Equitable screening*

Filip Tokarski

December 15, 2023

Abstract

I study the problem of a government providing benefits while considering the perceived equity of the resulting allocation. Such concerns are modelled through an equity constraint requiring that equally deserving agents receive equal allocations. I ask what forms of efficiency-enhancing screening are compatible with equity and show that while the government cannot equitably screen with a single instrument (e.g. payments or wait times), combining multiple instruments, which on their own favor different groups, allows it to screen while still producing an equitable allocation. I also show that using only payments to screen can lead to more equitable outcomes than using only wait times.

1 Introduction

In 2017, the French president Emmanuel Macron unveiled a set of environmental policies aimed at reducing carbon emissions. The plan involved a gradual increase in fuel taxes, including a significant hike in diesel and petrol prices the following year. However, the proposed tax increase sparked protests and riots which later expanded into the 'Yellow Vest' movement. The protesters, largely from rural and less affluent areas, objected to the policy on equity grounds. They claimed that households in poor financial standing were disproportionately paying the cost of decarbonization, and that the tax unfairly burdened those who couldn't cut down on driving or switch to greener alternatives.[1] Ultimately, mounting unrest led President Macron to suspend the planned fuel tax hike. Carbon taxes in other countries such as Canada[2] and Australia[3] provoked similar grievances.

This paper concerns problems where the policymaker is bound by such equity considerations. It focuses on the problem of a government allocating goods (or bads) such as vaccines, affordable housing, basic food items, or emission reductions. I model equity concerns using a *merit function* specifying societal perceptions of how entitled each agent is to the allocated good. For instance, in the case of emission reductions, rural households reliant on cars for transportation would be more deserving a right to emit than urban ones. I then introduce an equity constraint requiring that agents with equal merit receive the same allocation.

*I am grateful to Piotr Dworczak, Joey Feffer, Mike Ostrovsky, Sam Wycherley, Paula Onuchic, Annie Liang and Ben Golub for their helpful comments.

[1] https://www.nytimes.com/2018/12/06/world/europe/france-fuel-carbon-tax.html

[2] https://www.theglobeandmail.com/politics/article-federal-carbon-tax-changes-rural-atlantic-canada

[3] https://www.theguardian.com/environment/2014/jul/17/australia-kills-off-carbon-tax

I then ask what policy designs that are compatible with equity. Standard economic logic suggests that public programs can be made more efficient through costly screening—if agents need to pay or wait to get the benefit, only those who need it will do so. Indeed, the literature studying the optimal design of such programs (e.g. Condorelli (2013), Akbarpour et al. (2020)) finds that the optimal mechanism often involves screening recipients, e.g. through payments. For instance, Akbarpour et al. (2020) find that, under fairly general conditions, the government should sell the good (at possibly non-market prices) and then redistribute the revenue to poorer agents.

However, different agents with the same need for the good might find waiting or paying burdensome to different extents. Consider for instance the problem of distributing Covid-19 vaccines studied by Akbarpour et al. (2023). They show that the optimal mechanism combines priorities to vulnerable groups with a market mechanism under which one can pay to be vaccinated early. A poor person with severe health conditions might nevertheless be less willing to pay to get a vaccine early than would a rich person of similar health. Consequently, we might worry that efficiency-enhancing screening could lead to inequitable allocations. The authors themselves acknowledge that such mechanisms might provoke backlash on fairness grounds. Similar concerns have been raised by both academic philosophers (Kass (1997), Walzer (1983)) and the public.[4] Indeed, in many cases governments forgo screening in favor of mechanisms based purely on observables, or lotteries.[5] Examples of goods allocated by lottery include primary and secondary school places (Stone (2008)), public housing (Elster (1989), p.63) and US green cards.

I therefore ask how (if at all) the government can screen agents when allocations are subject to equity constraints. I look at screening using only payments (which are less costly to the rich), only waiting (less costly to the poor) and both of these instruments at once. In the former two cases, equitable screening is impossible. However, when the designer uses both payments and waiting, she has significant freedom to screen despite equity constraints.

This paper relates to work examining how moral sentiments constrain market designers and policymakers. Roth (2007) discusses how the repugnance of certain transactions precludes the use of markets in settings where they would be efficient. The literature following Abdulkadiroğlu and Sönmez (2003) models fairness concerns in matching markets through assigning *priorities* to agents. It then studies matching mechanisms that eliminate *justified envy*—a notion capturing perceived injustice. Finally, Dessein et al. (2023) argue that many US colleges switched to test-optional admissions to reduce public scrutiny of their admission decisions. However, no existing work studies equity constraints in mechanism design

[4]https://www.statnews.com/2020/12/03/how-rich-and-privileged-can-skip-the-line-for-covid19-vaccines

[5]One might argue, along the lines of Weitzman (1977), that governments eschew screening not due to fairness concerns, but because screening is not optimal when the government has distributional concerns. For instance, if a government allocating affordable housing screened using payments, the allocation would be biased towards wealthier agents. This would be suboptimal for a government with redistributive preferences. There are two responses to this argument. First, the government could use a screening instrument other than payments whose cost is *negatively* correlated with wealth; in Subsection 3.2 I study screening with waiting, which plausibly has this feature. Secondly, as observed by Akbarpour et al. (2020), screening with payments is still often optimal even in such settings. This is because the government can generate revenue from selling the good to rich agents and then redistribute it to the poor who value money highly. These observations suggest a deeper reason for the frequent absence of screening in such programs.

problems of the sort this paper considers. In doing so, my paper also relates to the literatures on algorithmic fairness in Computer Science, and on discrimination in Economics, which attempt to conceptualize bias, unfairness and discrimination (see Alves et al. (2023) and Onuchic (2022) for respective surveys). However, in both cases, researchers focus on problems of classification or statistical inference, and hence do not account for the strategic behavior of agents. By contrast, the purpose of this paper is to study mechanisms that are fair after accounting for strategic responses.

