in one attribute, e.g. medical quality, is more important for ex-post utility than high quality in the other attribute. Providers can shift resources to increase expected quality in some attribute. Patients rationally focus on attributes depending on signal precision and beliefs about the providers' resource allocations. When signal precision is such that patients focus on the less important attribute, any Perfect Bayesian Nash Equilibrium is inefficient. Increasing signal precision can reduce welfare, as the positive effect of better provider selection is overcompensated by the negative effect that a shift in patient focusing has on provider quality choice. We discuss the providers' strategic reporting incentives and reporting policies. Under optimal reporting, signals about the important attribute are always published. However, banning reporting on less important attributes might be necessary.

3.1 Introduction

Health care services have multiple relevant quality dimensions. When choosing doctors, hospitals or taking decisions about nursing homes, patients care about medical quality on the one hand, and may take non-medical quality factors such as general appeal of the doctor's office or hospital environment, short waiting times and interpersonal skills of the staff on the other hand into account. Some of these dimensions are difficult to observe, measure, evaluate and communicate, whereas others can be observed and measured with fairly high precision. For instance, selected mortality rates or Coronary Artery Bypass Graft (CABG) rates provide only an imprecise signal of hospital medical quality.[1] Contrary to that,

[1] Iezzoni (1997) shows that report card rankings may vary profoundly according to the chosen risk adjusters. Thus, if patients do not have information about the risk adjusters used,

information brochures with pictures of patient rooms and sample dinner menus provide fairly accurate signals for the hotel attributes of the hospital environment. In Germany, for instance, the public feedback platform *Arztnavigator* provides detailed information of patient feedback on doctor's practice rooms, waiting times, and the doctor's and staff's friendliness and communication skills.[2]

In this paper, we address the question of which quality dimensions patients rationally focus on when the signals they receive about the qualities of the dimension before deciding on a provider have different precision, and what this focussing implies for the provision of quality and welfare. In particular, we are concerned with settings where patients value quality differences in one attribute, e.g. medical quality of the service, more than quality differences in the other attributes, e.g. the hotel properties of hospitals or nursing homes, but the quality signal in the more important attribute is less precise.

Interestingly, empirical research indicates that public reporting of clinical quality scores has a positive but only weak effect on patients' provider choice.[3] One reason might be that patients are skeptical about the accuracy of these quality measures. Furthermore, other quality dimensions might play an important role for the choice of health care providers. Goldman and Romley (2008) analyze the role of amenities alongside treatment quality measures on hospital choice for Californian data. They show that various measures of treatment quality of hospitals (e.g. mortality rates) have only a small effect on patient demand while improvements in amenities strongly raise demand. Furthermore, patients' perceptions of reputation and specialty medical services as well as satisfaction with a prior hospital stay significantly affect hospital choice. Among these, satisfaction with a prior stay may thereby be driven partly by non-medical factors. Fornara, Bonaiuto, and Bonnes (2006) e.g. show that hospital users' perceived quality of care improves when the humanization degree of the hospital environment increases.[4] Regarding the demand response, Dafny and Dranove (2008) report that the effect of health plan

there is significant noise. According to Dranove (2000), Medicare Hospital Compare identifies only a small percentage of hospitals as having mortality rates significantly above or below the mean. Thus, although quality reports become increasingly available through e.g. report cards or public feedback platforms, the signals that patients receive through these about medical quality are often still fairly imprecise through an inherent difficulty of observing and measuring and interpreting medical quality accurately.

[2]See Arztnavigator (2015).

[3]See e.g. Dranove (2000) and the discussion therein.

[4]For environmental factors, Arneill and Devlin (2002) conducted a study where they showed participants slides of doctors' waiting rooms and then asked what quality of care participants expected. Arneill and Devlin (2002) find that a significantly higher perceived quality of care for waiting rooms that are nicely furnished, light, contain artwork and are warm versus waiting rooms that are dark, have outdated furnishings, contain no artwork or poor quality reproductions and are cold in appearance.

report cards on Medicare beneficiaries is driven by responses to patient satisfaction scores, while other more objective quality measures did not affect enrollment decisions.

An important concern in this context is whether a potentially strong demand response to non-medical quality attributes such as amenities, interpersonal skills or perceived high quality environment leads to a suboptimal quality of care. This would be the case if medical quality is more important to generate patient welfare than all other dimensions of care - such that quality of care should be high on the clinical quality dimension -, but health care providers do not provide sufficiently high quality in the clinical dimension as patient demand is more responsive to quality differences in other dimensions. However, why should patients respond more to quality differences in other dimensions than medical quality if medical quality is the important dimension in terms of their realized utility? Generally, why would patients focus on an attribute that is less important in terms of consumption utility?

Our starting point is the observation that many quality dimensions can only be observed imperfectly ex-ante, and that the precision of information about quality varies across dimensions. In particular, we model provider competition when patients observe attribute quality of a two-attribute health service only imperfectly. Providers can allocate given resources across the attributes in order to increase expected quality in either one or the other attribute. A patient's utility gain from an increase in quality in one attribute is larger than in the other attribute, thus representing the situation where high quality in the medical treatment dimension is more important for patient welfare than amenities. Patients receive a binary signal about realized quality in each attribute from each provider before deciding on a provider.

We first define rational focusing on attributes: A patient focuses on an attribute if a high quality signal in this attribute drives her provider choice. We say that focusing is strong if this holds for any combination of beliefs that the patient might have about the underlying resource allocation decisions of the providers, whereas there is focusing, but not strong, if this holds for beliefs that are symmetric across providers. With this definition, we can describe a patient's focus on quality attributes depending on the precision of quality signals in the attributes.

We show that equilibria exist in which providers invest in the less important attribute. This occurs if the quality signal in this attribute is more precise than in the other attribute to the extent that patients focus on this attribute. Equilibrium is unique under strong focusing. If signal precisions are such that patients' focus is on the less important attribute, all Perfect Bayesian Nash equilibria are inefficient. Increasing signal precision, e.g. by introducing a signal in the less important attribute, can reduce welfare. This occurs if the positive effect of better provider selection due to higher signal precision is overcompensated by the negative effect that the shift in patient focusing, induced by the change in signal precision, has on provider quality choice. We derive conditions under which an increase in signal precision leads to an unambiguous welfare loss.

In the literature on health care reporting, the adverse effect of information that has been emphasized is providers' patient selection incentives (Dranove, Kessler, McClellan, and Satterthwaite, 2003), i.e., turning away the sickest patients because of providers' concerns about their 'ratings'. We point to a further effect that may result from the increase in information on other quality dimensions through e.g. public feedback platforms alongside the increased public reporting of medical quality: If information becomes relatively more precise on less important attributes, patients may focus on these, with adverse consequences for quality provision and welfare.

Feng Lu (2012) analyzes the impact of public reporting of some quality measures on quality in the reported and unreported dimensions. Feng Lu (2012) finds that after the introduction of public reporting, scores of quality measures improve along the reported dimensions, but significantly deteriorate along the unreported dimensions.[5] Feng Lu (2012) furthermore finds no evidence that there was a decrease in quality-related inputs, suggesting a reallocation of resources. Note that in our model, public reporting only has an effect on the resource allocation if it increases the relative precision of quality signals that patients receive in these attributes, and only if the effect is strong enough to shift patient focus.

Our analysis also allows to derive optimal reporting policies. Reporting in our framework is the sending of informative but noisy signals about realized quality with exogenous precision before quality is realized. In order to compare reporting policies including voluntary reporting, we change the baseline model in the following way: Whether patients receive signals (with exogenous precision) in certain attributes now depends on a strategic reporting decision by providers. We show

[5] Contrary to that, Werner, Konetzka, and Kruse (2009) find that overall both unreported and reported care in nursing homes improved following the launch of public reporting. Improvements in unreported care were particularly large among facilities with high scores or that significantly improved on reported measures. Low-scoring facilities experienced no change or worsening of their unreported quality of care. In our model, the technology is such that expected qualities in the dimensions are substitutes and not complements.

that if the more important attribute is not too important, in the unique equilibrium under strategic reporting providers invest in the less important attribute and only publish signals in this attribute. Thus, not only resource allocation, but also reporting might be inefficient. However, if the more important attribute is sufficiently important, it might also be the case that providers invest in the important attribute and only report in the important attribute although there would be patient focusing on the less important attribute if patients received signals in all attributes. Mandating full reporting might be then be welfare-reducing. Under optimal reporting, signals in the important attribute are always published, however, it might be necessary to control reporting in attribute 2. In particular, a ban on reporting in attribute 2 might have to be imposed.

3.2 Related Literature

Focusing. We define rational focusing via the precision of signals that patients receive about attributes in an environment with imperfect quality information. A patient evaluates signals according to her expected utility for any given beliefs. We say that she focuses on an attribute if, for given ranges in feasible outcomes, the difference in the precision of signals is such that the difference between signal value and expected outcome in this attribute is, compared to the other attribute, low. Focusing here is thus different from focusing and salience models (Bordalo, Gennaioli, and Shleifer, 2013, Koszegi and Szeidl, 2013) that assume that there is an exogenous difference between decision utility and consumption utility. In Koszegi and Szeidl (2013) e.g., under perfect information, focus weights of attributes in decision utility depend positively on the range of feasible outcomes in attributes.

Multi-attribute goods. The literature on markets with multi-attribute goods and quality investment is scarce. Bar-Isaac, Caruana, and Cuñat (2012) analyze monopoly provision of a two-attribute good where quality is imperfectly observable. Contrary to our set-up with exogenous information, they consider active consumers who choose which information to acquire. Customers are heterogenous in their valuation for attributes and can assess quality at a cost. The monopolist can invest in an increase of the probability of high quality in one attribute. A reduction in the consumers' costs of acquiring information on the other attribute may then reduce quality investment: The decrease in costs of assessment shifts the consumer that is indifferent between assessing one or the other dimension towards the first attribute, reducing demand and thereby quality investment. The direct positive welfare effect of reduced assessment costs may then be dominated by the negative investment effect leading to a reduction in overall consumer welfare. Closest to our work is Dranove and Satterthwaite (1992). In Dranove and Satterthwaite (1992), competing manufacturers sell goods through retailers where retail price is random and customers are heterogenous in their valuation for quality. Customers observe prices and quality only with noise and search retailers using an optimal sequential search rule. An increase in the precision of the price observation may then decrease welfare through the indirect effects of a change in the customers' search: Prices fall, but quality is reduced as well. If the latter effect is stronger, increasing precision of the price observation reduces consumer welfare. In contrast, we model a market with homogeneous consumers that benefit more from high quality in one attribute than in the other. Instead of searching, customers receive signals from all providers. We show under what conditions on signal precision and beliefs the customers' focus is on the less important attribute and derive the welfare consequences. Furthermore, we discuss strategic reporting by providers and optimal reporting policies. While the workings in our model show some analogy to the logic of the multitasking literature as in Holmstrom and Milgrom (1991), the modelling and conclusions are however different. In the multitasking literature, effort substitutability implies complementarity of the optimal (linear) incentive pay for tasks.[6] Better information in the sense of a reduction in the noise of the performance improves the tailoring of incentive pay and does not have a negative value for the principal. In contrast, we consider a market for a multi-attribute service where consumers receive noisy signals about realized quality by competing providers. The key contractual incompleteness in this market is that attributes cannot be separately priced such that consumers do not separately evaluate expected quality and utility differences in each attribute and that consumers cannot commit to ignore signals. Better information in the sense of increasing signal precision may then decrease welfare, as it is individually rational for customers to focus too strongly on signals in the less important attribute.

Health care quality under imperfect information and quality reporting. Gravelle and Sivey (2010) analyze competition between hospitals under

[6]Kaarboe and Siciliani (2011) analyze optimal contracting between a purchaser and a partly altruistic provider of health services within the multitasking framework where one quality dimension is verifiable whereas the second is not. They show that provider altruism with respect to health benefit can lead to overall complementarity of qualities even if they are substitutes on the effort cost side such that high powered incentives may be optimal.

fixed prices where patients receive imperfect signals about quality, which is one-dimensional. Hospitals have different quality cost functions and can set quality. Gravelle and Sivey (2010) show that when patients choose the hospital that sends a higher signal, better information in the sense of a reduction in the variance of the noise term may reduce quality of both hospitals if quality costs are sufficiently different.[7]

Most of the literature on quality information considers reporting in the form of disclosure of known, realized quality. Sun (2011) analyzes a monopolist's voluntary disclosure for a multiple-attribute good, where the attributes are a vertical and horizontal quality. When vertical quality is known, horizontal quality might not be disclosed. This is since a monopolist benefits by disclosure through attracting consumers nearby at the cost of deterring consumers far away. When vertical quality is low, the benefit outweighs the cost. The higher the quality, the more likely the consumer is to buy the product without disclosure such that when quality is high enough, the monopolist tries to cover the entire market at a high price without disclosure. Board (2009) analyzes disclosure incentives for a one-dimensional good under competition with heterogeneous firms. If a high-quality firm discloses, competitors must trade off the increase in competition and resulting fall in price if they also disclose with the reduction in perceived quality by consumers, if they do not. Nondisclosure by some high-quality firms thus generates positive externalities for low-quality firms who may pool with them and take advantage of raised consumer expectations. Board (2009) shows that the welfare effects of mandatory disclosure are complex, consumer surplus however rises if firms are sufficiently close in quality that the overall effect is increased competition. Contrary to that, we do not model quality disclosure, but reporting as a decision of publishing signals before quality is realized. Providers voluntarily never report in all attributes, since reporting in their weak attribute, i.e. the one they did not invest in, gives them a competitive disadvantage.

Quality reporting as a policy instrument in the context of healthcare is considered in Glazer and McGuire (2006). Glazer and McGuire (2006) study competition among health plans under adverse selection and fixed prices. They show that averaged quality reports, instead of full reports, can remedy adverse selection incentives, since averaging quality across dimensions and reporting only the average enforces pooling in health insurance. Less information in the form of averaged quality reports thus mitigates the problem of cream-skimming of good patients

[7]Patient demand is however not consistent with that of rational Bayesian agents, see the discussion in Shelegia (2012).

with tailored quality packages. The right weights for quality averaging may then implement efficient outcomes. Whereas Glazer and McGuire (2006) consider a common value set-up and fixed prices, Ma and Mak (2014) compare full quality reporting to average quality reporting under private values and price setting by a monopolist. Ma and Mak (2014) show that qualities and prices under an imposed average quality report generate higher consumer welfare than full quality report, as it restrains the firm's price-quality discrimination strategies. In our model, suppressing quality information in the form of banning reporting in some dimensions might be optimal since this shifts the patients' demand towards the quality dimensions that matter more to generate welfare.

3.3 Model

We consider a two-attribute health service $q = (q_1, q_2)$ with $q_i \in \{h, l\}$ for $i = 1, 2$ where h stands for high quality and l for standard quality respectively. Two providers A and B provide the service. The provider compensation is a uniform, exogenously set fee $P > 0$ per unit of service provided.[8] Quality cannot be contracted on.

Quality is stochastic. Providers can allocate resources in order to achieve high expected quality in either one or the other attribute.[9] In particular, each provider $j \in \{A, B\}$ has fixed resources which are symmetric across providers, and makes a resource allocation decision $a^j \in \{0, 1\}$. For any $a^j \in \{0, 1\}$ the realization probabilities for high quality in one attribute are

a^j	$\mathbb{P}(q_1 = h)$	$\mathbb{P}(q_2 = h)$
1	$1 - p$	p
0	p	$1 - p$

with $p \in (0, \frac{1}{2})$. Quality levels are realized independently for each attribute. With this technology, we say provider j invests in attribute 1 (2) if he sets $a^j = 1$ ($a^j = 0$). The lower p, the larger is the probability that high quality is realized in the attribute a provider invests in.

[8] Fees cannot be set separately for attributes. The fixed, exogenous fee reflects e.g. regulated prices or negotiated prices between health plans and providers for the service in their network.

[9] For a potential split of resources see discussion in section 3.8.

The assumptions made about how quality realization depends on the resource allocation incorporates two symmetries: First, a symmetric impact of resource allocation on quality realization across attributes. This is in order to make attributes perfectly symmetric on the technology side, as our focus is on differences across attributes on the demand side. Second, the modelling implies symmetry across high and low quality realization. The second one is mainly used for simplification. It particularly implies that the probability that high quality is realized in attribute i if invested in i equals the probability that low quality is realized if invested in the other attribute. Both symmetries are discussed in detail in section 3.8 where we also argue why giving up those symmetries basically preserves our results. Variable costs of providing the service are set to 0. Providers maximize expected profit, which will be equal to maximizing market share since the fee for the service is fixed.

