Constraints on Matching Markets Based on Moral Concerns *

Katharina Huesmann† & Achim Wambach‡

November 2016

Abstract

Many markets ban monetary transfers. Rather than exogenously imposing this constraint, we introduce discrimination-freeness as a desideratum based on egalitarian objectives. Discrimination-freeness requires that an agent's object assignment is independent of his wealth. We show that money cannot be used to Pareto-improve ordinal and money-free assignments without violating discrimination-freeness. Furthermore, if a discrimination-free assignment of objects and money is implementable then the respective object assignment is also implementable without money. Once money can be used outside a market designer's control, further restrictions than only money-freeness might be required to address discrimination concerns.

JEL classification: D47, D63, I00

Keywords: repugnance, inequality, market design, matching markets

*For helpful comments and insightful discussions we thank Sandro Ambuehl, Sophie Bade, Felix Bierbrauer, Philipp Heller, Hartmut Kliemt, Fuhito Kojima, Ulrike Malmendier, Timo Mennle, Al Roth, Alexander Westkamp, seminar participants in Cologne and Berlin, and conference participants at SMYE 2012, Frontiers of Market Design 2012, EEA-ESEM 2012, Summer School on Matching Problems, Markets and Mechanisms 2013, 25th Summer School in Economic Theory, Conference on Economic Design 2015, and ESWC 2015. Financial support from the German Science Foundation (DFG) through the research unit "Design and Behavior", the University of Cologne, and the Cologne Graduate School in Management, Economics and Social Sciences is gratefully acknowledged.

†University of Cologne, huesmann@wiso.uni-koeln.de.

‡Centre for European Economic Research (ZEW) and University of Mannheim, wambach@zew.de.

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 by law. Selling 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 monetary transfers clearly exist in certain markets.[2] According to Satz (2010), *"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."*[3] 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 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 money is available but

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

[2]See, for instance, Kahneman, Knetsch, and Thaler (1986), Frey and Pommerehne (1993), and Roth (2007).

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

wealth-independent access to goods is a desideratum.[4] Thereby, we contribute to an understanding of the link between not using money for the assignment of goods and satisfying wealth-independent access to goods. 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 by some wealth endowment. 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. We first show that any Pareto-improvement of an ordinal object assignment that relies on monetary transfers violates discrimination-freeness. This explains a desire to prohibit ex-post trades of objects based on discrimination concerns. If preferences and wealth are private information, the set of object assignments a market designer can implement (possibly with monetary transfers) without violating discrimination-freeness equals the set of object assignments a market designer can implement if money must not be used. To incorporate any preference information beyond rank order lists requires that the market designer can condition the mechanism on the agents' wealth endowments such that an agent's ex-post wealth is independent of his wealth endowment. Assigning objects without using monetary 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 differences in endowments. Using money-free assignments, however, might not sufficiently address discrimination concerns if money can be used to improve access to a good outside the market designer's control.

Intense discourses about money in various markets reveal the importance

[4]Inequality is not the only argument used by opponents of monetary transfers in certain markets. However, other arguments are not in our focus here. We furthermore do not intend at this point to work on the question on which markets inequality is desired.

of studying motivations behind the desire to ban monetary transfers. In the US, compensation of kidney donors is continuously discussed. Germany faced a back and forth in charging tuition fees at universities.[5] The content of the Universal Declaration of Human Rights supports our conjecture that by requiring discrimination-free access we address a deeper desire than restricting monetary 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.[6] Furthermore, empirical findings suggest that our definition of discrimination-freeness captures settings in which monetary incentives are considered unethical: Ambuehl, Niederle, and Roth (2015) show 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 et al., 2015). To the best of our knowledge, our work 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

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

[6]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 [...]".

each agent and determines monetary transfers.[7] It is called *discrimination-free*, if the object assignment of an agent does not depend on his wealth endowment. While an agent's ranking of objects is assumed to be wealth-independent, his willingness to pay for objects depends on wealth.[8] Thereby, a high willingness to pay for an object can 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. For quasi-linear preferences, the desire to trade between two agents does not depend on wealth and therefore no discrimination concerns occur.[9]

In our analysis, we investigate the implications of requiring discrimination-freeness instead of banning monetary transfers on the design of social choice functions. Can money be used to Pareto-improve money-free mechanisms without violating discrimination-freeness? Can money 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 ordinal and money-free social choice functions cannot be Pareto-improved on by using monetary transfers without violating discrimination-freeness.[10] This is because on the one hand, the amount of

[7]Thereby, we concentrate on deterministic social choice functions. We discuss probabilistic assignments in section 5.

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

[9]Our results are not driven by budget constraints. In Section 5 we discuss that additionally assuming budget constraints or restricting endowments hardly changes our analysis. We also discuss that incorporating budget constraints but assuming a constant marginal utility of money does not yield the same results as our specification of preferences. This is because the amount an agent is willing to accept to give up an object is not affected by the budget.

[10]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 compensating an agent for a worse object becomes larger if the agent gains wealth. On the other hand, the money all other agents are willing to pay for an object improvement is bounded. It particularly implies that allowing any kind of ex-post trades results in violating discrimination-freeness. We show that discrimination-freeness and Pareto-efficiency are not exclusive, but to satisfy both, wealth-dependent monetary transfers are needed.

Second, we consider a setting where types are private information and dominant strategy implementability of the social choice function is required.[11] We show that then 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. Thereby, the toolkit to implement object assignments if money is allowed but discrimination-freeness is desired corresponds to the one available if money must not be used. Again, wealth effects are crucial for the result: the lower an agent's wealth the more important for him is the monetary difference of assignments. 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. It is not Pareto-efficient within the set of all social choice functions (with money).

We find that if wealth endowments are public information, only a social choice function that fully eliminates an agent's potential endowment differences can exploit information beyond his object ranking for his assignment.

Third, even if a mechanism does not use money, discrimination concerns can arise once money could improve access outside the assignment procedure. An example based on Schummer (2000a) are bribing opportunities in the sense

[11]The main intuitions also hold if Bayes-Nash implementation is required. In particular, Corollary 1 remains to hold.

that one agent could pay another agent to misreport his preferences such that both are better off. By Schummer (2000a), for quasilinear preferences, an object assignment is bribe-proof (i.e., no such bribing opportunities exist) if and only if an agent's assignment does not depend on other agents' preferences. The result can be transferred to our setting of nonlinear preferences. The analysis of bribes can be interpreted more generally as using money outside a centralized mechanism to influence one's access to a good. For instance, one might attend a private school instead of a centrally assigned public school. Or, being wealthy allows to buy a house that is in the neighborhood of a popular school and therefore raises the chances to receive a place at this school.