The rest of the paper is structured as follows. I first introduce the model of public provision with an equity constraint. Section 3 then discusses the forms of screening that are feasible in this environment. In Section 4 I ask what allocation rules can be equitably implemented if the government also observes the agents' wealth. Section 5 discusses a relaxed version of the equity constraint; Section 6 concludes.

2 Model

The government allocates goods $x \in [\underline{x}, \bar{x}] \subseteq \mathbb{R}$ to agents with types $(\alpha, \beta) \in \Theta \subseteq \mathbb{R}^2$. Θ is open, connected and bounded. β represents an agent's value for the good and α represents her value for money (higher α means the agent is poorer). I consider two screening instruments—payments and wait times. Payments, $p \in \mathbb{R}$, are more burdensome for poorer agents (higher α), while waiting, $q \in \mathbb{R}_+$, is costlier for richer agents. Utility is given by:

$$U[\alpha, \beta; x, p, q] = v(\beta, x) - w(\alpha, p) - z(\alpha, q).$$

Assumption 1. *The following conditions hold on the closure of* Θ*:*

1. $v_\beta > 0$, $v_x > 0$, $w_\alpha > 0$, $w_p > 0$, $z_\alpha < 0$, $z_q > 0$, *where the derivatives are continuous.*
2. $v_{\beta x} > 0$; $\quad w_{\alpha p} > 0$; $\quad z_{\alpha q} < 0$.

I assume that both α and β are private information (I discuss this assumption in Section 4). The government therefore chooses allocations $x : \Theta \to [\underline{x}, \bar{x}] \subseteq \mathbb{R}_+$, $p : \Theta \to \mathbb{R}$, $q : \Theta \to \mathbb{R}_+$ subject and IC and IR constraints (I write $y = (x, p, q)$):

$$\text{for all } (\alpha, \beta) \in \Theta, \quad U[\alpha, \beta; y(\alpha, \beta)] \geq \max_{(\alpha^a, \beta^a) \in \Theta} U[\alpha, \beta; y(\alpha^a, \beta^a)], \tag{IC}$$

$$\text{for all } (\alpha, \beta) \in \Theta, \quad U[\alpha, \beta; y(\alpha, \beta)] \geq v(\beta, x_o) - w(\alpha, 0) - z(\alpha, 0), \tag{IR}$$

where $x_o \in [\underline{x}, \bar{x})$ is the outside option. I call allocations satisfying (IC) and (IR) *implementable*.

The government also faces an equity constraint which I model using an exogenous *merit function* $\eta : \Theta \to \mathbb{R}$. I assume $\eta(\alpha, \beta)$ is differentiable and strictly increasing in both arguments. That is, agents are more entitled to receiving the good if they value it more or if they are poorer (as richer agents can more easily satisfy their needs without government assistance). I denote the extreme values of $\eta(\alpha, \beta)$ on Θ by $\underline{\eta} = \inf_{(\alpha,\beta)\in\Theta} \eta(\alpha, \beta)$ and $\bar{\eta} = \sup_{(\alpha,\beta)\in\Theta} \eta(\alpha, \beta)$. Since Θ is connected and $\eta(\alpha, \beta)$ is continuous, for every $\eta^a \in (\underline{\eta}, \bar{\eta})$

there is some (α^a, β^a) such that $\eta(\alpha^a, \beta^a) = \eta^a$. The equity constraint then says that all agents with the same merit should receive the same allocation:[6]

$$\eta(\alpha^a, \beta^a) = \eta(\alpha^b, \beta^b) \quad \Longrightarrow \quad x(\alpha^a, \beta^a) = x(\alpha^b, \beta^b). \tag{E}$$

I will call any allocation satisfying (E) *equitable*. Note also that an allocation $x(\alpha, \beta)$ is equitable if and only if it can be written in the form $x(\alpha, \beta) = \hat{x}(\eta(\alpha, \beta))$ for some $\hat{x} : (\underline{\eta}, \overline{\eta}) \to [\underline{x}, \overline{x}]$.

While I have not assumed that agents with higher merit have to get higher allocations, this will still be the case for any implementable equitable allocation:

Lemma 1. *If $x(\alpha, \beta) \equiv \hat{x}(\eta(\alpha, \beta))$ is implementable, then $\hat{x}(\cdot)$ is weakly increasing.*

Proof. Any implementable $x(\alpha, \beta)$ has to be implementable on the subset of agents with $\alpha = \alpha^a$, for any α^a. We can write the utility of such agents as $v(\beta, x(\alpha^a, \beta)) - t(\beta)$, where $t(\beta) := w(\alpha^a, p(\alpha^a, \beta)) + z(\alpha^a, q(\alpha^a, \beta))$. This is a one-dimensional quasi-linear screening problem so any implementable allocation has to be weakly increasing in β. Hence, $x(\alpha, \beta)$ must be weakly increasing in β for every α. Finally, an equitable allocation must take the form $\hat{x}(\eta(\alpha, \beta))$. Since η is strictly increasing in β, it follows that $\hat{x}(\cdot)$ has to be increasing. □

3 Implementable equitable allocations

I now ask what forms of screening are compatible with equity. I consider the sets of implementable equitable allocations in three cases: when the government uses only payments to screen, when it uses only waiting, and when it combines both screening instruments.

3.1 Screening with payments

I first show that the government cannot equitably screen using only payments.