There is a continuum of patients C in the market with mass 1. Each patient $c \in C$ receives utility $u(q)$ from utilizing a health service with quality $q = (q_1, q_2)$ that is additively separable in attributes, i.e. $U(q) = \sum_{i=1}^{2} u_i(q_i)$.[10] We assume that the utility gain from high quality versus standard quality is higher in the first attribute than in the second attribute, i.e.

$$\theta \equiv \frac{u_1(q_1 = h) - u_1(q_1 = l)}{u_2(q_2 = h) - u_2(q_2 = l)} > 1.$$

Thus, high quality in attribute 1 is more important to generate increases in patient utility than high quality in attribute 2, in the following we refer to this property when we say that attribute 1 is the important attribute. In many health care applications, attribute 1 could be thought of as the medical quality, whereas attribute 2 is the friendliness and attentiveness of the staff and comfort of the amenities. Standard quality in the attribute medical quality could then be interpreted as the cure of a health problem with a certain probability of adverse side or medium term effects from the service, whereas high quality is cure of the health problem with a lower associated probability of adverse side or medium term effects from the service. We normalize consumption utility of standard quality in both attribute to zero ($u_1(q_1 = l) = u_2(q_2 = l) = 0$) and high quality in the second

[10] Thus, patients are homogeneous in their valuation of the health care service. We will discuss heterogeneous patients in Section 3.8. Note that $U(q)$ can be interpreted as an expected utility level patients face once q is realized. This reflects a setting where providers with quality level q but might not serve constantly q but quality levels varying around q with expectation q.

attribute to 1 ($u_2(q_2 = h) = 1$).[11] This implies $u_1(q_1 = h) = \theta > 1$. Each patient's utility from abstaining from utilizing the service is $\underline{u} < 0$. The fee P for utilizing the health care service is paid for by a patient's health insurance such that $u(q)$ gives the net utility of consuming the health service for the patient.[12]

Patients cannot perfectly observe the quality levels q^A and q^B of provider A and B respectively. They however receive signals about realized quality in the attributes from each provider before deciding on a provider. Each patient receives signals $s^j = (s_1^j, s_2^j) \in \{ll, lh, hl, hh\}$, $j \in \{A, B\}$. Attribute signals s_i^j are generated with error ϵ_i with $\epsilon_i = \mathbb{P}(s_i = h \mid q_i = l) = \mathbb{P}(s_i = l \mid q_i = h) < \frac{1}{2}$, we write $\epsilon = (\epsilon_1, \epsilon_2)$. For better readability we write s^j for the signal a patient c receives instead of s_c^j. We furthermore might use $s = s^j$ as long as it is clear from the context. We do not impose any assumptions on the correlation of signals across patients, i.e. we allow signals to be independently distributed as well as to be correlated.[13] Note that we do not model aggregation of signals across patients. One interpretation of the set-up could however be that there is aggregation, e.g. via a feedback platform, and through the aggregation all patients receive a signal in attribute i with error ϵ_i as above. The notion that the signal precisions differ across attributes could then be driven by the fact that, regarding medical quality, there are only few reports about actual medical quality being published, whereas aggregation of patient feedback about amenities, staff and perceived quality leads to a more precise overall signal for these other attributes.

In our basic model, patients do not observe the providers' resource allocation decisions. To evaluate signals from providers, each patient has beliefs $b^j \in \{0, 1\}$ about the resource allocation a^j, $j \in \{A, B\}$. Again, we omit c as an index for each patient. Given any belief, patients update their belief about the quality of the service from providers according to Bayes' rule. We denote the expected utility that a patient faces at provider j when she has belief b^j about the provider's

[11] With this normalization we do not loose any generality since for our analysis we will always compare two expected utility levels such that only the size of θ will play a role for the provider selection of the patients and net welfare effects.

[12] Health insurers here are exogenous to contracting. Alternatively, instead of a health insurer paying the fee we could assume that the utility of not utilizing the health service is sufficiently low.

[13] It therefore includes the case that all patients receive the same signals. This shows that with the current set-up, we could also write the model as a representative patient that receives signals generated as above instead of a continuum of patients. We choose the continuum for the discussion of heterogeneous patients in Section 3.8.

resource allocation and receives signal $s^j = (s_1^j, s_2^j)$ by $U_s[s^j|b^j, \epsilon]$.[14] When receiving signal s^A from provider A and signal s^B from provider B a patient then chooses provider A if

$$U_s[s^A|b^A, \epsilon] > U_s[s^B|b^B, \epsilon]$$

Ties are broken equally. For ϵ fixed we write $(s|b) \succ (s'|b')$ if $U[s|b, \epsilon] > U[s'|b', \epsilon]$, i.e. when observing signal s with underlying belief b a patient faces a higher expected utility than when observing signal s' with underlying belief b'.

To summarize, the timing of the game is as follows:

Stage 1: Provider A and provider B simultaneously decide on their resource allocation a^A and a^B, respectively. Patients do not observe resource allocations.

Stage 2: For each provider the quality level in both attributes is realized.

Stage 3: Each patient receives identically distributed attribute signals $s_i^j \in \{h, l\}$ on q_i^j for all $i \in \{1, 2\}$ and $j \in \{A, B\}$ on realized quality.

Stage 4: Each patient chooses a provider.

Stage 5: Patient utility from utilizing the health service is realized.

Given the set-up, maximizing profits for providers corresponds to maximizing the probability of being selected as provider. In the following, we analyze perfect Bayesian Equilibria (PBE) in pure strategies and discuss potential mixing strategies in Section 3.8. We require patient beliefs to be consistent with the providers' resource allocations in equilibrium.

3.4 Focusing on Attributes

A patient receives two signals s, one from each provider. Which provider will the patient choose? Assume that one of the signals, say from provider A, indicates standard quality in the first and high quality in the second attribute, i.e. $s^A = lh$. The signal from provider B indicates high quality in the first and standard quality in the second attribute, i.e. $s^B = hl$. Whether the signal of high quality in the first or in the second attribute is decisive for the patient's provider choice now does not only depend on θ, the relative ex-post importance of high quality

[14]In this formulation, the belief does not have to be correct. However, in equilibrium we require beliefs to be consistent with actions.

in attribute 1, but also on the relative attribute signal precisions, for any given beliefs and technology parameter p. Thus, it might well be the case that if signals are hl for provider B and lh for provider A, the patient chooses provider A. This particularly implies that she picks provider A whenever provider A's signal indicates high quality in the second attribute and provider B's signal indicates low quality in the second attribute. Then, the signal of high quality in attribute 2 drives patient choice and we say that the patient focuses on attribute 2. This is generalized and formalized in the following definition of focusing.

Definition 3.1 (Focusing on Attributes). *Fix ϵ, p and θ. A patient...*

(i) ...focuses *on attribute i if for any two signals $s^j = (s_1^j, s_2^j)$ and $s^k = (s_1^k, s_2^k)$ with $s_i^j = h$ and $s_i^k = l$ and symmetric beliefs $b^j = b^k \in \{0,1\}$ signal s^j yields higher expected utility, i.e. $(s^j|b^j) \succ (s^k|b^k)$ for all $b^k = b^j \in \{0,1\}$.*

(ii) ...strongly focuses *on attribute i if for any two signals $s^j = (s_1^j, s_2^j)$ and $s^k = (s_1^k, s_2^k)$ with $s_i^j = h$ and $s_i^k = l$ and any beliefs $b^j, b^k \in \{0,1\}$ signal s^j yields higher expected utility, i.e. $(s^j|b^j) \succ (s^k|b^k)$ for all $b^k, b^j \in \{0,1\}$.*

Since $(hh|b) \succ (s|b)$ for all $s \neq hh$ and $(s|b) \succ (ll|b)$ for all $s \neq ll$, the definition implies that focusing on attribute 1 is equivalent to $(hl|b) \succ (lh|b)$ for all beliefs b and focusing on attribute 2 is equivalent to $(lh|b) \succ (hl|b)$ for all beliefs b. It analogously holds with any beliefs b and b' for strong focusing.

Note that for any given p and θ, whether patients that maximize their expected utility focus on an attribute or not only depends on the signal technology. This is because the requirements have to hold for all potential (symmetric) beliefs. In particular, the definition of focusing is not linked to equilibrium beliefs. Patient focusing is thus a direct property of the signal technology and not of equilibrium behavior.[15]

Focusing behavior as defined above is rational in the sense that patients maximize their expected utility given beliefs and update according to Bayes' rule. Thus, focusing here is different from focusing or salience in the behavioral economics literature (Bordalo et al., 2013, Koszegi and Szeidl, 2013) where there is an exogenous wedge between decision utility and consumption utility. Inefficiency will occur in our model via demand focusing that is nevertheless perfectly rational. Note that the focusing definition could however easily be adjusted to incorporate

[15] If patients were able to observe a^j we could replace b^j and b^k by a^j and a^k in the definition of focusing. Again, focusing does not depend on the equilibrium action.

other, potentially non-rational decision rules where patients update differently or do not maximize expected utility. The focusing definition can also naturally be applied in more general product market settings.

Focusing on attributes depends on the signal error $\epsilon = (\epsilon_1, \epsilon_2)$, the investment technology p and the utility weight θ of attribute 1. Intuitively, the smaller the signal error in one attribute keeping the signal precision in the other attribute fixed, the more informative the signals are in this attribute and the more likely it is that there is focusing on this attribute. The utility factor $\theta > 1$ implies that high quality provided in attribute 1 is more important than high quality provided in attribute 2. Hence, if signal precision in attribute 1 is not lower than in attribute 2, patients focus on attribute 1. However, conversely, if signal precision in attribute 2 is higher than in attribute 1, patients might focus on attribute 2 if θ is small enough. Generally, we can divide the attribute signal error space into focusing areas for given p and θ. The following lemma describes the separating lines for the focusing areas.

Lemma 3.1. *Fix p and $\theta > 1$. Then there exist continuous and increasing functions $f^{s1} \leq f^{12} \leq f^{s2}$ with $f^i : [0, \frac{1}{2}] \to [0, \frac{1}{2}]$, $i \in \{s1, 12, s2\}$, that divide the signal error space $[0, \frac{1}{2}]^2$ into focusing areas. A patient...*

- ...strongly focuses on attribute 2 *iff* $\epsilon_1 > f^{s2}(\epsilon_2)$. *There is $\epsilon_2^* < \frac{1}{2}$ such that f^{s2} strictly increases on $[0, \epsilon_2^*]$ and $f^{s2}(\epsilon_2) = \frac{1}{2}$ for all $\epsilon_2 \geq \epsilon_2^*$. $\epsilon_2^* > 0$ iff $\theta < \frac{1}{1-2p}$.*

- ...focuses on attribute 2 *iff* $\epsilon_1 > f^{12}(\epsilon_2)$ *and* focuses on attribute 1 *iff* $\epsilon_1 < f^{12}(\epsilon_2)$. *$f^{12}$ strictly increases in ϵ_2. Furthermore, $0 < f^{12}(0) < f^{12}(\frac{1}{2}) = \frac{1}{2}$.*

- ...strongly focuses on attribute 1 *iff* $\epsilon_1 < f^{s1}$. *f^{s1} strictly increases in ϵ_2 and $0 < f^{s1}(0) < p < f^{s1}(\frac{1}{2}) < \frac{1}{2}$.*

For $\theta \to 1$ all functions converge to the 45-degree-line . For $\theta \to \infty$ the separating line of strong focusing on attribute 1 converges to p and all other functions converge to $\frac{1}{2}$.

Proof. See appendix. □

Figure 3.1 illustrates the separating lines for $p = 0.25$ and $\theta = 2$. Figure 3.2 illustrates the separating lines for again $p = 0.25$ but $\theta = 1.4$.

The two figures visualize how the focusing areas change when θ is varied. $\theta > 1$ implies that the area of focusing on attribute 1 is larger than the area of focusing

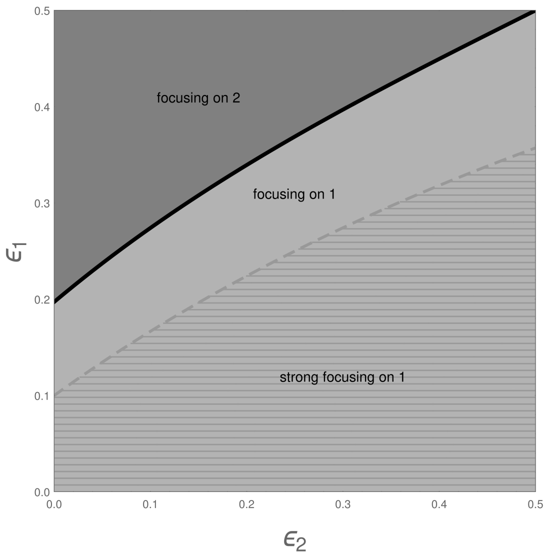

Figure 3.1: $p = 0.25$ and $\theta = 2$

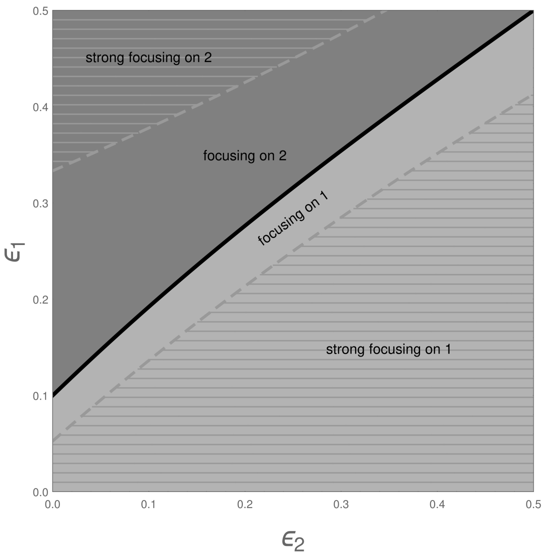

Figure 3.2: $p = 0.25$ and $\theta = 1.4$

on attribute 2. For large θ ($\theta > \frac{1}{1-2p}$, which is the case in figure 3.1), attribute 1 is important enough such that the area of strong focusing on attribute 2 vanishes completely. An area of focusing on attribute 2 exists independent of the magnitude of θ. However, this area becomes arbitrarily small for θ converging to infinity. For $\theta \to 1$, all separating lines converge to the 45-degree-line.

The lemma shows that for a fixed error in one attribute, lowering the error in the other attribute makes the signals in this attribute more important and might shift the focus of a patient towards this attribute. For any $\theta > 1$ and p we can choose ϵ_1 large enough such that lowering ϵ_2 results in a shift from focusing on attribute 1 to focusing on attribute 2. For the equilibrium and welfare analysis, we will be also interested in the conditions under which there is a shift from strong focusing on attribute 1 to focusing on attribute 2 when lowering ϵ_2. Graphically, this translates to finding a horizontal line such that this line crosses both the area of strong focusing on 1 and the area of focusing on 2. In our examples, for instance, this is the case for $\epsilon_1 = 0.25$. The following corollary provides a sufficient condition on θ to find such an ϵ_1.

Corollary 3.1. *Fix p and $\theta > 1$. There exist errors ϵ_1 such that by varying ϵ_2 the patients' focus shifts from focusing on attribute 1 to focusing on attribute 2.*

For $\theta < \overline{\theta} = \frac{1}{1-2p}$ there exist errors ϵ_1 such that by varying ϵ_2 the patients' focus shifts from strong focusing on attribute 1 to focusing on attribute 2.

Proof. See appendix. □

Particularly, by the monotonicity of the separating lines, for $\epsilon = (\epsilon_1, \epsilon_2)$ with ϵ_1 large enough, patients (strongly) focus on attribute 1 for large ϵ_2 and focus on attribute 2 for small ϵ_2. For θ close enough to 1 it is even possible to find ϵ_1 such that lowering ϵ_2 results in a shift from strong focusing on attribute 1 to strong focusing on attribute 2. However, the weaker conditions presented in the corollary will be sufficient for our further analysis.

3.5 Provider Quality Incentives and Equilibria

On the basis of the patients' focusing behavior we can analyze the providers' incentives to allocate their resources between attributes. We say that a strategy a^j of a provider j is dominant if for any patients' beliefs (b^A, b^B) and any strategy a^{-j} of the other provider, the strategy a^j is weakly better than any other strategy and strictly better for at least one combination of beliefs and the other provider's strategy. We call a^j strictly dominant if it is strictly better for all combinations of patients' beliefs (b^A, b^B) and the other provider's strategy a^B.

In the following we show that once patients focus on an attribute and the signal error in this attribute is lower than the signal error in the other attribute, it is a dominant strategy for a provider to invest in this attribute. If focusing is strong, it is even a strictly dominant strategy to invest in the respective attribute.

Proposition 3.1. *Let θ, p and $\epsilon = (\epsilon_1, \epsilon_2)$ be such that patients...*

(i) ....(strongly) focus on attribute 2. Then it is a (strictly) dominant strategy for any provider j to invest in attribute 2, i.e. $a^j = 0$.

(ii) ...(strongly) focus on attribute 1 and $\epsilon_1 < \epsilon_2$. Then it is a (strictly) dominant strategy for any provider j to invest in attribute 1, i.e. $a^j = 1$.

Proof. See appendix. □

The main idea of the proof is that for fixed beliefs of patients the resource allocation of the provider does not influence the expected utility of any patient when receiving a specific signal. This is because patients cannot observe the investment but perform the Bayesian updating when receiving the signal based on their belief. What changes when the provider selects a different investment strategy are the probabilities with which the signals are generated. If patients focus on one attribute and the signal error in this attribute is lower than in the other attribute, investing in this attribute generates "better" signals with higher probability than any other strategy. While focusing on attribute 2 already implies $\epsilon_2 < \epsilon_1$, we have to additionally condition on $\epsilon_1 < \epsilon_2$ when considering focusing on attribute 1.

One might wonder what optimal strategies are in case that there is focusing on attribute 1 but signal errors are such that $\epsilon_1 > \epsilon_2$. Focusing implies that for any fixed beliefs, hl yields higher expected utility than lh. However, investing in attribute 1 instead of investing in attribute 2 does not unambiguously produce better signals with higher probability as it is the case for $\epsilon_1 < \epsilon_2$ such that optimal provider strategies then depend on signals errors in more detail.[16]

The proposition implies that for strong focusing on attribute 2 it is a strictly dominant strategy for the providers to invest in attribute 2, i.e. it is strictly better for any strategy of the other provider and any combination of patients' beliefs. However, if focusing is not strong, providers might be indifferent between different resource allocations. This crucially depends on the beliefs of patients. For symmetric beliefs about the providers' resource allocations it is strictly better for the providers to invest in attribute 2 when patients focus on attribute 2. However, if patients have asymmetric beliefs, selection of the provider might be based only on the beliefs, ignoring the signals. Then providers are indifferent between different resource allocations. This might occur if patients believe that provider A invested in attribute 1 and provider B in attribute 2 and the parameters are such that patients choose provider A independent of the signals. For instance, $\epsilon = (\epsilon_1, \epsilon_2) = (\frac{1}{2}, 0)$ and $\theta > \frac{1}{1-2p}$ satisfy $(ll|b^A = 1) \succ (hh|b^B = 0)$ from which follows that patients ignore the signals and always select provider A anyway.

Proposition 3.1 directly implies that if patients focus on one attribute and the signal error in this attribute is lower than in the other attribute, investing in this attribute and corresponding beliefs is a Perfect Bayesian Equilibrium. Strong focusing (and $\epsilon_1 < \epsilon_2$ for focusing on attribute 1) implies uniqueness of the respective symmetric equilibrium. However, if focusing is not strong further equilibria might exist. Proposition 3.2 shows that the only further equilibria that might exist are asymmetric equilibria in which patients select the provider solely based on the beliefs and signals are irrelevant.