Overall, our results explain the desire to ban ex-post trades and the widespread 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. Our results furthermore raise awareness for mechanisms that seem to be discrimination-free because they do not explicitly involve transfers but that are not because money can be used outside the assignment procedure.

Related work. Our work relates to the literature on repugnance in markets, in particular the works on the desire of third parties to restrict monetary 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 monetary transfers into an economic model. Our definition of discrimination-freeness appears to be in line with what people judge as immoral according to Ambuehl et al. (2015).[12] While we concentrate on the concern for inequality, Ambuehl (2015) studies the concern of coercion in the context of financial incentives. He shows that monetary incentives can impact information demand of consumers.[13]

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 other agents' bundles. 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 endowment changes.

Key for our analysis is the non-linearity of preferences. Here, our model differs from the standard assumption of quasilinear preferences in many economic models. In some articles, the impact of non-linearities in preferences such as budget constraints, risk aversion or wealth effects, to the provision of indivisible goods is analyzed (see, e.g., Baisa, 2013, Che, Gale, and Kim, 2013, Garratt and Pycia, 2014, Maskin and Riley, 1984). Focus of that work, however, is rather on optimal mechanisms than on differences in outcomes depending on endowments.

[12] 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.

[13] There is furthermore 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.

Outlook. The remainder of the paper is organized as follows: in Section 2 we describe the basic model. In Section 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 4 we discuss discrimination-concerns that can arise through external factors even though the mechanism itself is money-free. We consider several extensions in Section 5 and conclude with Section 6.

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.[14]

Payoff Environment. Preferences of each agent i are described by an (indirect) 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 $A \in \mathbb{R}$. $\mathcal{U}$ denotes the set of all utility functions that are twice differentiable in wealth and that satisfy

1. *Strict and wealth independent ranking:* For any $\omega, \omega' \in \Omega$ and $A \in \mathbb{R}$

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

[14] $k \geq n$ ensures that there is an object for each agent. This simplifies the notation. It is straightforward to drop this assumption and adding n objects to Ω that indicate remaining unassigned. The setting and analysis do not change when allowing objects to have copies.

2. *Monotonicity and strict concavity:* For any $u_i \in \mathcal{U}$, $\omega \in \Omega$, and $A \in \mathbb{R}$

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

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

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

According to the first assumption, object rankings are strict and wealth does not influence how an agent ranks the objects.[15] The second assumption ensures that each agent has a finite willingness to pay for any object improvement.[16] The third assumption stated in words 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 takes ω. The compensation to accept for an object impairment thus becomes arbitrarily large for increasing wealth.

An example for a utility functions in $\mathcal{U}$ is $u_i(\omega, A) = v_i(\omega) + h_i(A)$ where $v_i : \Omega \to \mathbb{R}$ and $h_i : \mathbb{R} \to \mathbb{R}$ is twice continuously differentiable 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 that no

[15]Technically, wealth independent object rankings are already implied by the assumptions of continuity in wealth and strict rankings. Due to the importance of unique object rankings, we explicitly state this as an assumption.

[16]For any $u_i \in \mathcal{U}$, $A \in \mathbb{R}$, $\omega, \omega' \in \Omega$ where ω is preferred to ω' there exists a unique $M > 0$ such that $u_i(\omega, A - M) = u_i(\omega', A)$. For a formal proof see the proof of Proposition 1.

budget constraints exist.[17]

Each agent is endowed with an initial wealth $e_i \in \mathbb{R}$. $t_i = (u_i, e_i) \in T = \mathcal{U} \times \mathbb{R}$ denotes the *type* of each agent. 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. By the assumptions above, each $u_i \in \mathcal{U}$ uniquely implies an object ranking r_i. $T(r_i)$ denotes the set of all types that describe the same object ranking $r_i \in R$. R is the set of all possible object rankings. $T(u_i)$ 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. All types $T(u_i)$ agree on some object ranking r_i. Therefore, $T(u_i) \subset T(r_i)$. If two types in $T(u_i)$ disagree on what an object improvement is worth, it is due to heterogeneity in wealth. If two types in $T(r_i)$ disagree on what any object improvement is worth it can have two sources. First, cardinal appreciation might differ due to different utility functions. Second, endowments might differ.

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.

Each type $t_i = (u_i, e_i)$ uniquely defines preferences over individual outcomes. 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

[17] In Section 5 we discuss why this assumption is not critical for our analysis and explain which of our results do not rely on the assumption that the admissible wealth endowments of agents are not bounded from above.

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

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] We allow social choice functions not to be budget-balanced. For instance, money might be raised via taxes to fund resources. 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 focus on deterministic outcomes instead of lotteries over deterministic outcomes.[20]

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. $\varphi = (\sigma, m)$ 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.[21]

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

[20] 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 probabilistic assignments.

[21] For the definition of Pareto-efficiency, we restrict our attention to Pareto-improvements without extending the budget. By allowing to extend the budget Pareto-improvements are straightforward by transferring a positive amount of money to each agent. Therefore, budget-balanced social choice functions are the ones that are of interest when searching for Pareto-improvements.

We say that a social choice function $\varphi = (\sigma, m)$ is *implementable* if there exists a mechanism that has a dominant strategy equilibrium so that, for all type profiles, the equilibrium outcome is the outcome of the social choice function. By the revelation principle, when considering implementable social chioce functions we limit our attention to social choice functions for which 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, for any ordinal φ, $\varphi(t) = \varphi(t')$ if for all agents 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.

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 monetary 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 the factors determing how objects are allocated but does not a priori impose restrictions on monetary transfers.

Definition 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.*

To judge whether or not discrimination-freeness is satisfied it is sufficient to consider the outcome of each agent independent of other agents' outcomes.[22]

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 ω to object ω'. One agent is willing to pay more to receive object ω instead of object ω' than the other one. The willingness to pay is driven by preferences over bundles of objects and wealth as well as by endowments. The same holds for the amount of money the agents are willing to accept to give up the preferred object. A discrimination-free social choice function must not take account of the wealth effect but might regard utility effects.[23] A central question in our following analysis is to what extent discrimination-free social choice functions can use information about preferences to assign the objects.