Proposition 1. *If $q \equiv 0$, any equitable and implementable $x(\alpha, \beta)$ is the same for all $(\alpha, \beta) \in \Theta$.*

Proof. By Lemma 1, any implementable and equitable allocation $x(\alpha, \beta)$ can be written as $\hat{x}(\eta(\alpha, \beta))$, where $\hat{x}(\cdot)$ is weakly increasing. Note also that $p(\alpha, \beta) \equiv \hat{p}(\eta(\alpha, \beta))$ because identical allocations of x must require identical payments. Moreover, $\hat{p}(\cdot)$ has to be increasing—otherwise one could deviate and receive a weakly greater allocation of x for a strictly smaller p. However, an argument analogous to the proof of Lemma 1 tells us that any implementable $p(\alpha, \beta)$ has to be weakly decreasing in α. Since η is strictly increasing in α, it follows that $\hat{p}(\cdot)$ must be weakly decreasing, and hence constant. Such a payment scheme can only support a constant allocation. □

Intuitively, equity requires that poorer and richer agents of equal merit receive the same allocation, even though the richer agents have higher need. However, these richer agents with higher need have greater willingness to pay for the good (Figure 1). Therefore, any mechanism that sells the good will allocate more to richer agents, and hence violate equity.

[6] My model can capture lotteries when the allocation is binary or the utility for the good is linear: $v(\beta, x) = \beta \cdot x$. In the former case, we can interpret $x \in [0, 1]$ as the probability of being allocated the good.

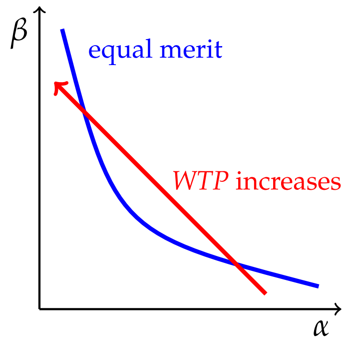

Figure 1

3.2 Screening with waiting

I now show that, generically, the government cannot screen equitably using only waiting. Intuitively, screening with waiting differs from screening with payments in that the allocation it produces is skewed *towards the poor*. That is, between two people with the same β, the poorer one will be more eager to wait to get the good, and hence will receive a higher allocation. Unlike in the case of screening with payments, this direction of bias is consistent with what the merit function requires. It is therefore less clear why screening with waiting is also incompatible with equity. Notice, however, that equity constraints impose requirements not only on the direction of the bias, but also on its exact form—the merit function $\eta(\alpha, \beta)$ specifies exactly the sets of types that have to be treated identically. However, with one screening device only, the designer will generically have 'too few degrees of freedom' to pool agents in this exact way. Proposition 2 formalizes this reasoning.

Definition 1. *A merit function is **generic** if the following conditions fail for every $\eta^a \in (\underline{\eta}, \overline{\eta})$:*

1. *There exist $x^a \in [\underline{x}, \bar{x}]$ and $q^a, k \in \mathbb{R}$ such that for all (α, β) with $\eta(\alpha, \beta) = \eta^a$:*

$$\frac{v_x(\beta, x^a)}{z_q(\alpha, q^a)} = k. \tag{1}$$

2. *There exist $x^a \neq x^b \in [\underline{x}, \bar{x}], q^a, q^b \in \mathbb{R}$ such that for all (α, β) with $\eta(\alpha, \beta) = \eta^a$:*

$$v(\beta, x^a) - z(\alpha, q^a) = v(\beta, x^b) - z(\alpha, q^b). \tag{2}$$

Proposition 2. *Suppose $p \equiv 0$. Then, if the merit function η is generic, any equitable and implementable $x(\alpha, \beta)$ is the same for all $(\alpha, \beta) \in \Theta$.*

While the proof is relegated to the appendix, I provide its key intuition. Notice $q(\alpha, \beta) \equiv \hat{q}(\eta(\alpha, \beta))$ since identical allocations of x have to come with equal wait times. Moreover, note that $\hat{q}(\cdot)$ has to be increasing—otherwise one could deviate and receive weakly more x for a strictly smaller q. Now, consider the case where $\hat{x}(\cdot)$ and $\hat{q}(\cdot)$ are smoothly increasing around some η^a. Then the FOCs of all agents with merit η^a must hold there:

$$\text{for all } (\alpha, \beta): \ \eta(\alpha, \beta) = \eta^a, \quad \frac{v_x(\beta, \hat{x}(\eta^a))}{z_q(\alpha, \hat{q}(\eta^a))} = \frac{\hat{q}'(\eta^a)}{\hat{x}'(\eta^a)}. \tag{3}$$

However, Condition 1 of Definition 1 says that (3) cannot hold for all such agents when $\eta(\alpha, \beta)$ is generic (Figure 2). Hence, agents will have a larger (smaller) *MRS* at that allocation, and will want to mimic agents with a slightly higher (lower) η. Condition 2 relates to a similar requirement for values of $\eta(\alpha, \beta)$ where $\hat{x}(\cdot)$ and $\hat{q}(\cdot)$ jump discontinuously.

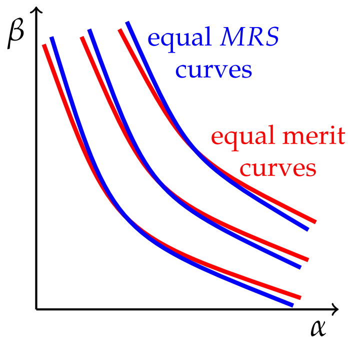

Figure 2

3.3 Screening with payments and waiting

I now let the government use both screening instruments at once. To keep the model tractable, I impose more structure:

$$U[\alpha, \beta; x, p, q] = \beta x - w(\alpha)p - z(\alpha)q.$$

Assumption 2. *The following conditions hold on the closure of* Θ*:*

1. β, $w(\alpha)$, $z(\alpha) > 0$.
2. $w(\alpha)$ *and* $z(\alpha)$ *are twice continuously differentiable.*
3. *The first, second and cross partials of* $\eta(\alpha, \beta)$ *exist and are continuous.*

As it turns out, using both screening devices allows for rich screening without violating the equity constraint. Intuitively, every amount of x can now come with a menu of payment options composed of different amounts of p and q. Since the designer has one screening device preferred by the poor and another one preferred by the rich, she can fine-tune such 'payment menus' to produce precisely the bias in allocation that equity requires. Loosely speaking, being able to compose such menus fixes the problem of 'too few degrees of freedom' we encountered when only wait times were used.