Proposition 3.2. *Let θ, p and $\epsilon = (\epsilon_1, \epsilon_2)$ be such that patients...*

[16] If a provider invests in attribute 1 instead of 2, on the positive side, signal hl is produced with a higher probability on the cost of signal lh. On the negative side, signal ll is produced with a higher probability on the cost of signal hh. The closer (ϵ_1, ϵ_2) to the 45-degree line, the large the positive and the smaller the negative effect is, since the difference in expected utilities of hl and lh increases and the differences in probabilities of producing hh compared to ll decreases.

(i) *...focus on attribute* 2*. Then* $(a^A, a^B) = (b^A, b^B) = (0, 0)$ *is a PBE. Any PBE with* $(a^A, a^B) = (b^A, b^B) \neq (0, 0)$ *is asymmetric, i.e.* $a^A \neq a^B$ *and patients select provider A if and only if* $a^A = 1$*. Strong focusing on attribute* 2 *implies that the symmetric PBE* $(a^A, a^B) = (b^A, b^B) = (0, 0)$ *is unique. Equilibrium is furthermore unique if, for a given* ϵ_i*, setting* $\epsilon_{-i} = \frac{1}{2}$ *implies strong focusing on attribute* i*, i.e. either patients strongly focus on attribute* 1 *once the signal in attribute* 2 *is uninformative or strongly focus on attribute* 2 *once the signal in attribute* 1 *is uninformative.*

(ii) *... focus on attribute* 1 *and* $\epsilon_1 < \epsilon_2$*. Then* $(a^A, a^B) = (b^A, b^B) = (1, 1)$ *is a PBE. Any PBE with* $(a^A, a^B) = (b^A, b^B) \neq (1, 1)$ *is asymmetric, i.e.* $a^A \neq a^B$ *and patients select provider A if and only if* $a^A = 1$*. Strong focusing on* 1 *implies uniqueness of the symmetric PBE* $(a^A, a^B) = (b^A, b^B) = (1, 1)$*.*

Proof. See appendix. □

For focusing on attribute 2, Proposition 3.2 shows that the equilibrium is not only unique under strong focusing, but also for signal errors that are such that there would be strong focusing on one attribute if the error for the other attribute would be set to $\frac{1}{2}$, i.e. if patients were not to receive an informative signal in this attribute. This is because, if an asymmetric equilibrium exists, with consistent beliefs signal ll from the provider with higher a is preferred to signal hh from the other provider. This continues to hold when e.g. increasing ϵ_2. Then, however, there is a contradiction with strong focusing, where hh is preferred to ll for any symmetric or asymmetric beliefs. The intuition for ϵ_1 is the same.

For the cases where multiple equilibria exist, note that only the symmetric equilibrium where both providers invest in the attribute that patients focus on is an equilibrium in dominant strategies of the providers. Therefore, it is robust with respect to perturbation in the patients' beliefs as the optimal strategy is independent of the beliefs. Furthermore, it is the only equilibrium where signals are informative for the patients such that they matter for their provider choice. Both reasonings might serve as a selection criterion for concentrating on symmetric equilibria.

Corollary 3.2. *Fix* θ *and* p *and consider* $\epsilon = (\epsilon_1, \epsilon_2)$ *such that patients focus on attribute* i *and* $\epsilon_i < \epsilon_{-i}$*. Then the symmetric equilibrium where both providers invest in attribute* i *is the only equilibrium in dominant strategies. It is furthermore the only equilibrium where signals are informative for patients.*

3.6 Welfare and Comparative Statics

We can now discuss the welfare consequences of the patients' focusing on attributes. Note that in the model, total provider surplus is fixed. For the welfare analysis, we will not consider the distribution of producer surplus between providers and henceforth concentrate on patient welfare. Thus, we will use the term welfare synonymous to patient welfare.

Now assume that quality (q^A, q^B) is realized for provider A and B (and is unknown by the patients). We denote by $U_q[(q^A, q^B)|(b^A, b^B), \epsilon]$ the expected utility of quality provision of a patient when quality (q^A, q^B) is realized and the patient, under beliefs (b^A, b^B), chooses providers to maximize her expected utility given signals when signals are generated with errors $\epsilon = (\epsilon_1, \epsilon_2)$. Denote by $W[(a^A, a^B)|(b^A, b^B), (\epsilon_1, \epsilon_2)]$ welfare if providers' resource allocations are $a = (a^A, a^B)$, patients have beliefs $b = (b^A, b^B)$, receive quality signals with error $\epsilon = (\epsilon_1, \epsilon_2)$ and choose providers maximing expected utility given signals and beliefs. Then

$$W[(a^A, a^B)|(b^A, b^B), (\epsilon_1, \epsilon_2)] = \sum_{q^B} \sum_{q^A} \mathbb{P}(q^A|a^A)\mathbb{P}(q^B|a^B)U_q[(q^A, q^B)|(b^A, b^B), \epsilon] \tag{3.1}$$

where $\mathbb{P}(q^j|a^j)$ is the probability that q^j is realized for resource allocation a^j.[17] There are two key drivers of welfare in the market: Firstly, a pure quality aspect, i.e. the expected consumption utility without considering signals, which is determined by the resource allocations. Secondly, a provider selection effect, i.e. selecting the provider whose quality realizations are high, which works through signal precision. This last one is important when considering the welfare effect of changes in signal precision, where a lower error c.p. improves selection based on true underlying quality. Before analyzing changes in the precision of the signals, we first look at welfare for a given signal precision.

Lemma 3.2. *Fix p and θ. For all $\epsilon = (\epsilon_1, \epsilon_2)$ investing in attribute 1 and corresponding beliefs yields higher welfare than investing in attribute 2 and corresponding beliefs, i.e.*

$$W[(1,1)|(1,1), (\epsilon_1, \epsilon_2)] > W[(0,0)|(0,0), (\epsilon_1, \epsilon_2)].$$

[17] Note that our welfare definition directly incorporates optimal demand side behavior given beliefs. We could of course define $U_q[(q^A, q^B)|(b^A, b^B), \epsilon]$ based on patients' actions more generally. We write welfare in this way to concentrate the analysis on the welfare effect of different provider resource allocations and patients beliefs. Note that in the welfare definition above, patients beliefs do not yet have to be correct, they only have to be correct when comparing welfare in equilibrium.

Proof. See appendix. □

Thus, independent of ϵ, if both providers invest in 1 (and patients have corresponding beliefs), welfare is higher than if both provider invest in 2 (and patients have corresponding beliefs). For $\epsilon_1 \leq \epsilon_2$, this is intuitive. For $\epsilon_1 > \epsilon_2$, there are some opposing effects. While, by investing in 1, providers increase the probability of quality $q^j = hl$ at the cost of $q^j = lh$ where hl yields higher utility than lh, for high ϵ_1 and low ϵ_2 patients can barely infer information about quality realization in attribute 1 from signals while they reasonably can for attribute 2. In aggregation, however, the quality effect dominates the signal precision effect and welfare is higher when providers invest in attribute 1.

We already know that if ϵ is such that patients strongly focus on attribute 2, in the unique PBE both providers invest in attribute 2 with corresponding patients beliefs. Thus, when patients strongly focus on attribute 2, the unique PBE is inefficient. Under focusing on attribute 2, from Proposition 3.2 any equilibrium that is not the equilibrium in which both providers choose $a = 0$ is asymmetric and provider j is chosen if and only if $a^j > a^{-j}$. I.e., except for the symmetric equilibrium with investment in attribute 2, in equilibrium a provider is chosen with probability 1, independently of the signals that the patients receive. Then, welfare in these equilibria is again lower compared to the situation where both providers invest in attribute 1 and patients hold the corresponding belief, as quality provision is partly inefficient, and there is no selection based on signals. This is summarized in Proposition 3.3 below.

Proposition 3.3. *Fix p and θ. If ϵ is such that patients focus on attribute 2, any PBE is inefficient.*

Proof. See appendix. □

The interesting question is whether increasing signal precision increases welfare. For ϵ_1 large enough we saw that by increasing the precision in the second attribute we might move from an equilibrium where both provider invest in attribute 1 to an equilibrium where both invest in attribute 2. From above, the latter is inefficient. The welfare effect when increasing signal precision is however not obvious as there are two effects. On the one hand, increasing signal precision might lead to a "worse" provision of quality. On the other hand, patients can better select the providers with high quality realizations. In the following we show that there exist parameter ranges such that increasing signal precision in attribute 2 for given

ϵ_1 unambiguously leads to a reduction in welfare if it induces a shift from both providers investing in attribute 1 to both providers investing in attribute 2.

Proposition 3.4. *Fix $\theta > \underline{\theta} = \frac{1-p-p^2}{1-2p}$, p and ϵ_1. Consider any ϵ_2 and ϵ_2' such that patients focus on attribute 2 for $\epsilon = (\epsilon_1, \epsilon_2')$. Then the following holds*

$$W[(0,0)|(0,0),(\epsilon_1,\epsilon_2')] < W[(1,1)|(1,1),(\epsilon_1,\epsilon_2)]$$

Proof. See appendix. □

There is a lower bound on θ which ensures that, even for a maximal improvement in welfare from increasing signal precision – which would be the case for a change from $\epsilon_2 = \frac{1}{2}$ to $\epsilon_2 = 0$ –, the effect of reducing expected quality in attribute 1 with the shift in investment dominates. Proposition 3.4 implies in particular that if a change in ϵ_2 causes a shift from an equilibrium where both providers invest in attribute 1 to an equilibrium where both providers invest in attribute 2, there is an unambiguous welfare loss. This is made precise in the following corollary.

Corollary 3.3. *Fix p and $\theta > \underline{\theta}$. Consider $\epsilon = (\epsilon_1, \epsilon_2)$ with $\epsilon_1 < \epsilon_2$ and $\epsilon' = (\epsilon_1, \epsilon_2')$ such that for ϵ patients focus on attribute 1 and for ϵ' patients focus on attribute 2. Then an increase in the signal precision of attribute 2 from ϵ_2 to ϵ_2' results in a welfare loss in the respective dominant strategy equilibrium.*

If, furthermore, $\underline{\theta} < \theta < \overline{\theta}$, consider $\epsilon = (\epsilon_1, \epsilon_2)$ with $\epsilon_1 < \epsilon_2$ and $\epsilon' = (\epsilon_1, \epsilon_2')$ such that for ϵ patients strongly focus on attribute 1 and for ϵ' patients focus on attribute 2. Then an increase in the signal precision of attribute 2 from ϵ_2 to ϵ_2' results in a welfare loss in equilibrium.

For $\epsilon = (\epsilon_1, \epsilon_2)$ and $\epsilon' = (\epsilon_1, \epsilon_2')$ such that patients focus on attribute 1 for $\epsilon = (\epsilon_1, \epsilon_2)$ and on attribute 2 for $\epsilon' = (\epsilon_1, \epsilon_2')$ multiple equilibria might exist. Therefore it is a priori not clear which equilibria are selected and thus whether a reduction in welfare occurs when lowering ϵ_2 to ϵ_2'. However, as discussed the symmetric equilibrium stands out as it is the only equilibrium in dominant strategies and robust with respect to perturbations in the beliefs. When only concentrating on equilibria in dominant strategies, for any $\theta > \underline{\theta}$ the welfare loss occurs when lowering ϵ_2 such that it induces a shift from focusing on attribute 1 to focusing on attribute 2.

From Corollary 3.1 we know that for $\theta < \overline{\theta} = \frac{1}{1-2p}$ there exist ϵ_1 such that for $\epsilon = (\epsilon_1, \frac{1}{2})$ patients strongly focus on attribute 1 and for $\epsilon = (\epsilon_1, 0)$ patients focus on attribute 2. Thus, there exists ϵ and ϵ' as described above. Furthermore, Corollary 2 and Proposition 3.2 showed that in this case the equilibria are unique. $\theta > \underline{\theta}$ ensures that there is a welfare loss.

3.7 Quality Reporting

So far we assumed that patients receive informative signals from each provider for all attributes. However, it might be a strategic choice of providers to send quality signals in attributes, e.g. via participation in evaluations and quality reporting, or establishment of an online feedback platform. From a policy perspective, it is important to understand which reporting policies induce optimal outcomes. When is it necessary to require providers to undertake quality reporting in certain attributes or ban reporting in others? In the following we first discuss strategic reporting of providers. We then analyze different reporting policies and compare them to strategic reporting by providers.

Strategic reporting. To incorporate strategic quality reporting by providers, we change the game in the following way: Whether patients receive signals about attribute quality now depends on a reporting decision by providers. Each provider can decide at the time of resource allocation for each attribute whether to send signals about quality or not.[18] We assume that a provider, when deciding about reporting, again cannot influence the precision of the signals. I.e., when reporting in attribute 1, the provider sends a signal about this attribute with error ϵ_1 and when reporting in attribute 2 he sends a signal about this attribute with error ϵ_2. The reason that he cannot influence the signal precision is again the general difficulty in observing, measuring and communicating quality in certain attributes. In terms of hospital quality, think of an external report or a platform where patients rate experienced quality in a hospital. While medical quality is rather difficult to evaluate, non-medical quality attributes are fairly easy to rate. Note that *not* reporting in attribute i is equivalent to a signal error of $\frac{1}{2}$ in attribute i.

Providers simultaneously decide on their resource allocation a and their reporting r, i.e. in which attributes they want to report signals. Patients now might not receive signals in some attribute, but they update their beliefs about resource allocations depending on whether they receive signals in attributes. To keep the game simple, we exploit Section 3.5's results and restrict attention to

[18]Crucial here is that providers do not know their quality at the time of deciding whether to take part in reporting. Thus, reporting is **not** signaling on realized quality.

strategies[19]

$$(a, r) \in \{(1, s_1), (0, s_2), (x(\epsilon), s_1 s_2), (1, none)\},$$

where s_1 (s_2) stands for reporting only on attribute 1 (2) and $s_1 s_2$ for reporting in both. Furthermore, $x(\epsilon) \in \{0, 1\}$ with $x(\epsilon) = 0$ if ϵ is such that patient focusing is on attribute 2 and $x(\epsilon) = 1$ if $\epsilon_1 < \epsilon_2$ (and therefore patient focusing is on attribute 1 when they receive signals in both attributes). Thus, we consider the cases that (i) no signals are sent (no reporting) and providers invest in attribute 1, (ii) a provider sends the signal in the attribute that he invested in, but not in the other attribute (partial reporting), and (iii) signals in both attributes are sent, and investments are in the attribute that patients focus on when receiving signals in both attributes, given ϵ (full reporting). Again we concentrate on pure strategy equilibria.

How do providers strategically report and invest? Assume that ϵ is such that if signals are sent in both attributes, there is focusing on 2. Now consider the situation that both providers report in both attributes and invest in attribute 2. Then, each provider is selected with probability $\frac{1}{2}$. Now assume a provider changes his reporting to only reporting in attribute 2, and not reporting in attribute 1. Then, this provider is selected with probability higher than $\frac{1}{2}$ when playing against the provider who is reporting in both attributes. This is because, since investments are in attribute 2, the provider reporting in both attributes sends a low quality signal in attribute 1 with probability higher than $\frac{1}{2}$, and since the signal is informative, in these cases the provider not reporting in attribute 1 is selected when the signal in the other attribute is the same. Thus, not reporting in the 'weak' attribute is a profitable deviation. This logic can be generalized to show that there are no equilibria with reporting in both attributes.

[19]Thereby we ensure the exclusion of implausible equilibria. For any combination of patient beliefs when the strategy space is not restricted, i.e. for any combination of reporting and resource allocation, any of the excluded strategies would be weakly dominated. For this note that we know from the results in Section 3.5 that receiving a signal only in one attribute i implies that investing in attribute i weakly dominates investing in the other attribute (keeping the signal structure constant). With restricting strategies, we can restrict patient beliefs accordingly and can thereby rule out implausible equilibria where dominated strategies are selected by the providers. If no signals are reported, a provider's action has no influence on any information the patient receive. In this case we assume that providers invest in 1 to avoid a point of discontinuity when considering receiving no signal in attribute 2 and facing signal errors ϵ_1 that are close to $\frac{1}{2}$.

Lemma 3.3. *Fix p and θ and consider ϵ such that patients either focus on attribute 2 or they focus on attribute 1 and $\epsilon_1 < \epsilon_2$. Then, an equilibrium in which both providers report in both attributes does not exist.*

Proof. See appendix. □

To determine equilibria, a crucial consideration is how patients choose providers when one provider sends only a signal in attribute 1 (and invests in attribute 1) and the other provider sends a signal only in attribute 2 (and invests in attribute 2). Although patients do not observe resource allocations directly, they can update their beliefs when receiving, respectively not receiving, signals. Then, if $(h \cdot |1) \succ (\cdot h|0)$ (i.e. a signal of high quality in attribute 1 and no signal in attribute 2 under belief 1 yields higher expected utility than a high quality signal in attribute 2 and no signal in attribute 1 under belief 0), the provider only sending a signal in attribute 1 is selected with probability greater than $\frac{1}{2}$. Then both providers sending a signal only in attribute 2 (with investing in 2) cannot be an equilibrium, as sending a signal only in attribute 1 is a profitable deviation. It is straightforward to show that

$$(h \cdot |1) \succ (\cdot h|0) \;\forall\, \epsilon \;\Leftrightarrow\; \theta > \theta^c = \frac{1-p}{1-2p}.$$

This particularly also says that if $\theta < \theta^c$ there exist ϵ, e.g. $\epsilon = (\frac{1}{2}, 0)$ and some neighborhood, such that $(\cdot h|0) \succ (h \cdot |1)$. Note that $\underline{\theta} < \theta^c < \overline{\theta}$ with $\underline{\theta}$ and $\overline{\theta}$ as defined in the previous sections. We can now describe equilibria under strategic reporting.

Proposition 3.5. *(i) Fix p and θ. For any ϵ such that $\epsilon_1 < \epsilon_2$ (and therefore patients focus on attribute 1), in the unique PBE providers invest in attribute 1 and report only on attribute 1.*

(ii) Fix p and $\theta > \theta^c$. Then there exist errors ϵ such that patients focus on attribute 2 when receiving signals in both attributes, however in the unique PBE providers invest in attribute 1 and report only on attribute 1.