3.1 Pareto-Efficiency

Markets without restrictions on money allow to transfer utility via money and thereby offer opportunities for Pareto-improvements via trades of objects and

[22] In contrast, classical fairness criteria like *envy-freeness* or *equal treatment of equals* make restrictions on how an agent evaluates another agent's outcome. An alternative approach to address discrimination-freeness in that spirit would be to demand that two agents who only differ according to their wealth endowment receive the same object assignment. However, it would require to consider probabilistic assignments. Our definition allows to concentrate on deterministic assignments with potential tie-breakers being determined ex-ante.

[23] If agents are budget constrained but preferences are quasilinear, wealth also plays a role. However, conclusions differ because wealth only impacts on what an agent is willing to pay for an object and does not impact on what an agent is willing to accept to give up an object. Furthermore, the willingness to pay of an agent is bounded from above. See also the discussion in Section 5.

money. Assume there are two objects ω and ω'. ω is assigned to an agent i, ω' is assigned to an agent j. Both agents prefer ω over ω'. In a world where monetary transfers are banned any assignment of the two objects cannot be Pareto-improved. However, both agents could gain when exchanging objects in turn for money. If the willingness to pay of agent j for swapping objects is higher than the money agent i is willing to accept for a swap, both agents can benefit from trade.

The central question for the following analysis is whether money can be used to realize Pareto-improvements compared to classical money-free matchings without violating discrimination-freeness. The following proposition implies that the answer is no. The intuition using the simple example above is that when increasing agent i's wealth endowment, agent i might not any more be willing to give up the preferred object in turn for an amount of money that agent j is willing to pay.

Proposition 1. *Consider any ordinal object assignment σ that is not Pareto-dominated by any other ordinal object assignment. Then, $\varphi = (\sigma, 0)$ is not a Pareto-efficient social choice function. Any budget-balanced social choice function $\varphi' = (\sigma', m')$ that Pareto-dominates φ discriminates.*

Proof. See appendix. □

By Proposition 1, monetary transfers cannot be used to realize Pareto-improvements of ordinal object assignments without violating discrimination-freeness. Ordinal object assignments thus are at the Pareto-frontier of discrimination-free social choice function whenever the object assignment cannot be Pareto-improved on by swapping objects. The proposition holds independent of the information setting. In particular, *even if* full information about types is available, money does not help to realize Pareto-improvement while respecting discrimination-freeness. Furthermore, allowing any kind of ex-post trades among agents is not discrimation-free. In a setting of incomplete information, the *Serial Dictatorship* mechanism where one agent after the other selects an object is an example for a discrimination-free mechanism that is implementable in dominant strategies and that is at the Pareto-frontier of all discrimination-free mechanisms. However, it is not Pareto-efficient within the set of all social choice function that admit monetary transfers.

A second implication of Proposition 1 is that ordinal object assignments without transfers are not Pareto-efficient. In particular, incentives for ex-post trades of objects and money exist. This raises the question of whether a social choice function can be both discrimination-free and Pareto-efficient. Indeed, efficiency and discrimination-freeness are not exclusive. Key for Proposition 1 is that by $\varphi = (\sigma, 0)$ wealth endowments of agents are not affected. A similar result can be obtained if φ is such that it preserves an agent's wealth status in the sense that an agent's ex-post wealth is unbounded in dependence of his own endowment but bounded in dependence of other agents' endowments. For instance wealth independent monetary transfers satisfy this criterion. On the other hand, by using social choice functions with monetary transfers that impact on the wealth status of agents, Pareto-efficiency can be obtained without using transfers.

As an example for a discriminaton-free and Pareto-efficient social choice function consider one that first redistributes wealth and then selects a welfare maximizing assignment. In a first step the endowment of each agent i is adjusted 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 and discrimination-free since an agent's object assignment is independent of his wealth endowment.[24] A construction of such a social choice function clearly depends on the information structure and requires substantial information about the agents' types. In the following section, we deal with the implications of incomplete information on the set of implementable social choice functions.

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 monetary transfers in a discrimination-free way. In the following we consider a setting of incomplete information and are interested in characteristics of discrimination-free and implementable social choice functions. It allows to deduce whether a market designer that can use monetary transfers but is restricted to discrimination-freeness has more freedom to allocate objects compared to being restricted to not using monetary transfers.

First, we assume that the whole type profile is private information (i.e., both utilities $(u_i)_{i \in N}$ and endowments $(e_i)_{i \in N}$). Later, we assume that endowments $(e_i)_{i \in N}$ are public information while utilities $(u_i)_{i \in N}$ are private information.

Proposition 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,*

[24]The social choice function described is not necessarily budget-balanced. However, if wealth endowments are drawn from a distribution such that the expected total endowment is $\overline{e}$, the mechanism above is budget balanced in expectation if each agent's wealth is adjusted to $\frac{1}{N}\overline{e}$. It is also feasible to construct a mechanism that it is Pareto-efficient and ex-post budget-balanced. The construction idea is the following: Allocate the objects such that for some specific endowment profile no Pareto-improvements are feasible even if monetary transfers are admitted. For any other endowment profile, it is then feasible to redistribute endowments of the agents such that the object allocation cannot be Pareto-improved without violating the budget constraint.

- $\sigma_i(t_i, t_{-i}) = \sigma_i(t'_i, t_{-i})$ *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.[25] □

To get an intuition for the proposition first note that discrimination-freeness and implementability of a social choice function imply that neither the object assignment nor the money assignment can be conditioned on endowments. Assume that the second part of the proposition, i.e., payments are independent of the type, does not hold. In particular, agent i's money assignment for being type t_i is smaller than for being type t'_i. Implementability implies that the respective object assignments need to differ as well. We can then construct an admissible utility function such that the outcome of t_i is the most preferred one for some wealth level and the outcome of t'_i is the most preferred for some other wealth level. This contradicts discrimination-freeness of φ and therefore agent i's payments cannot depend on his type. The restriction on σ (i.e., σ_i can 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 his rank order list for his object allocation contradicts implementability.