Proposition 3. *Let* $x(\alpha, \beta) \equiv \hat{x}(\eta(\alpha, \beta))$*, where* $\hat{x}(\cdot)$ *is increasing and twice differentiable. Then* $x(\alpha, \beta)$ *is equitable and implementable.*

Proof. To simplify the proof, I first reparametrize types:

$$\kappa = \frac{\beta}{z(\alpha)}, \quad \lambda = \frac{w(\alpha)}{z(\alpha)}, \quad \tilde{\Theta} = \left\{ \left(\frac{\beta}{z(\alpha)}, \frac{w(\alpha)}{z(\alpha)} \right) : (\alpha, \beta) \in \Theta \right\}.$$

Then every (α,β) corresponds to a unique (κ,λ). To see that, take (α^a,β^a), (α^b,β^b) such that:

$$\frac{\beta^a}{z(\alpha^a)} = \frac{\beta^b}{z(\alpha^b)}; \quad \frac{w(\alpha^a)}{z(\alpha^a)} = \frac{w(\alpha^b)}{z(\alpha^b)} \tag{4}$$

Suppose that $\alpha^a > \alpha^b$. Then, since $z_\alpha < 0$ and $w_\alpha > 0$,

$$\frac{w(\alpha^a)}{z(\alpha^a)} > \frac{w(\alpha^b)}{z(\alpha^b)},$$

which contradicts (4), so $\alpha^a = \alpha^b$. Then, the first equation in (4) gives $\beta^a = \beta^b$.[7]

Now, define:

$$\underline{\kappa} = \inf_{(\alpha,\beta)\in\Theta} \frac{\beta}{z(\alpha)}, \quad \bar{\kappa} = \sup_{(\alpha,\beta)\in\Theta} \frac{\beta}{z(\alpha)}, \quad \underline{\lambda} = \inf_{(\alpha,\beta)\in\Theta} \frac{w(\alpha)}{z(\alpha)}, \quad \bar{\lambda} = \sup_{(\alpha,\beta)\in\Theta} \frac{w(\alpha)}{z(\alpha)}.$$

Since Θ is bounded, Assumption 2 guarantees that $\infty > \underline{\lambda}, \underline{\kappa}, \bar{\lambda}, \bar{\kappa} > -\infty$, so $\tilde{\Theta}$ is bounded. Agents' utilities (up to scaling) in the reparametrized model are given by:

$$\tilde{U}[\kappa,\lambda;x,p,q] = \kappa x - \lambda p - q.$$

I will also write the payment rule and the merit function in (κ,λ) space:

$$\tilde{p}\left(\frac{\beta}{z(\alpha)}, \frac{w(\alpha)}{z(\alpha)}\right) \equiv p(\alpha,\beta), \quad \tilde{\eta}\left(\frac{\beta}{z(\alpha)}, \frac{w(\alpha)}{z(\alpha)}\right) \equiv \eta(\alpha,\beta). \tag{5}$$

Since $\eta(\alpha,\beta)$ was strictly increasing in β and α, $\tilde{\eta}(\kappa,\lambda)$ is strictly increasing in κ and λ. Moreover, Assumption 2 ensures that $\tilde{\eta}(\kappa,\lambda)$ has continuous first and second partials on the closure of $\tilde{\Theta}$.

Recall that any allocation of the form $\hat{x}(\tilde{\eta}(\kappa,\lambda))$ is equitable. By Proposition 2 in Rochet (1987), $\hat{x}(\tilde{\eta}(\kappa,\lambda))$ is implementable if there is a convex function $V(\kappa,\lambda) : \tilde{\Theta} \to \mathbb{R}$ such that:[8]

$$\frac{d}{d\kappa} V(\kappa,\lambda) = \hat{x}(\tilde{\eta}(\kappa,\lambda)). \tag{R}$$

Notice that (R) holds for the following $V(\kappa,\lambda)$:

$$V(\kappa,\lambda) = \int_{\underline{\kappa}}^{\kappa} \hat{x}(\tilde{\eta}(\tau,\lambda))\, d\tau + \zeta \int_{\underline{\lambda}}^{\lambda} \tau\, d\tau, \tag{6}$$

where $\zeta \in \mathbb{R}$. I now show this $V(\kappa,\lambda)$ is convex for some ζ. The Hessian of $V(\kappa,\lambda)$ is:

$$H(\kappa,\lambda) = \begin{bmatrix} \hat{x}'(\tilde{\eta}(\kappa,\lambda)) \cdot \tilde{\eta}_\kappa(\kappa,\lambda) & \hat{x}'(\tilde{\eta}(\kappa,\lambda)) \cdot \tilde{\eta}_\lambda(\kappa,\lambda) \\ \hat{x}'(\tilde{\eta}(\kappa,\lambda)) \cdot \tilde{\eta}_\lambda(\kappa,\lambda) & r(\kappa,\lambda) \end{bmatrix},$$

[7] Note this would no longer hold if both screening devices were less costly to the rich (or to the poor).

[8] Rochet (1987) assumes $\tilde{\Theta}$ is convex. However, I can add 'fictitious types' enlarging the type space to $(\underline{\kappa},\bar{\kappa}) \times (\underline{\lambda},\bar{\lambda}) \supseteq \tilde{\Theta}$ and implement an extended allocation on it that coincides with $\hat{x}(\tilde{\eta}(\kappa,\lambda))$ on $\tilde{\Theta}$.

where

$$r(\kappa,\lambda) = \int_{\underline{\kappa}}^{\kappa} \underbrace{\hat{x}''(\tilde{\eta}(\tau,\lambda))\cdot\tilde{\eta}_{\lambda}^{2}(\tau,\lambda) + \hat{x}'(\tilde{\eta}(\tau,\lambda))\cdot\tilde{\eta}_{\lambda\lambda}(\tau,\lambda)}_{:=P(\kappa,\lambda)}\, d\tau + \zeta.$$