(iii) Fix p and $\theta < \theta^c$. Then there exist errors ϵ such that in the unique PBE providers invest in attribute 2 and report only on attribute 2.

Proof. See appendix. □

Proposition 3.5 states that, under strategic reporting, there exist equilibria in which providers invest in an attribute and only publish quality signals in that respective attribute. Thus, it might be the case that not only resource allocation, but also information provision is inefficient. However, as the second part of Proposition 3.5 shows, if $\theta > \theta^c$, strategic reporting might even result in providers voluntarily withholding information in attribute 2 and investing in attribute 1, although ϵ is such that there would be focusing on 2.

To get an intuition for parts (ii) and (iii) of the proof, consider the extreme case of $\epsilon_1 = \frac{1}{2}$, e.g. there is no signal in attribute 1, and $\epsilon_2 = 0$, e.g. signals in attribute 2 are precise. Focusing on 2 when receiving both signals is therefore satisfied as for symmetric beliefs it always yields higher expected utility when receiving signal h in attribute 2 than signal l. Since $\epsilon_1 = 0$ only two strategies are relevant: reporting about attribute 2 or not. It is a strictly dominant strategy for a provider to withhold information about attribute 2 if and only if the expected utility for the patient is higher if resources are concentrated on attribute 1 but she receives no signal about the realization, i.e. $(1-p)\theta + p$, than if resources are concentrated on attribute 2 and she receives an exact signal about the realization in attribute 2, i.e. $p\theta + 1$. This holds if and only if $\theta > \theta^c = \frac{1-p}{1-2p}$. The proof in the appendix elaborates some more general conditions on ϵ for which the claims hold. Particularly, it shows that claim (ii) is not only satisfied in a neighborhood of $\epsilon = (\frac{1}{2}, 0)$ but also once $p > \frac{1}{3}$ and ϵ is such that patients focus on attribute 2 and $\epsilon_1 > p$. For claim (iii) it is crucial that ϵ is such that $(\cdot h|0) \succ (h \cdot |1)$.

Comparison of Reporting Policies. Since *not* reporting in attribute i is equivalent to a signal error of $\frac{1}{2}$ in attribute i, we can use of the previous sections to determine the welfare of potential outcomes with reporting and thus optimal outcomes.

Recall that $W[a|b, \epsilon]$ denotes expected (patient) welfare if providers' resource allocations are $a = (a^A, a^B)$, patients have belief $b = (b^A, b^B)$ and receive quality signals with errors $\epsilon = (\epsilon_1, \epsilon_2)$. Keeping the resource allocation constant and only improving signal precision by sending a signal, we have, by the simple selection effect, for any errors (ϵ_1, ϵ_2),

$$W[(1,1)|(1,1),(\epsilon_1,\epsilon_2)] > W[(1,1)|(1,1),(\epsilon_1,\frac{1}{2})],$$
$$W[(0,0)|(0,0),(\epsilon_1,\epsilon_2)] > W[(0,0)|(0,0),(\frac{1}{2},\epsilon_2)].$$

Furthermore, it holds that

$$W[(1,1)|(1,1),(\epsilon_1,\epsilon_2)] > W[(0,0)|(0,0)|(\frac{1}{2},\epsilon_2)],$$

since here the selection and resource allocation effect go in the same direction. For a selection and resource allocation effect going in opposite directions we know from Proposition 3.4, that if $\theta > \underline{\theta}$ and $\epsilon = (\epsilon_1, \epsilon_2)$ is such that patients focus on attribute 2, signal provision only in attribute 1 (with investing in 1) yields higher welfare than signal provision in both attributes and investing in 2, i.e.

$$W[(1,1),(1,1),(\epsilon_1, \frac{1}{2})] > W[(0,0),(0,0),(\epsilon_1, \epsilon_2)].$$

Put together, an equilibrium where both providers invest in attribute 2 and report only in attribute 2 is welfare dominated by an equilibrium in which providers invest in attribute 2 but report in both attributes. Whether investing in attribute 1 and reporting only in 1 dominates full reporting and investing in 2 depends on θ and ϵ. From section 3.6 we know that for $\theta > \underline{\theta}$ it holds for all ϵ. However, even for $\theta < \underline{\theta}$, as long as θ is not too small, reporting only in attribute 1 and both providers allocating resources in 1 might still yield higher welfare than full reporting with investment in attribute 2. In particular, for any given ϵ, there exists $\hat{\theta}(\epsilon) \leq \underline{\theta}$ such that for $\theta > \hat{\theta}(\epsilon)$, reporting in 1 and investing in 1 is the optimal outcome, and for $\theta < \hat{\theta}(\epsilon)$, full reporting and investing in 2 is the optimal outcome.[20]

With the analysis of equilibria under strategic reporting and the welfare considerations above, we can now compare welfare of different reporting policies. A reporting policy describes for each attribute whether signal reporting is voluntary, mandatory or banned. Whenever we call a policy *mandatory reporting in attribute* i or *banning reporting in attribute* i it implies that reporting in the other attribute is a voluntary decision of the providers.

Proposition 3.6. *(i) Fix p and θ. For any ϵ such that $\epsilon_1 < \epsilon_2$ (and therefore patients focus on attribute 1), mandatory full reporting is optimal and strictly increases welfare compared to voluntary reporting in both attributes.*

(ii) Fix p and $\theta > \theta^c$. Let ϵ be such that patients focus on attribute 2 when receiving signals in both attributes, and in the unique PBE under voluntary reporting in both attributes there is reporting only in attribute 1. Then, voluntary reporting in both attributes is already optimal. Banning reporting in attribute 2 as well as mandating reporting in attribute 1 are both optimal. Any policy mandating reporting in attribute 2 is not optimal.

(iii) Fix p and $\theta < \theta^c$. Let ϵ be such that reporting only in attribute 2 is the unique PBE under strategic reporting. For $\theta > \hat{\theta}(\epsilon)$, mandating reporting in

[20]Consider any ϵ and θ. If reporting in 1 and investing in 1 and dominates full reporting and investing in 2, it does as well for any $\theta' > \theta$. If full reporting and investing in 2 dominates reporting in 1 and investing in 1 it does as well for any $\theta' < \theta$. Since we face the first case for all $\theta > \underline{\theta}$ and the second one for θ close enough to one, we can find such $\hat{\theta}(\epsilon)$.

attribute 1 *strictly increases welfare compared to voluntary reporting in both attributes but is not necessarily an optimal policy. Banning reporting in attribute* 2 *is optimal. For* $\theta < \hat{\theta}(\epsilon)$*, mandating full reporting is optimal while voluntary reporting as well as banning reporting in attribute* 2 *are not optimal. Banning reporting in attribute* 2 *might even decrease welfare compared to voluntary reporting in both attributes.*

Proof. See appendix. □

Note that, directly implied by the welfare discussion above, reporting in attribute 1 is always part of an optimal reporting policy, even if the signal precision in attribute 1 is very low. The proposition above shows that different policies might lead to optimal reporting in equilibrium. For some parameter constellations it is even not necessary to intervene with a specific policy to reach optimal reporting as providers might already voluntarily withhold information in attribute 2 when desirable from a welfare maximizing perspective. This occurs despite focussing on attribute 2 when patients receive signals in both attributes.

Proposition 3.6 also shows that mandating reporting in 1 might require at the same time to regulate reporting in attribute 2. Depending on the parameters it might be necessary to ban signals in attribute 2 or to mandate them.

From our discussion about difficulties in measuring and communicating medical quality compared to other attributes of a health care service, a particularly relevant case is the situation where the signal is imprecise on attribute 1 but fairly precise on attribute 2. To emphasize this case, we will summarize the results for high ϵ_1 and low ϵ_2 in the following corollary.

Corollary 3.4. *Fix* p *and* θ*. Then for all errors* $\epsilon = (\epsilon_1, \epsilon_2)$ *close enough to* $(\frac{1}{2}, 0)$*, patients focus on attribute* 2 *and for*
(i) $\theta > \theta^c$*, voluntary reporting, mandatory reporting in attribute* 1 *as well as banning reporting in* 2 *are optimal policies. Mandatory full reporting is not optimal.*
(ii) $\underline{\theta} < \theta < \theta^c$*, banning reporting in attribute* 2 *is necessary and sufficient for optimal reporting.*
(iii) $\theta < \underline{\theta}$*, mandatory reporting on attribute* 1 *as well as mandatory full reporting are optimal policies, whereas banning reporting on attribute* 2 *is not.*

Thus, considering signals that are very precise in attribute 2 but very imprecise in attribute 1, for high and low θ mandating reporting in 1 is already optimal. $\theta > \theta^c$ implies for errors close enough to $\epsilon = (\epsilon_1, \epsilon_2)$ providers voluntarily withhold information about attribute 2 which corresponds to optimal reporting. For $\theta < \underline{\theta}$ optimal reporting is sending both signals. However, for errors close enough to $\epsilon = (\epsilon_1, \epsilon_2)$, providers voluntarily also send information in attribute 2 when information in attribute 1 is mandated. For intermediate θ, i.e. $\underline{\theta} < \theta < \theta^c$, it is not sufficient to mandate information in attribute 1 to yield optimal reporting. In this case it is necessary to control signals in attribute 2 by banning them.

For medical services, our results then imply that if the information structure is such that medical quality signals are imprecise but signals on amenities precise it depends on how important medical care compared to the other dimensions is whether or not optimal reporting includes attribute 2. Mandating information in attribute 1 yields optimal reporting except for some intermediate θ where an additional ban in attribute 2 is necessary.

3.8 Discussion

In this section, we will discuss the consequences of relaxing several modeling assumptions as well as extensions. We will first discuss (i) symmetries in the quality realization technology and (ii) symmetries across providers before discussing how to (iii) model the technology for splitting the resources or mixing strategies and the consequences of (iv) correlation in quality realizations. Finally, we discuss (v) observability of the providers' resource allocations for patients and (vi) heterogeneity in θ.

Symmetries in quality realization. To keep the model tractable, it incorporates two symmetries about how the resource allocation impacts the quality realization, (1) a symmetric impact of the resource allocation on quality realization across attributes and that (2) quality realization probabilities are symmetrically spread around $\frac{1}{2}$. We will shortly discuss both in the following. Both symmetries arise from the assumption

$$\mathbb{P}(q_1^j = h | a^j) = (1-p)a^j + p(1-a^j) = \mathbb{P}(q_2^j = h | 1 - a^j).$$

We do not need the symmetries for our qualitative results - the symmetries rather shift thresholds but do not change the qualitative claims. In the following we explain how the symmetries can be removed and the implications of allowing for asymmetries.

Symmetric impact of resource allocation on quality realization across attributes. We assume that for any resource allocation decision a^j the probability that high quality is realized in attribute 1 equals the probability of high quality realization in attribute 2 if resources are allocated according to $1 - a^j$, i.e. $\mathbb{P}(q_1 = h|a^j) = \mathbb{P}(q_2 = h|1 - a^j)$. The parameter p can be interpreted as a measure of how effective resources in both attributes are for quality realization. Our assumption therefore reflects a symmetry across the two attribute meaning that resources have the same impact of quality realization for both attributes.
One way to give up this assumption is to consider different parameters $p_1, p_2 \in (0, \frac{1}{2})$ for the effectiveness of the resource allocation for both attributes, particularly

$$\mathbb{P}(q_1 = h|a^j) = (1 - p_1)a^j + p_1(1 - a^j)$$
$$\mathbb{P}(q_2 = h|1{-}a^j) = (1 - p_2)a^j + p_2(1 - a^j).$$

Once p_1 is smaller than p_2 resources are more effective in attribute 1 than in attribute 2 on quality realization and the other way around. This additional asymmetry does not qualitatively change our results but would only add one additional asymmetry across attributes in addition to signal errors ϵ and relevance θ to our model. Thus, $p_1 < p_2$ would additionally favor investments in attribute 1 while $p_1 > p_2$ would favor investments in attribute 2. This produces a shift in the borders of focusing as well as when investing in one attribute is a dominant strategy. The smaller the difference between p_1 and p_2, the closer we come to the presented results. However, since this source of asymmetry across attributes is not the focus of our work we do not include it into our basic model while being aware that technological asymmetries across attributes exist in applications.

Quality realization probabilities symmetrically spread around $\frac{1}{2}$. The second symmetry behind our assumption on how a^j impacts quality realization is that the probability that high (low) quality in an attribute is realized investing in i and the probability that high (low) quality is realized investing in the other attribute add up to one, i.e. $\mathbb{P}(q_i = h|a^j) = 1 - \mathbb{P}(q_i = h|1 - a^j)$. It can be interpreted as a symmetry across low and high quality realization.
This symmetry can be given up by assuming instead

$$\mathbb{P}(q_1 = h|a^j) = a^j\overline{p} + (1 - a^j)\underline{p} = \mathbb{P}(q_2 = h|1 - a^j) \quad \text{with } \underline{p} < \overline{p}$$

Here, the probability $1 - p$ of high quality realization in the attribute a provider invested in is replaced by $\overline{p}$ and the probability p of high quality realization in the attribute the provider did not invest in is replaced by $\underline{p}$. Then, resources are still

equally effective in both attributes (see discussion point above), but probabilities of high quality realization in one attribute for a^j and $1 - a^j$ are not any more symmetrically spread around $\frac{1}{2}$. Particularly, if both $\overline{p}$ and $\underline{p}$ are rather high, the probability that high quality is realized in one attribute is high independently of whether the provider invested in the attribute or not (and the probability of low quality realization is low). For $\overline{p}$ and $\underline{p}$ both being rather low, the probability that high quality is realized in one attribute is low independently of the provider's action.
In the following we argue that our qualitative results do not change but only critical values for θ or ϵ might change. For this, we first consider how the error space is divided into focusing areas (see Lemma 3.1). For any fixed $(\underline{p}, \overline{p})$ and θ, we again can describe separating lines by monotonically increasing functions. Again, an area of focusing on attribute 2 and attribute 1 and an area of strong focusing on attribute 1 always exist. The area of strong focusing on attribute 2 exists if and only if $\theta < \frac{1}{\overline{p}-\underline{p}}$. Thus, the general characteristics of the separating lines for the focusing areas remain the same. The incentives for the providers do not change and thus, Proposition 3.1 can be formulated in the same way. Particularly, once patients focus on one attribute and the signal error in this attribute is smaller than in the other attribute, it is a weakly dominant strategy to invest in this attribute. Strong focusing implies strict dominance.

Considering welfare implications, what has to be adjusted is the critical value $\overline{\theta}$ above which the negative welfare effect of a shift in resources from attribute 1 to attribute 2 dominates the positive welfare effects from selection improvements when increasing signal precision in attribute 2. Particularly, $\underline{\theta} = \frac{1-(1-\overline{p})^2-\underline{p}}{\overline{p}-\underline{p}}$.

Symmetric providers. We consider symmetric providers in the sense that both face the same signal errors and the same realization probabilities for a resource allocation decision a^j. If we assumed asymmetric provider in the sense that they might differ in ϵ and p the main drivers of the model are the same. What changes is that the focusing areas of patients might differ across providers. However, if for both providers patients (strongly) focus on the same attribute there is no qualitative difference in the results except that the bounds for the critical θ might change. If for one provider the patient focuses on one attribute and for the other provider on the other attribute, only asymmetric equilibria might exist.

Splitting resources and mixing strategies. We let the providers choose among investing their fixed resources either in attribute 1 or in attribute 2, i.e. they choose $a^j \in \{0, 1\}$. A natural way to extend the set of strategies is to consider

divisible resources and allow providers to choose $a^j \in [0,1]$. We then interpret a^j as a share a^j of the resources being invested in attribute 1 while the other part is invested in attribute 2. The implications of allowing for a budget split crucially depend on how a budget split translates into quality realization probabilities.
Once the probability of high quality realization is concave enough in the share of resources invested in this attribute, splitting resources is not effective enough for high quality realizations and the game we considered in our basic model would basically remain the same. However, there are several other options of how to interpret a budget split in terms of quality realization probabilities, two of which we discuss below. In both cases, the multi-attribute character of the good is important. Furthermore, we show that at least for small θ and errors close enough to $(\epsilon_1, \epsilon_2) = (\frac{1}{2}, 0)$ our results continue to hold.

Mixing strategies. One way to interpret the budget split is interpreting it as mixing strategies $a^j \in \{0,1\}$. Then, quality realization for any $a^j \in [0,1]$ can be denoted as

$$\mathbb{P}(q_1q_2|a^j) = a^j\mathbb{P}(q_1q_2|a^j = 1) + (1 - a^j)\mathbb{P}(q_1q_2|a^j = 0).$$

Particularly, realization probabilities for an equal budget split of $a^j = \frac{1}{2}$ are $\mathbb{P}(ll) = \mathbb{P}(hh) = p(1-p)$ and $\mathbb{P}(hl) = \mathbb{P}(lh) = \frac{1}{2} - p(1-p)$. Since furthermore $\mathbb{P}(q_1 = h) = \mathbb{P}(q_2 = h) = \frac{1}{2}$, for $a^j \notin \{0,1\}$ quality realization in the attributes is not any more independent but negatively correlated (see also discussion about correlation).
Now we consider errors ϵ that are close enough to $(\frac{1}{2}, 0)$. For the extreme case of $\epsilon = (\frac{1}{2}, 0)$, signals in attribute 1 are uninformative while signals in attribute 2 are precise. However, in contrast to our basic model where quality realization is always independent in each attribute, a signal of high quality in attribute 2 is not unambiguously good. Whenever patients believed that a provider mixed, i.e. $a^j \in (0,1)$ a signal h in attribute 2 does not only indicate high quality in attribute 2, but, at the same time indicates that the probability of high quality in attribute 1 is lower than the probability of low quality in attribute 1.

As long as θ is very small, $s_2 = h$ yields higher expected utility than $s_2 = l$. This implies that concentrating resources on 2 is a dominant strategy and our previous results remain valid. However, if θ is very large, depending on the beliefs of patients, $s_2 = l$ might yield higher expected utility than $s_2 = h$. Then, providers have an incentive to concentrate their resources on attribute 1 and it is not valid any more that for errors close enough to $(\frac{1}{2}, 0)$, concentrating resources on attribute 2 is a dominant strategy.