By Proposition 2, a market designer cannot exploit an agent's cardinal preference information for his object assignment if implementability and discrimination-freeness are desired. An agent's money assignment might depend on

[25] The proof presented is more complex than needed for the domain $\mathcal{U}$ of utility functions. However, it reveals that Proposition 2 even holds if the domain of utility function is modified such that every agent 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. See also Section 5.

other agents' object rankings but must not depend on any of his own type characteristics. Since any agent i's report therefore does not influence his money assignment m_i, executing only the object assignment but not the money assignment does not impact reporting incentives. Thereby, the toolkit of a market designer to assign objects in a discrimination-free way is restricted to the one of a market designer who must not use monetary transfers. This is formalized in the following corollary.[26]

Corollary 1. *Let $\varphi = (\sigma, m)$ be an implementable and discrimination-free social choice function. Then, $\varphi^0 = (\sigma, 0)$ is implementable and discrimination-free as well.*

The *Serial Dictatorship* mechanism where one agent after the other selects an object and where no money is used is an example for a social choice function that is implementable and discrimination-free. By Proposition 1 it is even at the Pareto-frontier of all discrimination-free social choice functions (that might use money).[27] Inefficiencies of such a mechanism without monetary transfers are therefore 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 Mechanism* proposed by Gale and Shapley (1962) or the *Top Trading Cycles Mechanism* (see, e.g., Abdulkadiroğlu and Sönmez, 2003). The *Top Trading Cycles Mechanism* is also at the Pareto-frontier of money-free mechanisms while the *Deferred-*

[26]Corollary even 1 holds for more general settings. It holds if Bayes-Nash implementability instead of implementation in dominant strategies is required. Furthermore, it also holds if social choice functions might assign probabilistic outcomes (see also Section 5).

[27]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.

Acceptance Mechanism is not.

Availability of Wealth Information. In the following we assume that wealth information is available while preferences over objects and wealth are unknown. Intuitively, this increases the scope for a market designer to use information about preferences. For instance, the market designer now might, in a first step of the mechanism, redistribute wealth. It turns out, that an agent's object assignment can be based on more information than only his object ranking *only if* his ex-post wealth is independent of his initial wealth.

Proposition 3. *Let $\varphi = (\sigma, m)$ be an implementable and discrimination-free social choice function. Wealth endowments $(e_i)_{i \in N}$ are public information. Fix any agent i and $t_{-i} \in T^{n-1}$. Then, either agent i's ex-post wealth $A_i = e_i + m_i(u_i, e_i, t_{-i})$ is constant in e_i for all $u_i \in \mathcal{U}$ or it holds that*

- $\sigma_i(t_i, t_{-i}) = \sigma_i(t_i', t_{-i})$ *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* $t_i, t_i' \in T(e_i)$ *and all endowments* $e_i \in \mathbb{R}$.

Proof. See Appendix.[28] □

The proof uses similar characteristics of preferences as the proof of Proposition 3 does. However, complexity increases compared to Proposition 2 as varying wealth might impact on monetary transfers. Suppose that ex-post wealth of an agent i is not independent of e_i. To show that monetary transfers are constant we assume the contrary and show that it leads to a contradiction. If monetary transfers are not constant, there are two wealth levels e_i and e_i'

[28] In analogy to the proof of Proposition 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 5.

such that agent i's ex-post wealth that is associated with some object ω 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 ω with the associated monetary transfer, and for the other wealth level he prefers another object ω' with the associated monetary transfer that he can reach by varying his report. This contradicts discrimination-freeness.

Note that if wealth information is known to the market designer, this wealth information might be used to design transfers without violating discrimination-freeness. An example are income-dependent taxes (for instance to fund the provision of a good).

To construct the utility function described above, it is needed that for any agent i and any parameter setting, agent i's ex-post wealth is not constant in his endowment. Only if this assumption is violated, a market designer might exploit information beyond rankings. As examples for mechanisms with constant ex-post wealth levels consider those mechanisms that first adjust each agent's wealth level to any predefined wealth level that is independent of his initial wealth (see, e.g., the mechanisms discussed in the previous section). Only if those severe wealth adjustments are performed, the market designer has all the flexibility to assign objects and monetary transfers as if discrimination-freeness is not imposed.

4 Discrimination Through Factors Outside the Mechanism

A wide range of real-world applications illustrates that many money-free mechanisms depend on factors that could be influenced with money. If the market designer cannot prevent such usage of money, wealth might impact on who gets what which in turn gives rise to discrimination concerns. In the following we discuss some examples of such factors.

Bribing. In our setting agents are not allowed to trade objects in turn for money after the assignment. However, a market designer might not be able to control monetary transfers among agents that take place before preferences are reported. In particular, bribing situation as described in Schummer (2000b) might exist: some agent could pay another agent to misreport his preferences such that both agents are better off compared to the assignment under truthful reporting.

For instance, consider the serial dictatorship where one agent after each other picks an object. This mechanism is implementable and discrimination-free. Now assume that the first agent has the same favorite object as the second agent. Once the amount of money the first agent is willing to accept for not picking his favorite object is lower than the willingness to pay of the second agent for receiving the favorite object, it makes both agents better off if the second agent bribes the first agent in order to state wrong preferences.

Therefore, even if a social choice function is discrimination-free, money might be used for bribes and, in turn, it allows to gain advantages through higher wealth. A way to circumvent discrimination concerns arising from bribes is to require that a social choice function is not only discrimination-free but also bribe-proof meaning that no bribing opportunities as described above occur (Schummer, 2000b). By Schummer (2000a), in a setting with quasilinear preferences over bundles of objects and transfers, bribe-proofness requires that each agent's outcome is independent of other agents' preferences.[29] The

[29] In (Schummer, 2000b) it is generalized to a broader class of quasilinear settings.

general idea is transferable to our setting with non-linear preferences. Therefore, if not only implementability and discrimination-freeness is required but also bribe-proofness to ensure that discrimination concerns cannot arise from bribing opportunities, the market designer is restricted beyond using transfers.

Remark 4. *Let φ be any implementable and discrimination-free social choice function. φ is bribe-proof if and only if $\varphi_i(t_i, t_{-i}) = \varphi_i(t_i, t'_{-i})$ for all $t_i \in T$ and $t_{-i}, t'_{-i} \in T^{n-1}$.*[30]

To require that an agent's assignments is independent of other agents' types heavily restricts the information about preferences a market designer can use to assign objects. This is at the expense of efficiency. In particular, if there are exactly as many object as agents, the allocation is essentially constant, i.e., types are irrelevant for the assignment. A simple lottery satisfies this condition. If more objects than agents exist it implies wastefulness: there is a type profile such that an object remains unassigned that is preferred to the assigned object by at least one agent.