It suffices to show that the determinant of H, and $H_{(1,1)}, H_{(2,2)}$ are positive everywhere. $H_{(1,1)}$ is positive since $\tilde{\eta}_{\kappa}(\kappa,\lambda) > 0$ and $\hat{x}(\cdot)$ is increasing by assumption. For the determinant, we want to show that for every $\kappa \in (\underline{\kappa},\bar{\kappa})$, $\lambda \in (\underline{\lambda},\bar{\lambda})$:

$$\hat{x}'(\tilde{\eta}(\kappa,\lambda))\cdot\tilde{\eta}_{\kappa}(\kappa,\lambda)\cdot\ r(\kappa,\lambda)\ -\ [\hat{x}'(\tilde{\eta}(\kappa,\lambda))\cdot\tilde{\eta}_{\lambda}(\kappa,\lambda)]^2 \geq 0. \tag{7}$$

If for some (κ,λ) we have $\hat{x}'(\tilde{\eta}(\kappa,\lambda)) = 0$, the inequality holds, so consider the opposite case. Then $\hat{x}'(\tilde{\eta}(\kappa,\lambda)), \eta_{\kappa}(\kappa,\lambda), \eta_{\lambda}(\kappa,\lambda) > 0$, so (7) is equivalent to:

$$r(\kappa,\lambda) \geq x'(\tilde{\eta}(\kappa,\lambda))\cdot\frac{\tilde{\eta}_{\lambda}(\kappa,\lambda)^2}{\tilde{\eta}_{\kappa}(\kappa,\lambda)},$$

$$\zeta \geq x'(\tilde{\eta}(\kappa,\lambda))\cdot\frac{\tilde{\eta}_{\lambda}(\kappa,\lambda)^2}{\tilde{\eta}_{\kappa}(\kappa,\lambda)} - \int_{\underline{\kappa}}^{\kappa} P(\kappa,\lambda)d\kappa.$$

Notice that it suffices to show that the RHS is uniformly bounded across κ. Recall that the first and second partials of $\tilde{\eta}$ and the first and second derivatives of $\hat{x}$ are continuous on the closure of $\tilde{\Theta}$. Since $\tilde{\Theta}$ is bounded, the first term on the RHS and $P(\kappa,\lambda)$ are uniformly bounded. Since κ is bounded by $\underline{\kappa},\bar{\kappa} \in \mathbb{R}$, the RHS is indeed uniformly bounded. Therefore, we can ensure (7) by choosing ζ high-enough. Moreover, we can choose ζ high enough to make $r(\kappa,\lambda) \geq 0$ everywhere. This also ensures $H_{(2,2)} \geq 0$ everywhere. □

4 Observable wealth

While I assumed that neither need nor wealth are observable, the government usually has some information about them. For instance, in the problem of vaccine allocation, age and medical history are good indicators of need. Similarly, tax data proxies for one's wealth (even if some income sources or assets remain unobserved). In such cases, agents' private information can be thought of as *residual uncertainty* after accounting for observables. Indeed, even public programs conditioned on earnings face substantial uncertainty over one's wealth.[9]

Still, what can the government do if it perfectly observes the protected characteristic? As it turns out, this gives it significant freedom to screen as long as it suitably adjusts the mechanism. Intuitively, the designer can now 'control for' the fact that some agents prefer the good because they are wealthy and screen purely based on need.

Proposition 4. *Suppose α is observable and the government uses either only money or only wait times to screen. Then $x(\alpha,\beta)$ is equitable and implementable if and only if $x(\alpha,\beta) \equiv \hat{x}(\eta(\alpha,\beta))$, where $\hat{x}(\cdot)$ is increasing.*

[9] https://thehill.com/regulation/administration/268409-outrage-builds-over-wealthy-families-in-public-housing/

Proof. Since α is observable, the allocation can be implemented separately for every value of α. Then, by the argument in the proof of Lemma 1, any implementable $x(\alpha, \beta)$ must be weakly increasing in β. Recall the equity constraint is satisfied if and only if $x(\alpha, \beta) \equiv \hat{x}(\eta(\alpha, \beta))$. Since $\eta(\alpha, \beta)$ is strictly increasing in β, it follows that $\hat{x}$ is implementable if and only if it is weakly increasing. □

5 Relaxing the equity constraint

My model of societal perceptions of equity was built around a single merit function. In reality, however, people often share general principles concerning equity and desert, but hold different opinions about finer trade-offs or ways in which these principles should be applied. For instance, Craxì et al. (2021) found that while healthcare workers about tended to agree which groups should get priority in the Covid-19 vaccine rollout, opinions differed on how these groups should be ranked. As a result, real equity constraints on policymakers would be less demanding than my analysis suggests. While Proposition 3 says that screening with multiple instruments is possible even under such strong restrictions, it is also interesting to ask how screening with only payments or only wait times compare when the equity constraint is relaxed. To that end, I develop a measure of *equity violation* capturing how far away a particular allocation is from satisfying the equity constraint (here, I interpret the merit function as a 'rough consensus' among the public). My measure assumes that every agent assesses the allocation's equity by looking at agents similar to herself, and comparing those with the same allocation as her to those with the same merit as her. In other words, she compares similar agents who *are* treated the same to those who *should* be treated the same. The further apart these two sets are, the more inequitable the allocation seems to her. Then, the degree to which the whole allocation violates equity is the size of the largest such 'local equity violation'.

Definition 2 (Equity violation)**.** *For every type $(\alpha, \beta) \in \Theta$, let $D(\alpha, \beta)$ be the set of directions in which the allocation is locally constant:*

$$D(\alpha, \beta) = \{d \in \mathbb{R}^2 : \nabla_d x(\alpha, \beta) = 0\}.$$

*Let the **local equity violation** for type (α, β) be:*

$$l(\alpha, \beta) = \begin{cases} \infty & \text{if } D(\alpha, \beta) = \varnothing, \\ \inf_{d \in D(\alpha, \beta)} \left| \arctan(d) - \arctan\left(-\frac{\eta_\alpha(\alpha, \beta)}{\eta_\beta(\alpha, \beta)}\right) \right| & \text{otherwise.}^{10} \end{cases} \tag{8}$$

*The **equity violation** of the allocation rule x, denoted $L(x)$, is its largest local equity violation:*

$$L(x) := \sup_{(\alpha, \beta) \in \Theta} l(\alpha, \beta).$$

[10] My results would be unaffected if arctan(·) was replaced with another bounded and strictly increasing function. However, using arctan(·) lets us visualize the size the local equity violation as an angle.