Budget split with keeping independent realization. An alternative way how a budget split $a^j \in [0,1]$ translates into quality realization probabilities is keeping the independent quality realization across attributes and defining the quality realization probability in attribute i as

$$P(q_i = h|a^j) = a^j \mathbb{P}(q_i = h|a^j = 1) + (1 - a^j)\mathbb{P}(q_i = h|a^j = 0).$$

The characteristics of the focusing areas generally remain the same, except that the areas of focusing on 1 and 2 will slightly shrink. For $\epsilon_1 = \frac{1}{2}$, $f^1(\frac{1}{2})$ decreases and $f^2(\frac{1}{2})$ increases compared to our basic model. Particularly, there will be an area where patients neither focus on attribute 1 nor on attribute 2. However, the areas of strong focusing remain exactly the same as before because $a^j \in \{0,1\}$ will be the extreme cases that define the borders.

What might not remain the same are investment incentives. This is because a budget split makes a quality realization of hh more likely compared to a concentration of resources on 1 or 2. For $a^j = \frac{1}{2}$ all possible quality levels are realized with equal probability. Particularly, the probability that hh (as well as ll) is realized is $\frac{1}{4}$ while it is $p(1-p)$ for investing in 1 or investing in 2. Thus, the probabilities for hh and ll increase when splitting the budget while the sum of the probabilities for hl and lh decrease.

However, as long as the signal errors are close enough to $\epsilon = (\frac{1}{2}, 0)$, it is a dominant strategy to concentrate resources on attribute 2. This is because putting more resources on attribute 1 has only marginal effects on signals in attribute 1 while putting more resources on attribute 2 significantly increases the probability of high quality signals in attribute 2 (when patients' beliefs are fixed and with it expected utilities of a specific signal). Thus, our results remain the same at least for errors close enough to $\epsilon = (\frac{1}{2}, 0)$.

Independent quality realization. We assume that quality is realized independently for both attributes. One might think of settings where quality realization in both attributes is correlated, i.e. the probability that high quality is realized differs depending on whether high or low quality was realized in the other attribute. This might be either a positive or negative correlation.

Consider the case of a positive correlation. Focusing can be defined analogously and for focusing on 2 it still holds that investing in 2 is a dominant strategy and equilibria are inefficient. The area for focusing on 2 might be even larger than for independent quality realization. However, whether increasing signal precision in attribute 2 results in a welfare loss depends on how strong the correlation is.

For strong correlations, the selection effect might always dominate the investment effect.

When there is a negative correlation, the mechanisms differ. Again, we can define focusing on attribute 2 as before. However, a signal hh might yield lower expected utility than a signal hl. For a low error in attribute 2, a high error in attribute 1 and a strong negative correlation, a signal hl is an indicator for high quality in attribute 1, while hh indicates low quality in attribute 1. Those effects might result in the area of focusing on 2 being smaller than before, particularly, patients might not always focus for $(\epsilon_1, \epsilon_2) = (\frac{1}{2}, 0)$. However, if it is an equilibrium that both providers invest in 2 the welfare loss when varying ϵ_2 might be even larger.

Observable Resource Allocation. In our model, patients have beliefs about the providers' resource allocations. In the following, we investigate how our results change if patients can observe the resource allocation, but still do not observe the realization of quality and again receives signals about it. The main difference to the case where the resource allocation is unobservable is that by choosing a particular a the providers now send additional information. This has the following effect: Under unobservable provider choice in Proposition 3.1, for a certain belief of a patient a change in a provider's action did not change the expected utility of a signal, but only the probabilities with which the signals are generated. When a is however observable, a change in a provider's action also changes the expected utility of a particular signal.
Then, for parameter constellations where investing in attribute 2 is a strictly dominant strategy under non-observability of provider choice, investing in attribute 1 might be a strictly dominant strategy once resource allocations are observable, since patients now update with the investment choice and demand shifts more strongly. If this is the case, the inefficiency from low expected quality in attribute 1 in equilibrium disappears once the resource allocations are observable. Whether this change occurs depends on the probability $e_2 = \epsilon_2(1-p)+p(1-\epsilon_2)$ that a low signal for attribute 2 is generated if the provider invests in attribute 2. For low e_2, i.e. if the probability that a high signal is generated in attribute 2 remains high, observability of investments does not influence the equilibrium outcome as investing in attribute 2 remains more profitable. However, for large e_2 the equilibrium might differ.

Proposition 3.7. *(Observable Resource Allocation) Fix $\theta < \frac{1}{1-2p}$. Let $\epsilon = (\epsilon_1, \epsilon_2)$ be such that patients strongly focus on attribute 2. Define $e_2 = \epsilon_2(1-p)+p(1-\epsilon_2)$. If $e_2 < 1 - \sqrt{\frac{1}{2}}$ investing in attribute 2 is a strictly dominant strategy such that the corresponding symmetric PBE is unique.*

If $e_2 > \frac{3-\sqrt{5}}{2}$ investing in attribute 1 is a strictly dominant strategy such that the corresponding symmetric PBE is unique.

Proof. See appendix. □

The intuition behind Proposition 3.7 is that if p or ϵ_2 are rather large (which implies that e_2 is rather large), investing in attribute 2 does not payoff for the provider as the probability that only a low signal in attribute 2 is generated is high. On the other hand, for non-observable resource allocations with given patients' beliefs, investing in attribute 2 might be a dominant strategy as for this only that ϵ_2 is small enough is crucial. If e_2 is intermediate such that it is not covered by the bounds presented in the proposition, it depends on the specific combination of the parameters whether investing in attribute 1 or investing in attribute 2 is strictly dominant.

From a welfare perspective, observable resource allocations could enhance efficiency in equilibrium as increasing precision in the less important attribute might not induce the negative resource allocation effect under observable resource allocations. However, it requires that e_2 is large enough. Applying it to our leading example of attribute 2 representing amenities etc., we rather expect a high probability that investments in this attribute are reflected in the signal, i.e. e_2 is low. Furthermore, it might be difficult for patients to interpret resource allocations directly.

Assumption of homogeneous θ. We set-up the model to particularly look at a situation where patients are homogeneous and all have a higher utility from high quality in one attribute, i.e. where results are not driven by heterogeneous patient valuations for attributes. However, patients might of course differ in the utility θ of high quality in the first attribute compared to high quality in the second attribute.[21] For different clinical areas different θ hold. For instance, θ for patients suffering from cancer should be rather high as clinical factors are

[21]Note that ex-post differences in θ, i.e. differences that occur after the decision for a provider, can be considered as being already incorporated in θ when interpreting utilities for each quality state as expected utilities. Reasons for ex-post heterogeneity includes e.g. differences in quality perception.

much more important than amenities. On the other hand, for births θ might be rather low as generally not many complications are expected. Our results than can be applied for each health area separately. In areas with a high θ investing in attribute 1 is an equilibrium while in areas with a low θ investing in attribute 2 might be an equilibrium.

Even within one area θ might differ among patients. Reasons might be differences in individual preferences or the severity of the individual patient's health case. Consider any signal error $\epsilon = (\epsilon_1, \epsilon_2)$ with $\epsilon_2 < \epsilon_1$. This implies that there is a threshold θ_2 such that for $\theta < \theta_2$ the patients strongly focus on attribute 2. It is clear that if for each patient $c \in C$, $\theta_k < \theta_2$ holds, investing in attribute 2 is a dominant strategy. Analogously, there is a threshold θ_1 such that if for each patient $c \in C$, $\theta_k > \theta_1$ holds, investing in attribute 1 is a dominant strategy. Generally, which effect dominates depends on the distribution of θ in the population. If the mass of patients whose θ is below (above) the respective critical thresholds is sufficiently large, then investing in attribute 2 (1) is an equilibrium outcome.

3.9 Conclusion

We model quality competition among health care providers in a market where health care services have multiple quality attributes and patients observe attribute quality only imperfectly before deciding on a provider. A patient focuses on a particular attribute if a high quality signal in this attribute drives her provider choice. Focusing is strong if this is the case for all combinations of beliefs that the patient has about the underlying resource allocations of providers. We show that, even if high quality in one attribute is less important in terms of patient utility, patients might focus on this attribute such that providers invest in quality improvement in this attribute. If signal precision is such that patients focus on this less important attribute, any equilibrium is inefficient. An increase in signal precision can then lead to a welfare reduction as the positive effect of a better provider selection from an increase in signal precision might be overcompensated by the negative effect that a shift in patient focusing has on provider quality choice. When providers can choose reporting in the form of sending informative signals strategically, we furthermore show that providers do not report in all attributes such that not only resource allocations, but also reporting might be inefficient.

In health care, there has been an increase in the availability of information about provider quality via e.g. quality reporting requirements or public feedback platforms. For hospital report cards, most empirical literature finds positive but small patient reactions to publicized quality information. Our model is fully consistent with the positive demand effect: if quality reporting reduces signal error only in the medical attribute, it unambiguously increases welfare if the effect is strong enough. However, reporting requirements or the increasing availability of public feedback platforms often also improve the precision of information about other dimensions. Better overall information about health care providers might however imply a higher relative precision of information in the less important quality attributes like the hotel properties of hospitals, with adverse effects on quality. For overall welfare, the quality reporting policy is crucial. While under optimal reporting signals in the more important attributes are always published, banning reporting in less important attributes might be necessary.

3.10 Appendix

Preliminaries

Before turning to the proofs we introduce a notation that will be helpful to calculate the expected utilities $U_s[s|1,\epsilon]$ and $U_s[s|0,\epsilon]$ when receiving a signal s, facing signal errors ϵ and having a belief $b = 1$ or $b = 0$.

Quality realizes independently for each attribute. Therefore, we can calculate the expected utilities separately for each attribute for $b \in \{0, 1\}$. To calculate and compare expected utilities the following function will be useful to us.

$$f(y,z) := \frac{yz}{yz + (1-y)(1-z)} \quad \text{for} \quad y \in [0,1], z \in (0,1)$$

The function $f(y,z)$ has the following properties

- $f(y,z) = f(z,y)$ and $f(y,z)$ is increasing in y and in z
- $f(y,z) + f(1-y, 1-z) = 1$
- $f(y,1-z) - f(y,z) = f(1-y,1-z) - f(1-y,z)$ is decreasing in z and symmetrically spread around $y = \frac{1}{2}$. For $z < \frac{1}{2}$ it is increasing in $y \in (0, \frac{1}{2})$ and decreasing in $y \in (\frac{1}{2}, 1)$, analogously for $z > \frac{1}{2}$ it is decreasing in $y \in (0, \frac{1}{2})$ and increasing in $y \in (\frac{1}{2}, 1)$.

To see how the function is related to expected utilities when observing signals consider any signal s_i about quality in attribute i, any corresponding signal error

ϵ_i and any belief $b \in \{0,1\}$ the patients might have. The expected utility U_{s_i} when observing $s_i \in \{l,h\}$ in attribute i then is

$$\begin{aligned}
U_{s_i}[s_i|b,\epsilon_i] &= \mathbb{P}(q_i = h|s_i, b)u_i(q_i = h) \\
&= \frac{\mathbb{P}(s_i|q_i = h)\mathbb{P}(q_i = h|b)}{\mathbb{P}(s_i|b)}u_i(q_i = h) \\
&= \frac{\mathbb{P}(s_i|q_i = h)\mathbb{P}(q_i = h|b)}{\mathbb{P}(s_i|q_i = h)\mathbb{P}(q_i = h|b) + \mathbb{P}(s_i|q_i = l)\mathbb{P}(q_i = l|b)}u_i(q_i = h) \\
&= f(y(s_i), z(i))u_i(q_i = h) \quad \text{with } y = \mathbb{P}(s_i|q_i = h) \text{ and } z = \mathbb{P}(q_i = h|b)
\end{aligned}$$

$y(s_i = h) = 1 - \epsilon_i$ and $y(s_i = l) = \epsilon_i$. z_i is the probability that high quality is served in attribute i, therefore $z_1 = 1 - z_2 = 1 - p$ if $b = 1$ and $z_1 = 1 - z_2 = p$ if $b = 0$. Thus, whenever patients receive a signal $s = s_1 s_2$ from any provider and have a belief $b \in \{0,1\}$ the expected utility if choosing this provider is as follows.

$$\begin{aligned}
U_s[s_1s_2|1,\epsilon] &= f(y(s_1), 1-p)\theta + f(y(s_2), p) \quad \text{with} \quad y(s_i = h) = 1 - \epsilon_i = 1 - y(s_i = l) \\
U_s[s_1s_2|0,\epsilon] &= f(y(s_1), p)\theta + f(y(s_2), 1-p) \quad \text{with} \quad y(s_i = h) = 1 - \epsilon_i = 1 - y(s_i = l)
\end{aligned}$$

Proof of Lemma 3.1

To define the separating lines for the areas of focusing, the difference in expected utilities when observing signal hl with underlying belief $b \in \{0,1\}$ and signal lh with underlying belief $b' \in \{0,1\}$ is crucial. It will be convenient to use beliefs about the probability x of high quality realization in attribute 1 instead of beliefs b about the resource allocation. Then, $x = 1-p$ for $b = 1$ and $x = p$ for $b = 0$. In the following, when we use b we refer to the beliefs $b \in \{0,1\}$ about the actions of the providers and when we use x we refer to corresponding beliefs $x \in \{p, 1-p\}$ about the high quality realization in attribute 1. We define

$$\begin{aligned}
g(x, x', \epsilon_1\epsilon_2) &= U_s[hl, b, \epsilon_1\epsilon_2] - U_s[lh, b', \epsilon_1\epsilon_2] && (3.2)\\
&= [f(1-\epsilon_1, x) - f(\epsilon_1, x')]\theta - [f(1-\epsilon_2, 1-x') - f(\epsilon_2, 1-x)] && (3.3)\\
&= [f(1-\epsilon_1, x) - f(\epsilon_1, x')]\theta - [f(1-\epsilon_2, x) - f(\epsilon_2, x')] && (3.4)
\end{aligned}$$

where f is defined in the preliminaries and the last inequality is implied by the characteristics of f.

The sign of g is important for the focusing of the patients since $(hl|b) \succ (lh|b) \Leftrightarrow g(b, b', \epsilon_1\epsilon_2) > 0$ and $(lh|b') \succ (hl|b) \Leftrightarrow g(b, b', \epsilon_1\epsilon_2) < 0$. Equation (3.4) together with the fact that $f(y,z)$ is strictly increasing in y for $z \neq 0$ we can deduce that

$g(b, b', \epsilon_1, \epsilon_2)$ ist strictly decreasing in ϵ_1 and strictly increasing in ϵ_2. Therefore, if for $(\epsilon_1^*, \epsilon_2^*)$ a patient (strictly) focuses on attribute 2 he (strictly) focuses on attribute 2 for all $(\epsilon_1, \epsilon_2^*)$ with $\epsilon_1 > \epsilon_1^*$ and $(\epsilon_1^*, \epsilon_2)$ with $\epsilon_2 < \epsilon_2^*$ as well. The same holds for (strict) focusing on attribute 1 with reversed signs.

Definition of the separating lines. We use the function g to describe the separating lines of the four focusing areas. For a fixed ϵ_2 define $\epsilon_1^*(x, x')$ as the unique root of $g(x, x', \cdot\epsilon_2)$ if existent and $\frac{1}{2}$ otherwise. If existent, the root is unique because of the monotonicity characteristics.

- $f^{s2}(\epsilon_2) = \max_{x,x'}\{\epsilon_1^*(x, x')|x, x' \in \{p, 1-p\}\}$
- $f^2(\epsilon_2) = \max_x\{\epsilon_1^*(x, x)|x \in \{p, 1-p\}\}$
- $f^1(\epsilon_2) = \min_x\{\epsilon_1^*(x, x)|x \in \{p, 1-p\}\}$
- $f^{s1}(\epsilon_2) = \min_{x,x'}\{\epsilon_1^*(x, x')|x, x' \in \{p, 1-p\}\}$

Once we show that $f^2 = f^1$ and define $f^{12} = f^1 = f^2$, the focusing behavior as described in the lemma follows by the definitions of the functions. For this note that $\epsilon_1 = 0$ implies $g = \theta - [f(1-\epsilon_2), x) - f(\epsilon_2, x')] > 0$ independent of the beliefs.

Characteristics of the separating lines. Since g is continuous and monotonically decreasing in ϵ_1 and increasing in ϵ_2 the functions f^i are continuous and increasing in ϵ_2. The more specific characteristics are as follows.

Focusing on 1 *or* 2: For any symmetric beliefs, g can be described by

$$g(x, x, \epsilon_1 \epsilon_2) = [f(1-\epsilon_1, x) - f(\epsilon_1, x)]\theta - [f(1-\epsilon_2, x) - f(\epsilon_2, x)].$$

$g(b, b, 0\epsilon_2) = \theta - [f(1-\epsilon_2, x) - f(\epsilon_2, x)] > 0$ and $g(b, b, \frac{1}{2}\epsilon_2) = 0 - [f(1-\epsilon_2, x) - f(\epsilon_2, x)] \leq 0$. Strict monotonicity of g in ϵ_1 therefore implies that there exists a unique root $\epsilon_1^*(x, x)$ such that $g(x, x, \epsilon_1^* \epsilon_2) = 0$. Furthermore, since $f(1-\epsilon_i, x) - f(\epsilon_i, x) = f(1-\epsilon_i, 1-x) - f(\epsilon_i, 1-x)$ the unique root ϵ_1^* of $g(b, b, \epsilon_1 \epsilon_2)$ is the same for $b = 0$ and $b = 1$. Thus, we can define $f^{12} = f^1 = f^2 = \epsilon_1^*(1-p, 1-p)$. For $\epsilon_2 = \frac{1}{2}$ we have $f^{12}(\frac{1}{2}) = \frac{1}{2}$.