The arguments are also transferable to setting with two market sides like patients and donors in the context of kidney donation. Bribes among patient and donor could become a concern.

Investing in priority. Many centralized assignment procedures use priorities as a substitute for preferences of a second side of the market. In school choice problems, for instance, each school ranks the students according to pri-

[30] In the Appendix A.2 we provide further details on bribes in the context of discrimination concerns. It includes the formal proof of the remark. Furthermore, we define *preserving discrimination-freeness under bribes* as an extension of discrimination-freeness to a setting with bribing situations. While bribe-proofness is a sufficient criteria to preserve discrimination-freeness under bribes we show that for nonbossy social choice functions (i.e., an agent cannot influence another agent's assignment without changing his own assignment) it is also necessary.

ority criteria. A priority criteria frequently used is the distance of a student's home to the school. Since those being able to afford high house prices have the choice where to live, students can gain priority at a preferred school by the means of money. Black (1999), for instance, showed that house prices are positively correlated with the quality of the school in the neighborhood. In the context of kidney transplants, organs are typically distributed to patients on a waiting list based on priority measures. Steve Jobs 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.[31] This approach required to be wealthy enough to be able to quickly move to any location, e.g., by private plane.

Investing in priorities can be interpreted as a special case of bribes among two sides of a market. As long as an assignment procedure depends on parameters that can be influenced via costly investments, wealth influences the assignment and discrimination concerns become relevant.

Co-existing private market. In addition to objects assigned via a centralized mechanism, there might exist further objects that are distributed through a private market. Once on the private market objects are distributed via prices, wealth impacts on the access to the objects which in turn gives rise to discrimination concerns. A classical example are private schools that charge admission fees. Interpreting the admission decision of a private school as its preference report illustrates that a co-existing private market is a special case of bribes. As long as a private market co-exists that charges prices or fees, discrimination concerns are not fully addressed.

[31] See 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).

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 applications.

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. For some arguments the ability to pay instead of the willingness to pay becomes crucial. Proposition 1 needs to be adjusted slightly because budget constraints do no longer imply that any social choice function with wealth independent monetary transfers is inefficient, but only implies it for social choice functions without monetary transfers.

Furthermore, considering a model where budget constraints occur but the agents have quasilinear preferences does not imply the same results as we obtained. 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 amount of money an agent is willing to accept to give up an object is independent of his wealth endowment. If a mechanism only uses positive payments, wealth endowments are not relevant any more and thus discrimination concerns do not arise. In particular, Proposition 2 and its Corollary 1 do not hold any more since a market designer might use positive payments as an incentive device to elicit private information about preference intensities. A market designer that can use payments but has to satisfy discrimination-freeness therefore has more possibilities to assign object than a market designer that must not use money.

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 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 we have to construct implementable and discrimination-free social choice functions. A further restriction of $\mathcal{U}$ then strengthens the results. It turned out that the proofs of Propositions 2 and 3 do not require 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 $h_i'' < 0$.[32] For Proposition 2 and Proposition 5 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 $\mathbb{R}$. Assuming some minimum endowment $\underline{e} \in \mathbb{R}$ does not impact on the general analysis since no result relys on the possibility of making an agent arbitrarily poor. Assuming a maximum endowment $\overline{e} \in \mathbb{R}$, impacts on Proposition 1 while it does not impact on the

[32]The condition $\lim_{A\to\infty} h_i'(A) = 0$ is only needed for Proposition 1 and can be dropped for the others.

other propositions. The main step of the proofs for the propositions 2 and 3 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 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 the 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.

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.

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 straightforward how to define discrimination-freeness. Sticking to our definition implies that even rankings must not play a role for the object distribution. An alternative approach is to treat the 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, in that case concerns for segregation rather than concerns for discrimination might become relevant.

Assigning Probability Shares. In our analysis we concentrate on deterministic outcomes and therefore take an ex-post perspective. Extending the model by assigning probability shares of objects might improve ex-ante efficiency since lotteries allow to exploit cardinal information about preferences. The definition of discrimination-freeness can be modified as the probabilistic object assignment of an agent being independent of his endowment. Then, in analogy to Proposition 1, if the object assignment σ is not ex-ante Pareto-dominated by any discrimination-free object assignment σ' there is no budget balanced discrimination-free social choice function that ex-ante Pareto-dominates $\varphi = (\sigma, 0)$. However, the resulting social choice function is still not Pareto-efficient within the set of social choice functions with monetary transfers.

To elicit private 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. Then, in contrast to Proposition 2, the assignment is not necessarily ordinal. However, Corollary 1, remains to hold: a market designer that is allowed to use monetary transfers but has to ensure discrimination-freeness does not have more possibilities to assign the object compared to a market designer that must not use transfers.

6 Conclusion

In this paper, we study the problem of assigning indivisible goods to consumers under the constraint of discrimination-freeness that requires a wealth-independent access to goods. We find that, independent of the information setting, money cannot help to Pareto-improve ordinal object assignment without violating discrimination-freeness. In a private information setting, whenever a mechanism cannot (or does not) fully eliminate potential differences in endowments, discrimination-freeness implies the same restrictions on the assignment of objects as banning monetary transfers does. Thereby, our model explains the severe restrictions on monetary transfers in certain markets like school choice procedures or organ donations based on inequality concerns.

Once an agent's assignment depends on factors that can be influenced with money outside a market designer's control, banning transfers cannot fully address discrimination concerns. For instance, if better schools are allocated in more expensive neighborhoods, living in a rather expensive neighborhood already implies better access to schools (Black, 1999). 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."*[33]

Though we cannot (and do not want to) deduce any advice as to whether

[33] See 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).

or not to ban monetary 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.

There is a branch of questions for further research that help to further differentiate our results. On which markets is discrimination-freeness a desideratum and why? What are trade-offs between discrimination-freeness and efficiency? How do preferences depend on wealth for specific real-world applications? Furthermore, moral concerns beyond discrimination-freenes might be important on certain markets. For instance, slippery-slope effects are often feared in the context of an introduction of monetary transfers. Another concern 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.[34]

References

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.

[34]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.

Baisa, B. (2013). Auction design without quasilinear preferences. *SSRN working paper no. 2311753*.

Black, S. E. (1999). Do better schools matter? Parental valuation of elementary education. *Quarterly Journal of Economics 114*(2), 577.