To build intuition, consider a smooth allocation that is strictly increasing in β and fix some type (α^a, β^a). Then all the types with allocations equal to that of (α^a, β^a) will lie on a smooth curve passing through (α^a, β^a). Figure 3a illustrates such a curve together with this type's *iso-merit curve*, that is, the set of types with the same merit as (α^a, β^a). Since both of these curves are smooth, we can compare 'how far apart' they are in the neighborhood of (α^a, β^a) by comparing their slopes there. The distance between these slopes, as measured by the angle between them, is the *local equity violation* at (α^a, β^a) (Figure 3b).

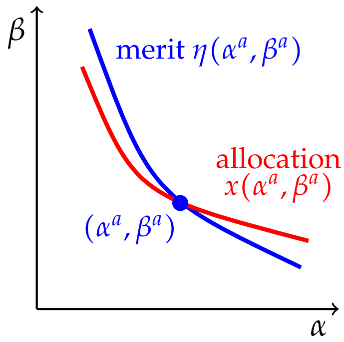

Figure 3a

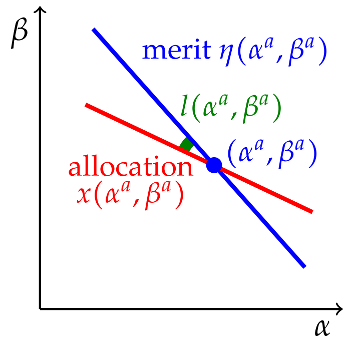

Figure 3b

We can ask which of the two screening instruments, when used on its own, will produce larger equity violations. A naïve (but incorrect) intuition suggests that screening with waiting will produce an allocation closer to the equitable one as it biases the allocation 'in the right direction'. However, as discussed in Subsection 3.2, screening with waiting will fail to 'match the shape' of the merit curve with the way it pools agents together. As it turns out, in some environments this 'shape effect' dominates the aforementioned 'direction effect'.

Screening with waiting will nevertheless produce smaller equity violations when the merit function depends on wealth α sufficiently strongly. The following proposition says that, as long as iso-merit curves are sufficiently flat everywhere, every separating allocation implemented using only payments will violate equity by more than some separating allocation implemented with only waiting.

Proposition 5. *Suppose wait times are bounded: $q \in [\underline{q}, \bar{q}]$ where $0 \in [\underline{q}, \bar{q}]$. Then there exists $M \in (-\infty, 0)$ with the following property: if for all $(\alpha, \beta) \in \Theta$ we have*

$$-\frac{\eta_\alpha(\alpha, \beta)}{\eta_\beta(\alpha, \beta)} > M,$$

then for every non-constant allocation rule x_p that is implementable with only payments, there exists a non-constant allocation rule x_q implementable with only waiting that produces a strictly smaller equity violation: $L(x_p) < L(x_q)$.

While the proof is relegated to the appendix, I illustrate its key intuition with the case of a smooth allocation that increases in β. Fix some type (α^a, β^a) and compare two allocations: x_q implemented using only payments, and x_p implemented using only waiting. Like before, the sets of agents with the same allocation as (α^a, β^a) will be smooth curves passing through that point. Moreover, the insights from Subsections 3.1 and 3.2 tell us that the curve for x_p will be upwards-sloping, while the curve for x_q will be downwards-sloping.

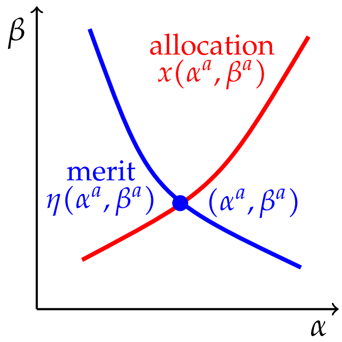

Figure 4a
Screening with only payments

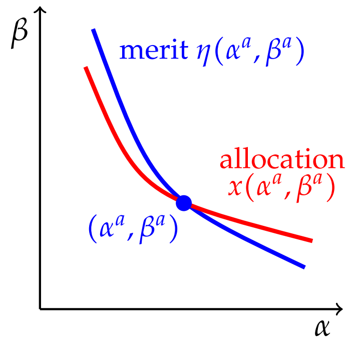

Figure 4b
Screening with only waiting

Let us now compare the local equity violations of from these two allocations. Figure 5a illustrates that if the iso-merit curve is sufficiently flat, its angle with the iso-$x_q(\alpha^a, \beta^a)$ curve will be smaller than that with the iso-$x_p(\alpha^a, \beta^a)$ one. If we can impose a sufficiently low uniform bound on the slopes of iso-merit curves, this will be true for every type, regardless of x^q. Figure 5b, on the other hand, illustrates why the result of Proposition 5 fails when iso-merit curves are not flat enough.