Since $\theta > 1$, for any ϵ_2 the function $g(x, x, \epsilon_1 \epsilon_2)$ can only be 0 if $\epsilon_1 \geq \epsilon_2$. Thus, $f^1 = f^2$ lies above the 45-degree line and patients focus on 1 for any errors with $\epsilon_1 \leq \epsilon_2$.

Strong focusing on 2*:* Consider $\theta = \frac{1}{1-2p}$. Then, for any beliefs x, x' we get $g(x, x', \frac{1}{2}0) = (x - x')\frac{1}{1-2p} - 1 \leq 0$ with equality for $x = 1-2p$ and $x' = p$.

Therefore, for $\theta > \frac{1}{1-2p}$ there always exist beliefs such that no root exist and therefore $f^{s2}(\epsilon_2) = \frac{1}{2}$ for all ϵ_2. No assume $\theta < \frac{1}{1-2p}$. Particularly, $0 < f^{s2}(0) < \frac{1}{2}$. and $f^{s2}(\frac{1}{2}) = \frac{1}{2}$. Define ϵ_2^* such that $g(1-p, p, \frac{1}{2}\epsilon_2^*) = 0$. Then the following holds: $0 < \epsilon_2^* < \frac{1}{2}$ and for all $\epsilon_2 > \epsilon_2^*$ no root of $g(1-p, p, \cdot\epsilon_2^*)$ exists and therefore $f^{s2}(\epsilon_2) = \frac{1}{2}$ for all $\epsilon_2 > \epsilon_2^*$.

Strong focusing on 1*:* For $\epsilon_1 = 0$ the function g is always larger than zero (independent of the belief and ϵ_2). For $\epsilon_1 = \frac{1}{2}$ we have $g(p, 1-p, \frac{1}{2}\epsilon_2) < 0$ independent of ϵ_2. Therefore, the minimum root of $g(x, x', \cdot\epsilon_2)$ is always larger than zero and smaller than $\frac{1}{2}$ which shows $0 < f^{s1}(\epsilon_2) < \frac{1}{2}$ for all ϵ_2. For $\epsilon_2 = 0$ the function g has the form

$$g(x, x', \epsilon_1\epsilon_2) = [f(1-\epsilon_1, x) - f(\epsilon_1, x')]\theta - 1.$$

The smallest root ϵ_1^* occurs for beliefs that minimize g. This is the case for $x = p$ and $x' = 1-p$. Therefore, $f^{s1}(0) = \epsilon_1^*$ with ϵ_1^* being the root of $g = f(1-\epsilon_1, p) - f(\epsilon_1, 1-p)\theta - 1$. This shows that $f^{s1}(0) < p$ because for $\epsilon_1 = p$ it is still negative. The same argument holds to show that $f^{s1}(\frac{1}{2}) > p$.

Remark. Comparable to focusing on 1 or 2 we can more explicitly specify the separating lines for strong focusing by defining $f^{s2} = \epsilon_1^*(1-p, p)$ if the root exists and $f^{s2} = \frac{1}{2}$ otherwise and $f^{s1} = \epsilon_1^*(p, 1-p)$.
For strong focusing on 2 note that we already know that strong focusing on 2 implies that $\epsilon_1 \leq \epsilon_2$. However, for all $\epsilon_1 \leq \epsilon_2$, $U_s[s|1, \epsilon] \leq U_s[s|0, \epsilon]$. This is because

$$(s|1) \succ (s|0) \quad \Leftrightarrow \quad [f(\epsilon_1, 1-p) - f(\epsilon_1, p)]\theta > f(\epsilon_2, 1-p) - f(\epsilon_2, p). \quad (3.5)$$

Here, we again exploited the characteristics of f described in the preliminaries. Therefore, $g(x, x', \epsilon)$ is maximal for $x = 1-p$ and $x = p$ and it is sufficient to find the root for this combination of beliefs.

For strong focusing on 1 errors can be such that $\epsilon_1 \leq \epsilon_2$ and for any fixed ϵ_2 we have $(s|1) \succ (s|0)$ for high ϵ_1 and $(s|0) \succ (s|1)$ for low ϵ_1. However, as (3.5) shows for any ϵ_2 fixed such that this ambivalence exists, there is a unique $\hat{\epsilon}_1$ such that for $\epsilon = (\hat{\epsilon}_1, \epsilon_2)$, $(hl|1) = (hl|0)$ which is equivalent to $(lh|1) = (lh|0)$. Furthermore, for $\epsilon = (\hat{\epsilon}_1, \epsilon_2)$ patients focus on 1 because the error in attribute 1 has to be smaller than the error in attribute 2. Then, for $\epsilon = (\hat{\epsilon}_1, \epsilon_2)$ the patient also strongly focuses on attribute 1 since $(hl|1) \succ (lh|1) = (lh|0)$ and $(hl|0) > (lh|0) = (lh|1)$. This implies that $f^{s_1}(\epsilon_2)$ lies above $\hat{\epsilon}_1$ and that the line of strong focusing can be defined as $f^{s1} = \epsilon_1^*(0, 1)$.

Dependence on θ. First, consider $\theta \to 1$. Then the function g converges to

$$g(x, x', \epsilon_1\epsilon_2) = [f(1-\epsilon_1, x) - f(\epsilon_1, x')] - [f(1-\epsilon_2, 1-x') - f(\epsilon_2, 1-x)]$$

For any beliefs x and x' the function g is zero if and only if $\epsilon_1 = \epsilon_2$ (it can be easily seen that it holds for $x = x'$. Analogously, it holds for $x = p$ and $x' = 1-p$ as well as for $x = 1-p$ and $x' = p$). Thus

$$(hl|0) = (lh|1) = (hl|1) = (lh|0).$$

Therefore, for $\theta \to 1$ the expected utilities when observing hl or lh are the same independent of the underlying beliefs and thus all separating functions converge to

$$f^{s2}(\epsilon_2) = f^{12}(\epsilon_2) = f^{s1}(\epsilon_2) = \epsilon_2.$$

Second, consider $\theta \to \infty$. For f^{s2} we have already seen that $f^{s2} = \frac{1}{2}$ for all $\theta > \frac{1}{1-2p}$. If $x = x'$ and θ is arbitrary high, the function g is always positive except for the case that $\epsilon_1 = \frac{1}{2}$. Therefore, $f^{12} = \frac{1}{2}$ as well. For other beliefs, the minimum ϵ_1 for which the function g with $\theta \to \infty$ is zero, is $\epsilon_1 = p$. Therefore, f^{s1} converges to $f^{s1}(\epsilon_1) = p$.

Proof of Corollary 3.1

The first part of the corollary is directly implied by the characteristics of the separating lines discussed in the previous lemma: If ϵ_1 is large enough, patients focus on 2 for $\epsilon_2 = 0$. For $\epsilon_2 = \frac{1}{2}$ they anyway focus on 1.

For the second part of the corollary is sufficient to show that for $\epsilon = (p, 0)$ and $\theta < \frac{1}{1-2p}$ the patient focuses on attribute 2. This is sufficient because the Lemma implies that for $\epsilon = (p, \frac{1}{2})$ the patient strongly focuses on attribute 1.

For $\epsilon = (p, 0)$ and any belief x we have to show that $(lh|x) > (hl|x)$ for $\theta < \frac{1}{1-2p}$. $(lh|1) > (hl|1)$ is equivalent to $f(p, 1-p)\theta + 1 < f(1-p, 1-p)\theta$. This is equivalent to $\theta < \frac{(1-p)^2+p^2}{1-2p}$. As for all p, $(1-p)^2 + p^2 > 1$ it is sufficient to choose $\theta < \frac{1}{1-2p}$.

Proof of Proposition 3.1

First, we show that for focusing on attribute i in combination with errors $\epsilon_i < \epsilon_{-i}$ and any beliefs and the other provider's strategy, investing in i is a weakly better strategy than investing in the other attribute. Second, we show that this implies weak dominance of investing in i, i.e. with the first part it remains to show that there is at least one combination of beliefs and the other providers' strategy such that investing i is strictly better than investing in the other attribute. Third, we show that strong focusing on i and $\epsilon_i < \epsilon_{-i}$ imply strict dominance, i.e. investing in attribute i is strictly better for all beliefs and the other provider's strategy.

Investing in i is weakly better than investing in $-i$. The main idea is that, independent of whether the provider invests in attribute 1 or attribute 2, the same signals are generated. For given beliefs, the expected utility of each possible signal does not depend on the allocation decision. What does depend on the allocation decision is the probability of each signal. For focusing on attribute i investing in i generates "better signals" (i.e. they yield a higher expected utility for patients) with higher probability compared to investing in the other attribute.

Focusing on attribute 1 and $\epsilon_1 < \epsilon_2$: Assume each patient has any belief (b^A, b^B) about the providers' strategy (possibly not the same for each provider and beliefs might differ across patients). Let provider B have any strategy (possibly not known to provider A). We have to show that it is a weakly dominant strategy for provider A to invest in attribute 1, i.e., $a^A = 1$.

Each patient either receives signal ll, lh, hl or hh from provider A. Independent of her belief b^A about provider A's resource allocation, focusing on 1 implies that each patient faces the following ordering of signals with respect to expected utilities if received from provider A:

$$(hh|b^A) \succ (hl|b^A) \succ (lh|b^A) \succ (ll|b^A).$$

The expected utility of a patient receiving s from provider A and having belief b^A is

$$U_s[s|b^A, \epsilon] = \sum_q u(q)\mathbb{P}(q|s, b^A, \epsilon).$$

Importantly, the allocation decision of the providers does not influence the expected utilities that patients with a belief b^A are facing when receiving a signal s. However, the probabilities of the signals depend on the allocation decision of the provider. For $a^A = 1$ and $\epsilon_1 \leq \epsilon_2$ the ordering is

$$\mathbb{P}(s = lh|1) < \mathbb{P}(s = ll|1) \leq \mathbb{P}(s = hh|1) < \mathbb{P}(s = hl|1).$$

For $a^A = 0$ this ordering is reversed with $\mathbb{P}(s = lh|0) = \mathbb{P}(s = hl|1)$, $\mathbb{P}(s = hh|0) = \mathbb{P}(ll|1)$, $\mathbb{P}(s = ll|0) = \mathbb{P}(s = hh|1)$ and $\mathbb{P}(s = hl|0) = \mathbb{P}(s = lh|1)$. Thus, for choosing $a^A = 1$ instead of $a^A = 0$ some part of the probability of $s = ll$ is shifted to hh, and from $s = lh$ to $s = hl$ (better signals have more weight). Therefore, for any allocation strategy of B, provider A is selected by any patient with weakly higher probability when choosing $a^A = 1$ instead of $a^A = 0$. This holds independent of the beliefs b about the allocation decision of A and B.[22]

Focusing on attribute 2: The approach is the same as above. Note that focusing on attribute 2 immediately implies that $\epsilon_1 > \epsilon_2$. If patients focus on attribute 2 the signal ordering for any belief b^A is the following

$$(hh|b^A) \succ (lh|b^A) \succ (hl|b^A) \succ (ll|b^A).$$

The signal probabilities for playing $a^A = 1$ have the ordering

$$\mathbb{P}(s = lh|1) < \mathbb{P}(s = hh|1) \leq \mathbb{P}(s = ll|1) < \mathbb{P}(s = hl|1).$$

Here we used that $\epsilon_1 > \epsilon_2$ holds. For choosing $a^A = 0$ the ordering reverses with $\mathbb{P}(s = lh|0) = \mathbb{P}(s = hl|1)$, $\mathbb{P}(s = hh|0) = \mathbb{P}(ll|1)$, $\mathbb{P}(s = ll|0) = \mathbb{P}(s = hh|1)$ and $\mathbb{P}(s = hl|0) = \mathbb{P}(s = lh|1)$. Thus, in this case $a^A = 0$ influences the signal probabilities such that better signals have higher probabilities.

Focusing implies weak dominance. Take any symmetric belief $(b^A, b^B) = (b, b)$ of the patients and any strategy a^B of provider B. We assume that parameters are such that patients focus on 2. It is then sufficient to show that it is strictly better for A to invest in 2 than to invest in 1.

Assume that B sends a signal $s^B = hl$. If A sends hl as well, A is selected with probability $\frac{1}{2}$. However, if A sends lh he is selected with probability 1. As choosing $a^j = 0$ instead of $a^j = 1$ shifts some of the probability of sending hl to sending lh (see first part of the proof), A can strictly increase his probability of being selected by choosing $a^j = 0$ instead of $a^j = 1$.

Arguments for focusing on 1 and $\epsilon_1 < \epsilon_2$ are the same.

Strong focusing implies strict dominance. We now show that for strong focusing on attribute 2 provider A strictly prefers to invest in attribute 2, independent of the beliefs and the resource allocation of provider B. The same arguments hold for strong focusing on attribute 1 and $\epsilon_1 < \epsilon_2$.

Assume that patients have any beliefs b^A and b^B and that provider B has chosen any a^B. Strong focusing implies that provider A is selected when sending signal ll while B sends hh. B is selected when sending hh while A sends ll. Now

[22]Note that for focusing on attribute 1 we needed $\epsilon_1 \leq \epsilon_2$ to conclude that more preferred signals are generated with higher probability. For $\epsilon_1 > \epsilon_2$ the effect is ambiguous. $a = 1$ still makes hl more probable on the cost of lh and leads to an increase in the expected profit of the provider. However, at the same time, ll is more probable on the cost of hh and therefore leads to a decrease in expected profit for the provider.

assume that both send ll or both send hh. We show that the probability that A is selected is the same in both cases and then show that this is sufficient to show that A is selected with strictly higher probability for $a^j = 0$ instead of $a^j = 1$.

Consider any beliefs $x, x' \in \{p, 1-p\}$ about the quality realization where $x = 1-p$ corresponds to a belief $b = 1$ and $x = p$ corresponds to a belief $b = 0$ (as discussed in the beginning of the proof of Lemma 3.1).

$(ll|x) \succeq (ll|x')$ is equivalent to $[f(\epsilon_1, x) - f(\epsilon_1, x')]\theta \geq [f(\epsilon_2, x) - f(\epsilon_2, x')]$. For $(hh|x) \succeq (hh|x')$ we just have to replace ϵ_i by $1 - \epsilon_i$. If $x = x'$, the inequality is satisfied both for ll and hh. For asymmetric x and x' the inequality for ll is equivalent to the one for hh. Therefore, the probability that A is selected if both providers send the signal hh equals the probability that A is selected if both providers send the signal ll.

This implies that A is strictly better off when choosing $a^A = 0$ instead of $a^A = 1$: First, assume that A is selected with probability 1 if both send hh or ll. Assume that B signals hh. By the proof of Proposition 3.1 we know that selecting $a^A = 0$ instead of $a^A = 1$ shifts probabilities from sending worse signals to better signals. In particular, from sending ll to sending hh. Since $\epsilon_2 < \epsilon_1$ the amount of probability shifted is not zero. If A sends ll, B is selected, if A sends hh, A is selected. Therefore, the shift in probabilities results in strict increase of the probability to be selected. Now assume that A is selected with probability less than 1 if both send hh or ll and assume that B signals ll. Then, A is selected with probability 1 when signaling hh but is selected with probability less then 1 when signaling ll. Here again, the shift in probabilities from ll to hh results in strict increase of the probability to be selected.

Proof of Proposition 3.2

We first show the parts of the proposition that claim that strong focusing on an attribute implies uniqueness of the PBE. Then we show that any further equilibria are asymmetric and who is selected in asymmetric equilibria. Finally we discuss the further conditions for uniqueness.

Strong focusing implies uniqueness. The uniqueness for strong focusing and corresponding errors is directly implied by the proof of the strict dominance of Proposition 3.1. For both providers it is - independent of the beliefs and the other provider's strategy - strictly better to invest in the attribute the patients strongly focus on.

Asymmetric equilibria. Assume that patients focus on attribute 2 and $(a^A, a^B) = (b^A, b^B) \neq (0,0)$ is an equilibrium. First, the equilibrium is not symmetric, i.e., $a^A \neq a^B$. This is because for symmetric beliefs investing in attribute 2 is strictly preferred by any provider to investing in attribute 1 (see above). Second, we want to show that provider A is selected with probability one if and only if $a^A > a^B$. Assume that $a^A > a^B$, i.e. $a^A = 1$ and $a^B = 0$. We want to show that this implies already $(ll|x^A) \succ (hh|x^B)$ which means that provider A is selected independent of the signal. Assume the contrary, i.e., $(hh|a^B)$ yields at least the same expected utility as $(ll|a^A)$. Then provider A has an incentive to deviate by choosing $a^A = 0$ instead of $a^B = 1$: From the proof of Proposition 3.1 we know $a^A = 0$ is weakly better than $a^A = 1$. If $(ll|x^A) \succ (hh|x^B)$ does not hold, it is also strictly better because if B sends hh and A sends ll, provider B is selected with strictly positive probability. If A sends hh and B send hh, on the other hand, provider B is never chosen because $b^A = 1$ and $b^B = 0$. As a shift from $a^A = 1$ to $a^A = 0$ generates signal hh with higher probability on the cost of sending signal ll and all other shifts in probabilities are weakly better as well it is strictly dominant for A to invest in attribute 2. This is a contradiction to the assumption that $a^A > a^B$ is the providers' strategy in equilibrium. Thus, if $(a^A, a^B) = (b^A, b^B) \neq (0,0)$ is a PBE and $a^A > a^B$, provider A is selected with probability one. On the other hand, if provider A is selected with probability one, $a^A > a^B$ has to hold.

The part for focusing on attribute 1 follows by the same arguments.

Further conditions for uniqueness. Assume that for $\epsilon = (\epsilon_1, \epsilon_2)$ patients focus on attribute 2 and strictly focuses on attribute 1 for $\epsilon' = (\epsilon_1, \frac{1}{2})$. Assume that for $\epsilon = (\epsilon_1, \epsilon_2)$ the equilibrium is not unique. Particularly, this implies that $(ll|1) \succ (hh|0)$ which is equivalent to

$$[f(\epsilon_1, 1-p) - f(1-\epsilon_1, p)]\theta > f(1-\epsilon_2, 1-p) - f(\epsilon_2, p) = 2f(1-\epsilon_2, 1-p) - 1.$$

The right hand side is decreasing in ϵ_2, therefore if $(ll|x^A) \succ (hh|x^B)$ holds for $\epsilon = (\epsilon_1, \epsilon_2)$ it holds as well when ϵ_2 increases and particularly for $\epsilon' = (\epsilon_1, \frac{1}{2})$. If the patient strongly focuses on attribute 1 for $\epsilon' = (\epsilon_1, \frac{1}{2})$ it is a contradiction because then $(hh|0) \succ (ll|1)$.