Che, Y.-K., I. Gale, and J. Kim (2013). Assigning resources to budget-constrained agents. *Review of Economic Studies*, 73–107.

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*.

Gneezy, U. and A. Rustichini (2000). A fine is a price. *Journal of Legal Studies 29*, 1–18.

Hylland, A. and R. Zeckhauser (1979). The efficient allocation of individuals to positions. *Journal of Political Economy*, 293–314.

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.

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.

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.

Sandel, M. J. (2012). *What money can't buy: the moral limits of markets.* Allen Lane, Penguin Group, London.

Satterthwaite, M. A. and H. Sonnenschein (1981). Strategy-proof allocation mechanisms at differentiable points. *Review of Economic Studies 48*(4), 587–597.

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.

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*.

Zargooshi, J. (2001). Quality of life of Iranian kidney. *Journal of Urology 166*(5), 1790–1799.

A Appendix

A.1 Proofs of the Propositions

Proof of Proposition 1

Fix any social choice function $\varphi = (\sigma, 0)$ such that σ is ordinal and not Pareto-dominated by any other ordinal object assignment. We prove Proposition 1 in three 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 some budget-balanced φ' Pareto-dominates 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 $(e_i)_{i\in N}$. We aim to find some $\overline{M} > 0$ such that for every agent i and any two objects ω and ω' with ω being preferred to ω' by agent i, it holds that $u_i(\omega, e_i - \overline{M}) \leq u_i(\omega', e_i)$. Then, agent i is not willing to pay more than

$\overline{M}$ for an improvement from ω' to ω. Since the set of agents and the set of objects is finite, it is sufficient to show that for any agent i preferring object ω over object ω' we can find $M > 0$ such that the inequality above holds. $M = M(i, \omega, \omega')$ might depend on i, ω and ω'. $\overline{M}$ can then be defined as $\overline{M} = \max_{i \in N, \omega, \omega' \in \Omega} M(i, \omega, \omega')$.

For any agent i preferring object ω over object ω' define $M > 0$ as the solution of the equation $u_i(\omega, e_i - M) = u_i(\omega', e_i)$. It remains to show that M exists and that it is well defined. First note, that if such an M exists, it has to be unique since $u_i(\omega, m)$ is strictly increasing in m. To show the existence, we use that $u_i(\omega, e_i) > u_i(\omega', e_i)$. Since $u_i(\omega, A)$ is strictly increasing in A and strictly concave in A, it has to hold that $u_i(\omega, A) \to -\infty$ for $A \to -\infty$. Therefore, for some M it holds that $u_i(\omega, e_i - M) = u_i(\omega, e_i)$.

φ' **discriminates.** Assume that there is a budget-balance social choice function φ' that Pareto-dominates φ. We show that the assumption that φ' is discrimination-free leads to a contradiction. Select some agent i that receives 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 uses 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 ω by φ and to ω' 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 ω' instead of ω by a monetary transfer $M(e_i)$. The money $M(e_i)$ that compensates agent i for receiving ω' instead of ω becomes arbitrarily large for increasing endowments. This is because by the assumption on utility functions $u_i \in \mathcal{U}$ the willingness to accept for receiving ω' instead of ω becomes arbitrarily large for increasing wealth.

The 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 $\varphi = (\sigma, 0)$ 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 with a balanced budget. For this, consider first a type profile $t = (t_i)_{i \in N}$ such that all agents have the same preferences $u \in \mathcal{U}$ and the same endowments $e \in \mathbb{R}$. $u \in \mathcal{U}$ and $e \in \mathbb{R}$ are such that each agent is willing to accept at least some $M > 0$ for any object impairment according to φ. More specifically, choose $u \in \mathcal{U}$ with $u(\hat{\omega}, A) = v(\hat{\omega}) + h(A)$ for $\hat{\omega} \in \Omega, A \in \mathbb{R}$ and some $h(\cdot)$ with $h' > 0$, $h'' < 0$ and $\lim_{A_i \to \infty} h' \to 0$. $v(\cdot)$ is specified such that if wealth A is small enough, an agent with preferences u is willing to accept at least M for any object impairment. Wealth endowments e such that everybody is willing to accept at least M for any object impairment.

Now consider the assignment of objects $\sigma(t)$. Select an agent that did not receive the most preferred object ω. Increasing agent i's endowment does not impact on wealth levels of the other agents. If agent i's endowment is high enough, say e^*, he is willing to pay at least M for any object improvement. All other agents are willing to accept M for any object impairment. Therefore, there are two agents that are both better off by trading objects in turn for money. Define $t^* = (t_i^*)_{i \in N}$ such that all agents except agent i have preferences $u \in \mathcal{U}$ and endowments e while agent i has preferences $u \in \mathcal{U}$ and an endowment e^*. $\varphi(t^*)$ thus can be Pareto-improved and φ is not Pareto-efficient.

Proof of Proposition 2

Fix any agent i and any type t_{-i} of the other agents. We omit t_{-i} in the notation. Suppose that $\varphi = (\sigma, m)$ is discrimination-free and implementable. First, we show that agent i's monetary transfer m_i is independent of his type t_i. Second, we show that his object assignment σ_i only depends on his preferences through his ordinal ranking.

m_i **is independent of** t_i**.** Suppose the contrary, i.e., there exist $t_i = (u_i, e_i)$ and $t'_i = (u'_i, e'_i)$ with $m_i(t'_i) < m_i(t_i)$. Implementability implies that for t_i and t'_i the objects they are assigned to differ. Furthermore, it 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 (ω, m) be the assignment in $\varphi(T)$ with the highest monetary assignment and (ω', m') any other outcome in $\varphi(T)$ with $m' < m$.

We now construct a utility function u^*_i such that for two wealth levels e^1_i and e^2_i agent i's object assignment differs for reporting $t^1_i = (u^*_i, e^1_i)$ and $t^2_i = (u^*_i, e^2_i)$. For any $e^1_i < e^2_i \in \mathbb{R}$, construct u^*_i by $u^*_i(\hat{\omega}, A) = v_i(\hat{\omega}) + h(A)$ for $\hat{\omega} \in \Omega, A \in \mathbb{R}$ with any $h : \mathbb{R} \to \mathbb{R}$ and $h' > 0$ and $h'' < 0$. $v_i : \Omega \to \mathbb{R}$ is such that $v_i(\omega') > v_i(\omega) > v_i(\hat{\omega})$ for all $\hat{\omega} \neq \omega, \omega'$ and

$$h(m + e^2_i) - h(m' + e^2_i) < v_i(\omega') - v_i(\omega) < h(m + e^1_i) - h(m' + e^1_i).$$

By construction, type (u_i^*, e_i^1) prefers outcome (ω, m) to all other outcomes that can be reached. An increase of the endowment from e_i^1 to e_i^2 results in preferring (ω', m') to (ω, m). By implementability of φ, the object assignment of an agent with preferences u_i^* has to depend on his wealth. This contradicts discrimination-freeness and thus completes the first part of the proof.