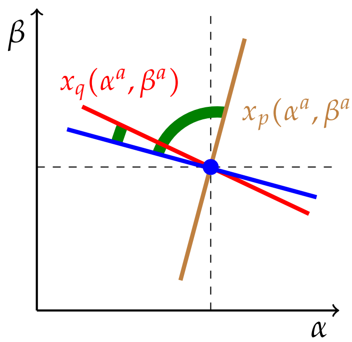

Figure 5a

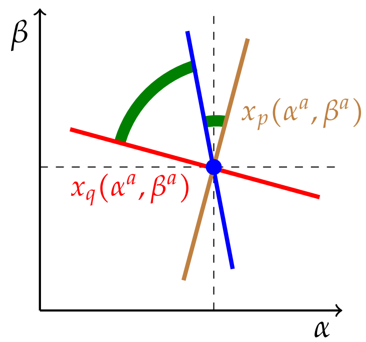

Figure 5b

This exercise offers broader qualitative lessons. Roughly speaking, Proposition 5 tells us that screening with waiting violates equity by less than screening with payments if merit depends more strongly on wealth than it does on need. However, the same will be true if agents vary a lot in wealth relative to how much they vary in need. To understand why, consider a mechanism screening with payments. If agents differ greatly in wealth (relative to how much they differ in need), most of the variance in allocations will be explained by preferences for money. This in turn means that the allocation will be heavily skewed towards the rich, which corresponds to a very flat iso-$x_p(\alpha^a, \beta^a)$ curve in Figure 5a. Consequently, the angle between this curve and the iso-merit curve will be wider.

6 Discussion

While my approach to modeling perceived equity is highly stylized, it offers general qualitative conclusions. First, every screening instrument will bias the allocation towards the group for which this instrument is less costly—this makes screening with payments problematic from an equity standpoint. Using a different instrument (like waiting) could reverse this bias, but the designer's control over the allocation would still be limited. Consequently, the resulting bias might still not satisfy the public. I show this problem can be solved by combining multiple screening instruments which on their own favor different social groups. Doing so gives the designer freedom to tinker with various groups' differential cost of the allocated good, and therefore to improve efficiency through screening while still producing an allocation that is seen as fair. I also highlight that screening with waiting need not always produce more equitable allocations than screening with payments. Waiting is, however, likely to be the more equitable instrument when society is highly averse to handing out benefits to wealthier people, or when the population's wealth heterogeneity is large relative to heterogeneity in terms of need for the good.

7 Appendix: omitted proofs

In what follows I write $(\alpha^a_\delta, \beta^a_\delta)$ for $(\alpha^a+\delta, \beta^a+\delta)$. I also write $x^a_\delta = x(\alpha^a_\delta, \beta^a_\delta)$, $p^a_\delta = p(\alpha^a_\delta, \beta^a_\delta)$ and $\eta^a_\delta = \eta(\alpha^a_\delta, \beta^a_\delta)$. I omit the subscript when $\delta = 0$.

Lemma 2. *Suppose there exists a decreasing sequence $\{\delta_i\}_i$ such that $\delta_i \to 0$ for which $\{x_{\delta_i}\}_i$ is strictly decreasing and $x^a_{\delta_i} \to x^a$. Then, for any (α^b, β^b) s.t. $x(\alpha^b, \beta^b) = x^a$ we have:*

$$\frac{v_x(\beta^b, x^a)}{z_q(\alpha^b, q^a)} \leq \frac{v_x(\beta^a, x^a)}{z_q(\alpha^a, q^a)}.$$

Proof. First, note that $\{q_{\delta_i}\}_i$ is strictly decreasing since strictly lower allocations must come with strictly lower wait times. I now show that $q^a_{\delta_i} \to q^a$. Indirect utility has to be continuous at every (α^a, β^a) (Milgrom and Segal (2002)), and so:

$$v(\beta^a, x^a) - z(\alpha^a, q^a) = \lim_{i\to\infty} \left\{ v(\beta^a_{\delta_i}, x^a_{\delta_i}) - z(\alpha^a_{\delta_i}, q^a_{\delta_i}) \right\}.$$

Note that $\lim_{i\to\infty} v(\beta^a_{\delta_i}, x^a_{\delta_i}) = v(\beta^a, x^a)$. Since $\alpha^a_{\delta_i} \to \alpha^a$ and z is continuous and strictly increasing in the latter argument, we have $q^a_{\delta_i} \to q^a$.

I now show that if the desired inequality fails, then $x(\alpha^b, \beta^b) \neq x^b$.

$$\frac{v_x(\beta^b, x^a)}{z_q(\alpha^b, q^a)} > \frac{v_x(\beta^a, x^a)}{z_q(\alpha^a, q^a)}. \tag{9}$$

By continuity of v_x and z_q, for i high enough we have that:

$$\frac{v_x(\beta^b, x^a)}{z_q(\alpha^b, q^a)} > \frac{v_x(\beta_{\delta_i}, x^a)}{z_q(\alpha_{\delta_i}, q^a)}. \tag{10}$$

Recall that $\{x^a_{\delta_i}\}_i$ and $\{q^a_{\delta_i}\}_i$ are strictly decreasing and tend to x^a and q^a. Hence, for all α, β:

$$\frac{v(\beta, x^a_{\delta_j}) - v(\beta, x^a)}{x^a_{\delta_j} - x^a} \to v_x(\beta, x^a); \quad \frac{z(\alpha, q^a_{\delta_j}) - z(\alpha, q^a)}{q^a_{\delta_j} - q^a} \to z_q(\alpha, q^a), \tag{11}$$

as $j \to \infty$. Fix i high enough that (10) holds. Then, by (11), for j high enough we have:

$$\frac{\frac{v(\beta^b, x^a_{\delta_j}) - v(\beta^b, x^a)}{x^a_{\delta_j} - x^a}}{\frac{z(\alpha^b, q^a_{\delta_j}) - z(\alpha^b, q^a)}{q^a_{\delta_j} - q^a}} > \frac{\frac{v(\beta_{\delta_i}, x^a_{\delta_j}) - v(\beta_{\delta_i}, x^a)}{x^a_{\delta_j} - x^a}}{\frac{z(\alpha_{\delta_i}, q^a_{\delta_j}) - z(\alpha_{\delta_i}, q^a)}{q^a_{\delta_j} - q^a}}.$$