Now assume that ϵ_1' is such that for $(\epsilon_1', \epsilon_2)$ patients strongly focus on attribute 2. Now assume that for (ϵ_1, ϵ_2) the equilibrium is not unique. This implies particularly $\epsilon_1 < \epsilon_1'$ and that $(ll|1) \succ (hh|0)$ which is again equivalent to

$$[f(\epsilon_1, 1-p) - f(1-\epsilon_1, p)]\theta < f(1-\epsilon_2, 1-p) - f(\epsilon_2, p)$$

The left hand side is increasing in ϵ_1. Thus, if it holds for any ϵ_1, it also holds for $\epsilon_1' > \epsilon_1$. This contradicts that for $(\epsilon_1', \epsilon_2)$ the patient strongly focuses on attribute 2 since then $(hh|0) \succ (ll|1)$.

Proof of Lemma 3.2

We fix any $\epsilon = (\epsilon_1, \epsilon_2)$ and therefore omit it in the following. We will first show that for given symmetric patients' beliefs with $b^A = b^B = b$ about the providers' resource allocation,

$$W[(1, a^B)|(b, b)] > W[(0, a^B)|(b, b)]$$

for any $a^B \in \{0, 1\}$, i.e., $W[(1, 0)|(b, b)] > W[(0, 0)|(b, b)]$ and $W[(1, 1)|(b, b)] > W[(0, 1)|(b, b)]$. From symmetry of $W[.]$ with respect to providers it then follows that $W[(1, 1)|(b, b)] > W[(0, 0)|(b, b)]$. Since for any symmetric beliefs patients make the very same selection of providers based on signals they receive, this then also implies that $W[(1, 1)|(1, 1)] > W[(0, 0)|(0, 0)]$.
Note that the only variables in

$$W[(a^A, a^B)|(b, b)] = \sum_{q^B} \sum_{q^A} \mathbb{P}(q^A|a^A)\mathbb{P}(q^B|a^B)U_q[q^A, q^B|b] \tag{3.6}$$

that depend on the resource allocation decision of provider A are $\mathbb{P}(q^A|a^A)$ for $q^A = hl$ and $q^A = lh$. This is because $\mathbb{P}(hh|a^A) = \mathbb{P}(ll|a^A) = (1-p)p$ for all a^A. Thus, we need to show that

$$\begin{aligned}
& \sum_{q^B} \mathbb{P}(q^B|a^B)[\mathbb{P}(hl|1)U_q[hl, q^B|b] + \mathbb{P}(lh|1)U_q[lh, q^B|b]] && (3.7)\\
& > \sum_{q^B} \mathbb{P}(q^B|a^B)[\mathbb{P}(hl|0)U_q[hl, q^B|b] + \mathbb{P}(lh|0)U_q[lh, q^B|b]] && (3.8)\\
\Leftrightarrow\ & \sum_{q^B} \mathbb{P}(q^B|a^B)[(1-p)^2 U_q[hl, q^B|b] + p^2 U_q[lh, q^B|b]] && (3.9)\\
& > \sum_{q^B} \mathbb{P}(q^B|a^B)[p^2 U_q[hl, q^B|b] + (1-p)^2 U_q[lh, q^B|b]] && (3.10)\\
\Leftrightarrow\ & \sum_{q^B} \mathbb{P}(q^B|a^B)U_q[hl, q^B|b] > \sum_{q^B} \mathbb{P}(q^B|a^B)U_q[lh, q^B|b] && (3.11)
\end{aligned}$$

For $q^B = lh$ and $q^B = hl$ we have $U_q[hl, q^B|b] \geq U_q[lh, q^B|b]$. Furthermore, $\mathbb{P}(hh|a^b) = \mathbb{P}(ll|a^b) = p(1-p)$ independent of a^b. It thus remains to show that

$$U_q[hl, hh|b] + U_q[hl, ll|b] > U_q[lh, hh|b] + U_q[lh, ll|b]. \tag{3.12}$$

Note that $U_q[q^A, q^B|b] = u(q^A)\mathbb{P}(q^A|q^A, q^B, b) + u(q^b)(1 - \mathbb{P}(q^A|q^A, q^B, b))$ where $\mathbb{P}(q^A|q^A, q^B, b)$ is the probability that q^A is chosen by the patient if quality levels q^A and q^B are realized, patient has belief b and the signal error is ϵ. Thus the previous inequality is equivalent to

$$
\begin{aligned}
& (u(hl) - u(hh))\mathbb{P}(hl|hl, hh, b) + (u(hl) - u(ll))\mathbb{P}(hl|hl, ll, b) && (3.13)\\
> \quad & (u(lh) - u(hh))\mathbb{P}(lh|lh, hh, b) + (u(lh) - u(ll))\mathbb{P}(lh|lh, ll, b) && (3.14)\\
\Leftrightarrow \quad & \theta\mathbb{P}(lh|lh, hh, b) + \theta\mathbb{P}(hl|hl, ll, b) > \mathbb{P}(hl|hl, hh, b) + \mathbb{P}(lh|lh, ll, b) && (3.15)
\end{aligned}
$$

As $b = b^A = b^B$ is the belief for both providers,

$$
\begin{aligned}
\mathbb{P}(hl|hl, hh, b) = \;& \mathbb{P}(ll|ll, lh, b) &= 1 - \mathbb{P}(lh|lh, ll, b) \qquad (3.16)\\
\mathbb{P}(hl|hl, ll, b) = \;& \mathbb{P}(hh|hh, hl) &= 1 - \mathbb{P}(hl|hl, hh, b) \qquad (3.17)
\end{aligned}
$$

Inserting this into the above inequality reduces the inequality to $\theta > 1$ which holds by definition of θ in our model.

Proof of Proposition 3.3

It is to show that for focusing on attribute 2, any PBE is inefficient. For strong focusing, this follows directly by the discussion above. For focusing, first consider the symmetric BNE where both providers invest in attribute 2. Then quality provision is inefficient by Proposition 3.3. Second, consider any other BNE $(a^A, a^B) = (b^A, b^B)$ with $a^A > a^B$. Proposition 3.2 showed that patients then choose provider A ignoring the signals sent. Thus, expected utility is $(1-p)\theta + p$ since $a^A = 1$. If both providers invest in attribute 1 and patients have corresponding beliefs, welfare is strictly higher as signals are then valuable to patients and by selection based on the signals they receive an expected utility higher than $(1-p)\theta + p$.

Proof of Proposition 3.4

First note that $W[a|b, (\epsilon_1, \epsilon_2)]$ is decreasing in both ϵ_1 and ϵ_2 (the more precise signals the better the patient can select). Therefore, for any ϵ_1 fixed it is sufficient to show the inequality for $\epsilon_2 = \frac{1}{2}$ and $\epsilon_2' = 0$ because this then implies that the inequality holds for any other ϵ_2 and ϵ_2'.

Denote

$$\Delta W_{10}(\epsilon_1) = W[(1,1)|(1,1), (\epsilon_1, \frac{1}{2})] - W[(0,0)|(0,0), (\epsilon_1, 0)].$$

We first show that $\Delta W_{10}(\frac{1}{2}) > 0$ and then show that this implies the inequality for all other ϵ_1.

To show that $\Delta W_{10}(\frac{1}{2}) > 0$ holds we explicitly calculate the expected utilities. For $a = (1,1)$, $\epsilon_2 = \frac{1}{2}$ and corresponding beliefs the signals are of no value for patients and therefore

$$W[(1,1)|(1,1),(\frac{1}{2},\frac{1}{2})] = (1-p)\theta + p.$$

For $a = (0,0)$, $\epsilon_2 = 0$ and corresponding beliefs $b = (0,0)$ the patient receives no signal in the first attribute and a precise signal in the second attribute. Thus, in the first attribute high quality is realized with probability p while in the second attribute high quality is realized with probability $1-p^2$ (the patient focuses on attribute 2 and therefore she only picks low quality in the second attribute if both providers realize low quality). Therefore

$$W[(0,0)|(0,0),(\frac{1}{2},0)] = p\theta + 1 - p^2.$$

This implies that $\Delta W_{10}(\frac{1}{2}) > 0$ is equivalent to $\theta > \frac{1-p-p^2}{1-2p}$.
Now we show that for all ϵ_1 such that patients focus on attribute 2 for $(\epsilon_1, 0)$, the welfare difference $\Delta W_{10}(\epsilon_1)$ decreases in ϵ_1, i.e. $\frac{\partial}{\partial \epsilon_1}\Delta W_{10}(\epsilon_1) < 0$. This then implies that $\Delta W_{10}(\epsilon_1) > 0$ for all ϵ_1 such that patients on attribute 2 for $(\epsilon_1, 0)$. The intuition of $\Delta W_{10}(\epsilon_1)$ decreasing in ϵ_1 is as follows: An improvement of the signal quality in the first attribute has a larger effect on expected utility if there is no signal in the second attribute ($\epsilon_2 = \frac{1}{2}$) compared to a precise signal ($\epsilon_2 = 0$). Thus, the welfare difference increases when ϵ_1 decreases.

For explicit calculation we calculate the partial derivative of the expected utilities separately. First, consider $W[(1,1)|(1,1),(\epsilon_1,\frac{1}{2})]$. Signals in the second attribute have no value for the patient. As quality is realized independently for both attributes the patient's expected utility in the second attribute is p. For the first attribute there are four different combinations of quality realization of the two providers. The patient faces high quality in the first attribute if both providers realize high quality (occurs with probability $(1-p)^2$) or if one of the providers realizes high quality and the other one standard quality (occurs with probability $2(1-p)p$) and the patient chooses correctly the provider with the high quality realization (which she does with probability $(1-\epsilon_1)$).[23] Thus, for the expected utility the following holds

$$W[(1,1)|(1,1),(\epsilon_1,\frac{1}{2})] = [2(1-\epsilon_1)(1-p)p + (1-p)^2]\theta + p.$$

[23]If A realizes h and B realizes l, A is chosen with probability $\frac{1}{2}$ if both send the same signal and with probability 1 if A sends h and B sends l. The overall probability that A is chosen is the $2\frac{1}{2}\epsilon_1(1-\epsilon_1) + (1-\epsilon_1)^2 = 1-\epsilon_1$.

Second, consider $W[(0,0)|(0,0),(\epsilon_1,0)]$. For this we consider all possible realizations of quality in the second attribute separately. $q_2 = (h,l)$ or $q_2 = (l,h)$ is realized with probability $2p(1-p)$. In both cases the signal of attribute 1 is irrelevant as the patient focuses on attribute 2 and has a precise signal in attribute 2. Thus the expected utility given realizations $q_2 = (h,l)$ or $q_2 = (l,h)$ is $\theta p + 1$ as utility in the first attribute is realized independent of quality in the second attribute.

If $q_2 = (h,h)$ or $q_2 = (l,l)$ is realized the selection of the provider is only based on the signal in the first attribute. If $q_1 = (h,l)$ or $q_1 = (l,h)$ high quality is selected with probability $(1-\epsilon_1)$. For $q_1 = (h,h)$ the patient selects high quality in attribute 1 with probability 1 and for $q_1 = (l,l)$ standard quality is selected.

Consolidation of those considerations gives

$$\begin{aligned} W[(0,0)|(0,0),(\epsilon_1,0)] &= 2(1-p)p(\theta p+1) && (3.18)\\ &+ (1-p)^2(\theta(p^2+2(1-p)p(1-\epsilon_1))+1) && (3.19)\\ &+ p^2(\theta(p^2+2(1-p)p(1-\epsilon_1)) && (3.20) \end{aligned}$$

where the first term represents expected utility of the patient if $q_2 = (h,l)$ or $q_2 = (l,h)$ is realized, the second if $q_2 = (h,h)$ is realized and the third if $q_2 = (l,l)$ is realized.
Now we can calculate $\frac{\partial}{\partial \epsilon_1}\Delta W_{10}(\epsilon_1)$ as

$$\frac{\partial}{\partial \epsilon_1}\Delta W_{10}(\epsilon_1) = -2(1-p)p + 2(1-p)^3 p + 2(1-p)p^3$$

This is always negative as $-2(1-p)p + 2(1-p)^3 p + 2(1-p)p^3 < 0$ is equivalent to $p - 1 < 0$ which always holds. Therefore, we showed that $\Delta W_{10}(\epsilon_1)$ decreases in ϵ_1.

Proof of Lemma 3.3

For the first part, consider ϵ such that patients focus on attribute 2 when they receive informative signals in both attributes. To show that in equilibrium providers never disclose information on both attributes, assume to the contrary that an equilibrium exists with reporting in both attributes by both providers. Both providers then invest in attribute 2. We show that if one provider deviates by disclosing information only in attribute 2 and investing in attribute 2, he is selected with a probability higher than $\frac{1}{2}$. This makes the deviation profitable.

Assume that provider A reports only in attribute 2 while B reports in both attributes (which implies that for both the belief is investing in attribute 2). A

either sends signal h or signal l in attribute 2, B either sends ll, hl, lh or hh. A is selected when signaling h in attribute 2 while B sends lh, ll or hl (the last two are due to focusing on 2). Furthermore, if A signals l and B signals ll, A is selected as well. Summing up the probabilities with which the signals are sent, provider A is selected with probability $(1-e_2)-(1-e_2)^2e_1+e_2^2(1-e_1)$ (with $e_i=\epsilon_i(1-p)+(1-\epsilon_i)p$). The term is strictly decreasing in e_1. e_1 is always smaller or equal to $\frac{1}{2}$ and the term is $\frac{1}{2}$ for $e_1=\frac{1}{2}$. Therefore, the probability that A is selected is larger than $\frac{1}{2}$.[24]

If ϵ is such that $\epsilon_1<\epsilon_2$ and it can be shown with the same arguments that if one provider is reporting only in attribute 1 and investing in attribute 1 and the other reports in both attributes and invests in attribute 1, the first is selected with a higher probability than the latter.

Proof of Proposition 3.5

(i) Assume that $\epsilon_1<\epsilon_2$ and both providers invest in 1 and report only in 1, patients have corresponding beliefs. We show that none of the providers has an incentive to deviate.

In the Lemma above we already showed that nobody has an incentive to deviate to report in both attributes. Furthermore, reporting only in attribute 2 and investing in attribute 2 is not a profitable deviation. This is because if A reports and invests in 2 and B reports and invest in 1, B is selected whenever sending signal h. This is because $\epsilon_1<\epsilon_2$ and $\theta>1$. However, he sends h with probability $(1-e_1)$ which is greater than $\frac{1}{2}$ (with $e_i=\epsilon_i(1-p)+(1-\epsilon_i)p$). The same argument holds, if A does not send any signal and B sends a signal only in attribute 1. Then, B is selected as well whenever sending h which occurs with a probability $\frac{1}{2}$. Therefore, there is no profitable deviation if both invest and report in attribute 1 which shows that this is an equilibrium.

It remains to show that this equilibrium is unique. Assume another equilibrium exists. If there is one provider that is selected with probability smaller than $\frac{1}{2}$ he has an incentive to deviate by copying the other provider's strategy. Therefore, in equilibrium both have to be selected with probability $\frac{1}{2}$. However, at least one of the providers, say B, necessarily has another strategy than investing in 1 and reporting in 1 (as we consider an equilibrium different to the one where both invest in 1 and report in 1). Then, due to the considerations above, A is selected with

[24]Except for $\epsilon_1=\frac{1}{2}$, but then there is no point in deciding about reporting in both attribute as there is anyway no signal to report in attribute 1.

probability greater than $\frac{1}{2}$ when investing in 1 and reporting in 1 and therefore has an incentive to deviate. This shows the uniqueness.

(ii) Consider any ϵ such that patients focus on attribute 2 when receiving both signals. Once $\theta > \theta^c$, $(h \cdot |1) \succ (\cdot h|0)$ holds (see considerations previous to the Proposition).

$(h \cdot |1) \succ (\cdot h|0)$ implies that reporting only in attribute 1 and investing in this attribute yields a selection probability greater than $\frac{1}{2}$ if the other provider reports only in attribute 2 and invests in 2. The same holds if the other provider does not report since then the one reporting and investing in 1 is as well always selected when signaling h in attribute 1. If, furthermore, it holds that if the other provider reports in both attributes and invests in attribute 2, investing in 1 and reporting in 1 yields a selection probability greater than $\frac{1}{2}$, we showed that investing and reporting in 1 is a PBE.

To show that there exists ϵ such that this holds, assume that provider A invests and reports only in attribute 1, and provider B reports in both attributes and invests in 2. We want to know for which ϵ the probability that A is selected is greater than $\frac{1}{2}$.

Note that A is always selected when sending h in attribute 1 and, on the same time, B either sends ll, hl or lh. For $\epsilon = (\frac{1}{2}, 0)$ provider A is also selected when B sends hh and A send h in attribute 1. Therefore, for all ϵ close enough to $\epsilon = (\frac{1}{2}, 0)$ provider A is selected with probability greater than $\frac{1}{2}$. Note that there are several other ϵ for which this holds. For instance, once $\epsilon_1 > p$, provider A also is selected when sending l in attribute 1 and provider B sends hl or ll. Then, once $p > \frac{1}{3}$ the total probability that A is selected is greater than $\frac{1}{2}$ which can be shown by explicit calculation.

It remains to show that investing and reporting only in 1 is a unique PBE. The arguments for this are exactly the same we saw in (i) for uniqueness.

(iii) First, we show that if $\theta < \theta^c$ we can choose $\epsilon = (\epsilon_1, \epsilon_2)$ such that $(\cdot h|0) \succ (h \cdot |1)$. Second, we show that $(\cdot h|0) \succ (h \cdot |1)$ is sufficient such that reporting only on attribute 2 and investing in attribute 2 with corresponding beliefs is a PBE. The uniqueness of the equilibrium then again follows by the same arguments as seen in (i).