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' have an incentive to deviate. Therefore, agent i's object assignment only depends on his rank order list of objects.

Proof of Proposition 3

Fix any agent i and any type t_{-i} of the other agents. We omit t_{-i} in the notation. Suppose that $\varphi = (\sigma, m)$ is discrimination-free and implementable. Assume that there exists some $u_i \in \mathcal{U}$ such that ex-post wealth $A_i = e_i + m_i(u_i, e_i)$ is not constant in e_i. It is sufficient to show that m_i is independent of u_i because it implies with the same arguments used for proofing Proposition 2 that σ_i is not sensitive to cardinal information of u_i. We assume that m_i is *not* independent of u_i and show that it results in a contradiction. The main step is to find 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 thus m_i has to be independent of u_i.

Construction of u_i^*. By assumption there exists some $u_i \in \mathcal{U}$ such that ex-post wealth $A_i = e_i + m_i(u_i, e_i)$ is not constant in e_i. Since m_i is not independent of preferences, there exists $e_i \in \mathbb{R}$ such that $m_i(u_i, e_i) \neq m_i(u_i', e_i)$

for some $u_i' \in \mathcal{U}$. Without loss of generality, $m_i(u_i, e_i) < m_i(u_i', e_i)$.[35] Choose e_i' such that $A_i = e_i + m_i(u_i, e_i) \neq e_i' + m_i(u_i, e_i') = A_i'$.

Define the choice set $C_{e_i}(\mathcal{U})$ as set of all bundles of objects and ex-post wealth available to agent i with wealth endowment e_i by varying his report (t_{-i} is still fixed). Formally,

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

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. Define $\omega = \sigma_i(u_i, e_i)$ and $\omega' = \sigma_i(u_i', e_i)$. $\omega \neq \omega'$ 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 (ω, A_i) and $(\omega', A_i + x)$ 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 (ω, A_i') and $(\omega', A_i' + x')$ 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 construct a utility function u_i^* such 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^*. By implementability, φ needs to assign different objects to an agent with preferences u_i^* for wealth endowments e_i and e_i'.

To construct u_i^*, we first consider $x' \leq 0$. Then consider any u_i^* such that ω is the most preferred object and ω' the second most preferred object, and $(\omega', A_i + x)$ is the most preferred bundle in $C_{e_i}(\mathcal{U})$. This is feasible with any utility function of the shape $u_i^*(\hat{\omega}, A) = v_i(\hat{\omega}) + h_i(A)$ for $\hat{\omega} \in \Omega, A \in \mathbb{R}$ and $h_i' > 0$, $h_i'' < 0$. Since ω is preferred over ω' and $x' \leq 0$, it holds that

[35] If $m_i(u_i, e_i) > m_i(u_i', e_i)$ the preferences u_i^* can be constructed following the same idea.

$u_i^*(\omega, A_i') > u_i^*(\omega', A_i' + x')$. Therefore, the most preferred bundle in $C_{e_i'}(\mathcal{U})$ does not entail object ω'. This contradicts discrimination-freeness.

Second, consider $x' > 0$. Again, consider a utility function of the shape $u_i^*(\hat{\omega}, A) = v_i(\hat{\omega}) + h_i(A)$ for $\hat{\omega} \in \Omega, A \in \mathbb{R}$ and $h_i' > 0$, $h_i'' < 0$. 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 $v_i(\cdot)$ such that object ω is the most preferred object and object ω' the second most preferred one.

For $h_i(A_i + x) - h_i(A_i) < h_i(A_i' + x') - h_i(A_i')$ choose $v_i(\omega)$ and $v_i(\omega')$ such that

$$h_i(A_i + x) - h_i(A_i) < v_i(\omega) - v_i(\omega') < h_i(A_i' + x') - h_i(A_i').$$

For all other objects that might be entailed in bundles of $C_{e_i}(\mathcal{U})$ assume that the distance in valuation to objects ω and ω' are large enough, such that those bundles are never preferred bundles in $C_{e_i}(\mathcal{U})$ for u_i^*. Then, $(\omega, A_i + x)$ 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 ω. This contradicts discrimination-freeness.

For $h_i(A_i + x) - h_i(A_i) > h_i(A_i' + x') - h_i(A_i')$ choose $v_i(\omega)$ and $v_i(\omega')$ such that

$$h_i(A_i + x) - h_i(A_i) > v_i(\omega) - v_i(\omega') > h_i(A_i' + x') - h_i(A_i')$$

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 ω and ω' are large enough, such that those bundles are never preferred bundles in $C_{e_i}(\mathcal{U})$ for u_i^*. Then, $(\omega', A_i + x)$ 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 ω'. This contradicts discrimination-freeness.

A.2 Preserving Discrimination-Freeness under Bribes

We first define bribing in the spirit of Schummer (2000b): an agent has an incentive to bribe another agent if paying another agent to state false preferences makes both agents better off.

Definition 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.

To account for consequences of bribes for the assignment of objects and monetary transfers, the definition of discrimination-freeness needs to be adjusted. Our original definition of discrimination-freeness requires that the object an agent receives is independent of his wealth level. If a social choice function is not bribe-proof, many bribing incentives might exist. Who receives which object then depends on which bribes are executed. For any social choice function $\varphi = (\sigma, m)$, any agent i and any type profile $t \in T$ we define $\sigma_i^B(t) \subset \Omega$ as the set of all objects that agent i might receive when being bribed because the respective outcomes all serve him with at least the same utility as the assignment without bribes does. Formally, $\omega \in \sigma_i^B(t)$ 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. If φ is bribe-proof, then $\sigma_i^B(t)$ contains only $\sigma_i(t)$. We say that a social choice function preserves discrimination-freeness under bribes if for any agent i, σ_i^B does not depend on agent i's wealth. Thereby, we reflect a desire to avoid that a change in an agent's wealth influences the set of objects an agent may receive when being bribed. The following definition formalizes the requirement of preserving discrimination-freeness under bribes.

Definition 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 $t_{-i} \in T^{n-1}$,*

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

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 an agent's outcome being independent of other agents' types. Nonbossiness of the social choice function makes bribe-proofness a necessary condition for preserving of discrimination-freeness under bribes. In line with Satterthwaite and Sonnenschein (1981) we call a social choice function φ *nonbossy* if for any agent i, $t_i, t'_i \in T$, and $t_{-i} \in T^{n-1}$, $\varphi_i(t_i, t_{-i}) = \varphi_i(t'_i, t_{-i})$ implies $\varphi(t_i, t_{-i}) = \varphi(t'_i, t_{-i})$. Therefore, an agent cannot change another agent's outcome (by reporting different preferences) without changing his own.[36]

Proposition 5. *Consider an implementable and discrimination-free social choice function φ. φ is bribe-proof if and only if $\varphi_i(t_i, t_{-i}) = \varphi_i(t_i, t'_{-i})$ for all $t_i \in T$ and $t_{-i}, t'_{-i} \in T^{n-1}$.*

[36] It is not focus of this work to discuss the desirability of requiring nonbossiness. See, for instance, Thomson (2014), for a detailed and critical discussion.

Suppose φ is additionally nonbossy. Then, φ preserves discrimination-freeness under bribes if and only if $\varphi_i(t_i, t_{-i}) = \varphi_i(t_i, t'_{-i})$ for all $t_i \in T$ and $t_{-i}, t'_{-i} \in T^{n-1}$.

The equivalence of bribe-proofness and an agent's assignment being independent of other agents' types is closely related to Schummer (2000a) and Schummer (2000b) that consider bribes when preferences are quasilinear. We transfer the general idea to our utility domain with non-linear preferences.[37] Bribe-proofness becomes a necessary condition for preserving discrimination-freeness 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 can be quasi-free because there might be an agent who is indifferent between two reports, but his report influences the outcome of another agent.

Proof of the Proposition. Assume that φ is implementable and discrimination-free.

Bribe-proofness $\Leftrightarrow$ $\boldsymbol{\varphi_i(t_i, t_{-i}) = \varphi_i(t_i, t'_{-i})}$**:** The part "$\Leftarrow$" is straight forward. If an agent's type does not influence another agent's outcome it never

[37] We show the equivalence for social choice functions that are discrimination-free. The results of Schummer (2000a) and Schummer (2000b) for a quasilinear setting suggest that the equivalence holds as well in a non-linear setting if discrimination-freeness is not required. However, we are primarily interested in discrimination-freeness and therefore do not further elaborate on this.

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.

It remains to show that bribe-proofness implies $\varphi_i(t_i, t_{-i}) = \varphi_i(t_i, t'_{-i})$. 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 it holds that $\varphi_i(t_j, t_{-j}) I_i \varphi_i(t'_j, t_{-j})$. Second, we show that it implies $\varphi_i(t_j, t_{-j}) = \varphi_i(t'_j, t_{-j})$. $\varphi_i(t_i, t_{-i}) = \varphi_i(t_i, t'_{-i})$ then follows by induction.

1. To show that bribe-proofness implies $\varphi_i(t_j, t_{-j}) I_i \varphi_i(t'_j, t_{-j})$ assume the contrary, i.e, φ is bribe-proof but there is some $t_{-j} \in T^{n-1}$ fixed (and omitted in the following) 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 agent i has an incentive to bribe another agent.

Continuity of the preferences in money implies the existence of $\delta > 0$ such that $(\sigma_i(t'_j, t_{-j}), m_i(t'_j, t_{-j}) - \delta) P_i \varphi_i(t_j, t_{-j})$ (i would pay δ to change type t_j's report from t_j to t'_j). Consider now a utility function u^*_j that represents the same ordinal ranking as u_j does and a wealth level e^*_j such that

$$(\sigma_j(t'_j, t_{-j}), m_j(t'_j, t_{-j}) + \delta)) P^*_j \varphi_j(t^*_j, t_{-j}) \text{ with } t^*_j = (u^*_j, e^*_j)$$

This construction is feasible for the following reasons. 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 agent j's report (an implication of discrimination-freeness), agent j's object assignment needs to be weakly worse than the one for reporting t^*_j. By exploiting that $m_j(t'_j) + \delta > m_j(t^*_j)$, for any u^*_j such that $u^*_j(\omega, A) = v_j(\omega) + h(A)$ for $\omega \in \Omega, A \in \mathbb{R}$ with $h' > 0$ and $h'' < 0$ it is feasible to choose $v_j(\cdot)$ and e^*_j such that the equation above is satisfied.

By the discussion above, $(\sigma_i(t'_j, t_{-j}), m_i(t'_j, t_{-j}) - \delta) P_i \varphi_i(t_j, t_{-j})$ 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 $\tau < \delta$. This holds because j is indifferent between reporting t_j or t^*_j since both types represent the same ranking. 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.

2. To show that $\varphi_i(t_j, t_{-j}) I_i \varphi_i(t'_j, t_{-j})$ implies $\varphi_i(t_j, t_{-j}) = \varphi_i(t'_j, t_{-j})$ assume again the contrary: For any agents i and j, $t_{-ij} \in T^{n-2}$ fixed (and omitted in the following), 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

$$(\omega, m) = \varphi_i(t_i, t_j) \neq \varphi_i(t_i, t'_j) = (\omega', m').$$

It implies that $\omega \neq \omega'$ and $m \neq m'$ because otherwise, agent i cannot be indifferent. Without loss of generality assume that $m > m'$. 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 $(\omega, m) P^*_i (\omega', m')$. Since φ is implementable and discrimination-free, reporting t_i and reporting t^*_i need to yield the same outcome for agent i. Therefore,

$$\varphi_i(t^*_i, t_j) = (\omega, m) \text{ and } \varphi_i(t^*_i, t'_j) = (\omega', m').$$

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 (ω, m) is strictly preferred over (ω', m').

Nonbossy Social Choice Functions: By the first part of the proposition bribe-proofness is equivalent to $\varphi_i(t_i, t_{-i}) = \varphi_i(t_i, t'_{-i})$ for all $t_i \in T$ and $t_{-i}, t'_{-i} \in T^{n-1}$. Furthermore, if φ is discrimination-free and bribe-proof it 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 object 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. □