1. Choice of ϵ: For $(\cdot h|0) \succ (h \cdot |1)$ it holds that

$$(\cdot h|0) \succ (h \cdot |1) \Leftrightarrow p\theta + f(1-\epsilon_2, 1-p) > f(1-\epsilon_1, p)\theta + (1-p).$$

The left hand side is decreasing in ϵ_2 and the right hand side is decreasing in ϵ_1. So the error for which the inequality is the easiest to fulfill is $\epsilon = (\frac{1}{2}, 0)$. For this error the inequality transfers to $\theta < \frac{1-p}{1-2p}$. Thus, only if $\theta < \frac{1-p}{1-2p} = \theta^c$ there exists an $\epsilon = (\epsilon_1, \epsilon_2)$ such that $(\cdot h|0) \succ (h \cdot |1)$. We just showed that at least for $\epsilon = (\frac{1}{2}, 0)$ it is the case which implies that there exists a neighborhood of $\epsilon = (\frac{1}{2}, 0)$ such that it holds for all ϵ in this neighborhood.

2. Investing and reporting only in attribute 2 with corresponding beliefs form a PBE. Consider ϵ such that $(\cdot h|0) \succ (h \cdot |1)$ holds and assume that both providers invest in attribute 2 and report only in attribute 2. Then both providers are selected with probability $\frac{1}{2}$. In the following we show that for any provider there is no incentive to deviate.

Assume that provider A deviates by not reporting in 2 but only in 1. If A discloses information on 1 and B on 2 then by $(\cdot h|0) \succ (h \cdot |1)$, B wins whenever generating a signal h in the second attribute the probability of which is larger than $\frac{1}{2}$ since B invests in 2. Therefore, provider A does not have any incentive to deviate to reporting in 1.
Now assume that provider A deviates by reporting in both signals and investing in 2. Again, B wins whenever generating signal h in the second attribute - except for A generating hh. On the other hand, B also is selected when generating l in the second attribute and A generates ll. Thus, B wins with probability $(1 - e_2) - (1 - e_2)^2 e_1 + e_2^2(1 - e_1)$. Here $e_i = \epsilon_i(1 - p) + (1 - \epsilon_i)p$ is the probability that an l signal is generated if the investment is in attribute i. The term decreases in e_1. Inserting $e_1 = \frac{1}{2}$ then shows that B wins with at least a probability of $(1 - e_2) - (1 - e_2)^2 \frac{1}{2} + e_2^2 \frac{1}{2} = \frac{1}{2}$. Therefore, A has no incentive to deviate.
Finally, assume that provider A deviates by not reporting at all. Then again, B wins whenever B sends signal h in the second attribute by $(\cdot h|0) \succ (h \cdot |1)$. Therefore, A has no incentive to deviate.

Proof of Proposition 3.6

The proof combines the results of Proposition 3.5 and the welfare discussion.
For part (i) note that it is optimal if signals in both attributes are reported (and with it providers then invest in 1). Voluntary reporting leads to reporting only in attribute 1.

To discuss parts (ii) and (iii) we only consider signal errors ϵ such that if receiving both signals, patients focus on attribute 2. Optimal reporting is then such that it induces that in equilibrium either both providers invest in attribute 1 and report only in attribute 1, or both providers invest in attribute 2 and report in both attributes. The first is desired if

$$W[(1,1)|(1,1),(\epsilon_1,\frac{1}{2})] > W[(0,0)|(0,0),(\epsilon_1,\epsilon_2)],$$

the latter if the reverse holds.

First, consider $\theta > \theta^c$ and ϵ such that in the unique PBE there is reporting only in attribute 1. From Proposition 3.5 we already know that this is the case if ϵ_1 is high enough and ϵ_2 is low enough since it holds for $\epsilon = (\frac{1}{2}, 0)$ (furthermore, it holds for all ϵ with focusing on 2 as long as $\epsilon_1 > p$ and $p > \frac{1}{3}$). $\theta > \theta^c$ implies $\theta > \underline{\theta}$ and therefore the optimal policy has to induce an equilibrium where both providers invest in 1 and report only on 1. Thus, voluntary reporting is already optimal, while any policy mandating reporting in attribute 2 is not optimal. Mandating reporting only in 1 or banning reporting in 2 yields the same outcome.

Second, consider $\theta < \theta^c$ and ϵ such that disclosing only in attribute 2 is the unique PBE. Again, by the proposition above, this holds if ϵ_1 is high enough and ϵ_2 is low enough. For $\hat{\theta}(\epsilon) < \theta < \theta^c$, it is desirable that signals are sent only in attribute 1. This can be achieved by banning reporting on attribute 2. For ϵ_1 high enough and ϵ_2 low enough, mandatory reporting in 1 is not an optimal policy since then providers would additionally report about attribute 2. However, mandatory reporting in 1 yields higher welfare than voluntary reporting since voluntary reporting in both attributes leads to reporting only in 2 in equilibrium. Banning reporting in 2 leads in equilibrium to only reporting in 1, therefore it is optimal.

For $\theta < \hat{\theta}(\epsilon)$ it is desirable that information about both attributes is available. For ϵ_1 high enough and ϵ_2 low enough, mandating reporting in attribute 1 is already an optimal policy since providers voluntarily report about attribute 2. In this case, banning reporting in attribute 2 is not optimal. Banning reporting in 2 might even decrease welfare compared to voluntary reporting in both attributes. This occurs whenever voluntary reporting yields reporting only in attribute 2 and, on the same time, reporting only in attribute 1 is associated with lower welfare than reporting only in attribute 2. This occurs if θ is close enough to 1 because then, the better selection effect in the second attribute dominates any potentially better resource allocation effect such that receiving information only on attribute 2 and investment in 2 yields higher welfare than receiving information only about attribute 1 but providers invest in attribute 1.

Proof of Proposition 3.7

Fix p, θ and ϵ as considered in the proposition. We are interested in the winning probability of provider A if B invests in 1 and A invests in attribute 2 when investments are observable (i.e. the patient has also the corresponding beliefs). If the probability of A winning is larger than one half, it is a strict dominant strategy to invest in attribute 2. If it is smaller than one half it is a strict dominant strategy to invest in attribute 1.

To assess the winning probabilities we explicitly consider for which signal combinations A wins. If B sends a signal with $s_2 = l$ and A sends a signal with $s_2 = h$ (which occurs with probability $(1 - e_2)^2$) the patient selects provider A as she strongly focuses on attribute 2. The only other cases where A might win are the signal combinations $(s^A, s^B) = (hh, lh)$ and $(s^A, s^B) = (hl, ll)$ (whether or not A is selected depends again on the parameters). In all other cases B is selected. This follows by the fact that if the same signals are generated provider B is selected and all other remaining signal combinations are implied either by strong focusing or by B winning for the same signals.
Therefore, provider A is selected at least with probability $(1 - e_2)^2$ and at most with probability $(1 - e_2)^2 + 2e_1^2 e_2(1 - e_2)$.

Thus, if $(1 - e_2)^2 > \frac{1}{2}$ investing in attribute 2 is a strictly dominant strategy which holds for all $e_2 < 1 - \sqrt{\frac{1}{2}}$.

If $(1 - e_2)^2 + 2e_1^2 e_2(1 - e_2) < \frac{1}{2}$ investing in attribute 1 is a strictly dominant strategy which is equivalent to $e_2 > \frac{3-\sqrt{5}}{2}$.

BIBLIOGRAPHY

Abdulkadiroglu, A., P. Pathak, A. Roth, and T. Sonmez (2006). Changing the Boston school choice mechanism. *NBER working paper no. 11965*.

Abdulkadiroğlu, A., P. A. Pathak, and A. E. Roth (2005). The New York City high school match. *American Economic Review*, 364–367.

Abdulkadiroğlu, A. and T. Sönmez (1999). House allocation with existing tenants. *Journal of Economic Theory 88*(2), 233–260.

Abdulkadiroğlu, A. and T. Sönmez (2003). School choice: A mechanism design approach. *American Economic Review 93*(3), 729–747.

Ambuehl, S. (2015). An offer you can't refuse. *Unpublished manuscript, Stanford University*.

Ambuehl, S., M. Niederle, and A. E. Roth (2015). More money, more problems? can high pay be coercive and repugnant? *American Economic Review 105*(5), 357–60.

American Medical Association Council on Ethical and Judicial Affairs (1995). Ethical considerations in the allocation of organs and other scarce medical resources among patients. CEJA Report K-A-93.

Arneill, A. B. and A. S. Devlin (2002). Perceived quality of care: The influence of the waiting room environment. *Journal of Environmental Psychology 22*(4), 345–360.

Arztnavigator (2015). AOK-Arztnavigator. Available at https://weisse-liste.arzt-versichertenbefragung.aok-arztnavi.de/, accessed on 12/17/2015.

Atkinson, A. B. (1973). How progressive should income tax be. *Essays in Modern Economics*, 90–109.

Baisa, B. (2013). Auction design without quasilinear preferences. *SSRN working paper no. 2311753*.

Bar-Isaac, H., G. Caruana, and V. Cuñat (2012). Information gathering externalities for a multi-attribute good. *Journal of Industrial Economics 60*(1), 162–185.

BBC (2013). German organ donor scandal doctor goes on trial. Available at http://www.bbc.com/news/world-europe-23753872, accessed on 12/17/2015.

Black, S. (1999). Do better schools matter? Parental valuation of elementary education. *Quarterly Journal of Economics 114*(2), 577.

Board, O. (2009). Competition and disclosure. *Journal of Industrial Economics 57*(1), 197–213.

Bordalo, P., N. Gennaioli, and A. Shleifer (2013). Salience and consumer choice. *Journal of Political Economy 121*(5), 803–843.

Che, Y.-K., I. Gale, and J. Kim (2013). Assigning resources to budget-constrained agents. *Review of Economic Studies*, 73–107.

CNN (2009). Did Steve Jobs' money buy him a faster liver transplant? Available at http://edition.cnn.com/2009/HEALTH/06/24/liver.transplant.priority.lists/, accessed on 12/17/2015.

Condorelli, D. (2013). Market and non-market mechanisms for the optimal allocation of scarce resources. *Games and Economic Behavior 82*, 582–591.

Dafny, L. and D. Dranove (2008). Do report cards tell consumers anything they don't already know? The case of medicare HMOs. *RAND Journal of Economics 39*(3), 790–821.

Dranove, D. (2000). Health care markes, regulators, and certifiers. In *Handbook of Health Economics*, Volume 2. North-Holland.

Dranove, D., D. Kessler, M. McClellan, and M. Satterthwaite (2003). Is more information better? The effects of report cards on health care providers. *Journal of Political Economy 111*(3), 555–588.

Dranove, D. and M. A. Satterthwaite (1992). Monopolistic competition when price and quality are imperfectly observable. *RAND Journal of Economics*, 518–534.

Feng Lu, S. (2012). Multitasking, information disclosure, and product quality: Evidence from nursing homes. *Journal of Economics & Management Strategy 21*(3), 673–705.

Fornara, F., M. Bonaiuto, and M. Bonnes (2006, December). Perceived hospital environment quality indicators: A study of orthopaedic units. *Journal of Environmental Psychology 26*(4), 321–334.

Frey, B. S. and F. Oberholzer-Gee (1997). The cost of price incentives: An empirical analysis of motivation crowding-out. *American Economic Review*, 746–755.

Frey, B. S. and W. W. Pommerehne (1993). On the fairness of pricing—an empirical survey among the general population. *Journal of Economic Behavior & Organization 20*(3), 295–307.

Gale, D. and L. S. Shapley (1962). College admissions and the stability of marriage. *American Mathematical Monthly*, 9–15.

Garratt, R. and M. Pycia (2014). Efficient bilateral trade. *SSRN working paper no. 2444279*.

General Comment No. 14 (2000). *Commitee on Economic, Social and Cultural Rights*.

Gibbard, A. (1973). Manipulation of voting schemes: a general result. *Econometrica*, 587–601.

Gill, J., J. Dong, and J. Gill (2014). Population income and longitudinal trends in living kidney donation in the united states. *Journal of the American Society of Nephrology 58*(58), 0–75.

Glazer, J. and T. G. McGuire (2006, March). Optimal quality reporting in markets for health plans. *Journal of Health Economics 25*(2), 295–310.

Gneezy, U. and A. Rustichini (2000). A fine is a price. *Journal of Legal Studies 29*, 1–18.

Goldman, D. P. and J. A. Romley (2008). Hospitals as hotels. *NBER working paper no. 14619*.

Gravelle, H. and P. Sivey (2010). Imperfect information in a quality-competitive hospital market. *Journal of Health Economics 29*(4), 524–535.

Hellwig, M. F. (2005). A utilitarian approach to the provision and pricing of excludable public goods. *Journal of Public Economics 89*(11), 1981–2003.

Hellwig, M. F. (2007). The undesirability of randomized income taxation under decreasing risk aversion. *Journal of Public Economics 91*(3), 791–816.

Hellwig, M. F. (2010). Utilitarian mechanism design for an excludable public good. *Economic Theory 44*(3), 361–397.

Holmstrom, B. and P. Milgrom (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics and Organization 7*, 24.

Huesmann, K. and W. Mimra (2015). Quality provision and reporting when health care services are multi-dimensional and quality signals imperfect. *CER-ETH working paper no. 15/221*.

Huesmann, K. and A. Wambach (2015). Constraints on matching markets based on moral concerns. *CESifo working paper no. 5356*.

Hylland, A. and R. Zeckhauser (1979). The efficient allocation of individuals to positions. *Journal of Political Economy*, 293–314.

Iezzoni, L. (1997). The risks of risk adjustment. *Journal of the American Medical Association 278*(19), 1600–1607.

Kaarboe, O. and L. Siciliani (2011). Multi-tasking, quality and pay for performance. *Health Economics 20*(2), 225–238.

Kahneman, D., J. L. Knetsch, and R. Thaler (1986). Fairness as a constraint on profit seeking: Entitlements in the market. *American Economic Review*, 728–741.

Kimball, M. S. (1990). Precautionary saving in the small and in the large. *Econometrica*, 53–73.

Knetsch, J. L. and J. A. Sinden (1984). Willingness to pay and compensation demanded: Experimental evidence of an unexpected disparity in measures of value. *The Quarterly Journal of Economics*, 507–521.

Koszegi, B. and A. Szeidl (2013). A model of focusing in economic choice. *Quarterly Journal of Economics 128*(1), 53–104.

Ma, C.-t. A. and H. Y. Mak (2014). Public report, price, and quality. *Journal of Economics & Management Strategy 23*(2), 443–464.

Maskin, E. and J. Riley (1984). Optimal auctions with risk averse buyers. *Econometrica*, 1473–1518.

Mellström, C. and M. Johannesson (2008). Crowding out in blood donation: was Titmuss right? *Journal of the European Economic Association 6*(4), 845–863.

Myerson, R. B. and M. A. Satterthwaite (1983). Efficient mechanisms for bilateral trading. *Journal of Economic Theory 29*(2), 265–281.

New York Times (2011). Does school choice improve education? Available at http://www.nytimes.com/2011/12/12/opinion/does-school-choice-improve-education.html, accessed on 12/17/2015.

Richard, T. (1970). *The gift relationship: From human blood to social policy.* Allen and Unwin London.

Roth, A. E. (2007). Repugnance as a constraint on markets. *Journal of Economic Perspectives 21*(3), 37–58.

Roth, A. E. (2008). What have we learned from market design? *The Economic Journal 118*(527), 285–310.

Roth, A. E. (2015). *Who Gets What and Why - The New Economics of Matchmaking and Market Design.* Eamon Dolan / Houghton Mifflin Harcourt.

Roth, A. E., T. Sönmez, and M. Ünver (2004). Kidney exchange. *Quarterly Journal of Economics*, 457–488.

Sandel, M. (2012). *What money can't buy: the moral limits of markets.* Allen Lane, Penguin Group, London.

Satterthwaite, M. and H. Sonnenschein (1981). Strategy-proof allocation mechanisms at differentiable points. *Review of Economic Studies 48*(4), 587–597.

Satterthwaite, M. A. (1975). Strategy-proofness and arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory 10*(2), 187–217.

Satz, D. (2010). *Why some things should not be for sale: the moral limits of markets.* Oxford University Press, New York.

Schummer, J. (2000a). Eliciting preferences to assign positions and compensation. *Games and Economic Behavior 30*(2), 293–318.

Schummer, J. (2000b). Manipulation through bribes. *Journal of Economic Theory 91*(2), 180–198.

Shelegia, S. (2012). Imperfect information in a quality-competitive hospital market. A comment on Gravelle and Sivey. *MPRA working paper no. 42121*.

Sun, M. (2011). Disclosing multiple product attributes. *Journal of Economics & Management Strategy 20*(1), 195–224.

The Conversation (2014). How Germany managed to abolish university tuition fees. Available at http://theconversation.com/how-germany-managed-to-abolish-university-tuition-fees-32529, accessed on 12/17/2015.

Thomson, W. (2011). *Fair allocation rules*, Volume 2. Elsevier, North-Holland, Amsterdam, New York.

Thomson, W. (2014). Non-bossiness. *RCER working paper no. 586*.

Universal Declaration of Human Rights (1948). *UN General Assembly*.

Werner, R. M., R. T. Konetzka, and G. B. Kruse (2009). Impact of public reporting on unreported quality of care. *Health Services Research 44*(2p1), 379–398.

Zargooshi, J. (2001). Quality of life of Iranian kidney. *Journal of Urology 166*(5), 1790–1799.

CURRICULUM VITAE

Personal Details

Name	Katharina Huesmann (née Busch)
Date of birth	October 16, 1982
Place of birth	Rheine, Germany

Education

2010–today	Ph.D. Student, Cologne Graduate School, University of Cologne, Germany
2002–2007	Diplom in Mathematics, University of Münster, Germany
2002	Abitur, Gymnasium Dionysianum, Rheine, Germany

Professional Experience

2014–today	Research Assistent, Department of Economics, University of Cologne, Germany
2008–2014	Consultant, McKinsey & Comp., Germany (on educational leave for PhD studies)

Köln, December 18, 2015