In particular, take $j > i$ and notice that increasing i further relaxes the inequality. Hence:

$$\frac{\frac{v(\beta^b, x^a_{\delta_j}) - v(\beta^b, x^a)}{x^a_{\delta_j} - x^a}}{\frac{z(\alpha^b, q^a_{\delta_j}) - z(\alpha^b, q^a)}{q^a_{\delta_j} - q^a}} > \frac{\frac{v(\beta_{\delta_j}, x^a_{\delta_j}) - v(\beta_{\delta_j}, x^a)}{x^a_{\delta_j} - x^a}}{\frac{z(\alpha_{\delta_j}, q^a_{\delta_j}) - z(\alpha_{\delta_j}, q^a)}{q^a_{\delta_j} - q^a}}. \tag{12}$$

Now, by revealed preference we have:

$$v(\beta_{\delta_j}, x^a_{\delta_j}) - z(\alpha_{\delta_j}, q^a_{\delta_j}) \geq v(\beta_{\delta_j}, x^a) - z(\alpha_{\delta_j}, q^a) \implies \frac{\frac{v(\beta_{\delta_j}, x^a_{\delta_j}) - v(\beta_{\delta_j}, x^a)}{x^a_{\delta_j} - x^a}}{\frac{z(\alpha_{\delta_j}, q^a_{\delta_j}) - z(\alpha_{\delta_j}, q^a)}{q^a_{\delta_j} - q^a}} \geq \frac{q^a_{\delta_j} - q^a}{x^a_{\delta_j} - x^a}.$$

Combining the latter inequality with (12) gives:

$$\frac{\frac{v(\beta^b, x^a_{\delta_j}) - v(\beta^b, x^a)}{x^a_{\delta_j} - x^a}}{\frac{z(\alpha^b, q^a_{\delta_j}) - z(\alpha^b, q^a)}{q^a_{\delta_j} - q^a}} > \frac{q^a_{\delta_j} - q^a}{x^a_{\delta_j} - x^a} \implies v(\beta^b, x^a_{\delta_j}) - z(\alpha^b, q^a_{\delta_j}) > v(\beta^b, x^a) - z(\alpha^b, q^a).$$

Hence, $x^b(\alpha, \beta) \neq x^a$. □

7.1 Proof of Proposition 2

Fix a generic merit function $\eta(\alpha, \beta)$. Note that indirect utility has to be continuous at every (α^a, β^a) (Milgrom and Segal (2002)). By continuity of η, v and z, this implies:

$$v(\beta^a, \hat{x}_+(\eta^a)) - z(\alpha^a, \hat{q}_+(\eta^a)) = v(\beta^a, \hat{x}_-(\eta^a)) - z(\alpha^a, \hat{q}_-(\eta^a)). \tag{13}$$

If $\hat{x}(\cdot)$ were discontinuous at η^a, (13) would have to hold for every type with merit η^a, which would violate Condition 2 of Definition 1. Hence, $\hat{x}(\cdot)$ has to be continuous on $(\underline{\eta}, \overline{\eta})$.

Fact 1. *There exists $\eta^a \in (\underline{\eta}, \overline{\eta})$ and a decreasing sequence $\{f_i\}_i$ s.t. $f_i \to \eta^a$, $\hat{x}(f_i) \to \hat{x}(\eta^a)$ and $\{\hat{x}(f_i)\}_i$ is strictly decreasing.*

Proof. $\hat{x}(\cdot)$ is not constant, so there exist $\eta^b, \eta^c \in (\underline{\eta}, \overline{\eta})$ s.t. $\eta^b < \eta^c$ and $\hat{x}(\eta^b) < \hat{x}(\eta^c)$. Let $\eta^a = \sup\{\eta : \hat{x}(\eta) = \hat{x}(\eta^b)\}$. Since $\hat{x}(\cdot)$ is continuous, $\hat{x}(\eta^a) = \hat{x}(\eta^b)$. Now, take any decreasing sequence $\{e_i\}_i$ s.t. for every i, $e_i \in (\underline{\eta}, \overline{\eta})$ and $e_i \to \eta^a$. Since $\hat{x}(\cdot)$ is monotonic, $\hat{x}(e_i) \leq \hat{x}(e_j)$

whenever $i > j$. Also, by continuity of $\hat{x}(\cdot)$, $\hat{x}(e_i) \to \hat{x}(\eta^a)$ and, by the construction of η^a, $\hat{x}(\eta^a) < \hat{x}(e_i)$ for every i. Since $\hat{x}(\cdot)$ is continuous, there exists a subsequence $\{f_i\}_i$ of $\{e_i\}_i$ for which $\hat{x}(f_i)$ is decreasing strictly. $\square$

Take such η^a, $\{f_i\}_i$ and fix any (α^b, β^b) s.t. $\eta(\alpha^b, \beta^b) = \eta^a$. By Fact 1, there exists a decreasing sequence $\{f_i\}_i$ s.t. $f_i \to \eta^a$ and $\hat{x}(f_i) \to \hat{x}(\eta^a)$. Now, by the continuity and monotonicity of $\eta(\alpha, \beta)$, for all i high enough there exist δ_i such that $\eta(\alpha^b_{\delta_i}, \beta^b_{\delta_i}) = f_i$, $\delta_i \to 0$ as $f_i \to \eta^a$, and $\{\delta_i\}_i$ is decreasing. Therefore, such a sequence exists for any type with merit η^a. Lemma 2 then tells us that all (α^b, β^b) with merit η^a must have same value of $v_x(\beta^b, x^a)/z_q(\alpha^b, x^a)$, which violates Condition 1 of Definition 1.

7.2 Proof of Proposition 5

Suppose we screen only with wait times ($p \equiv 0$) and the allocation is not constant.

Fact 2. *There exists (α^a, β^a) such that either a) for every $\delta > 0$, $x^a_\delta > x^a$, or b) for every $\delta < 0$, $x^a_\delta < x^a$.*

Proof. I first show that if $\alpha^a \geq \alpha^b$ and $\beta^a \geq \beta^b$, with at least one inequality holding strictly, then $x^a \geq x^b$. Suppose that $x^a < x^b$; then $q^a < q^b$, or else both types would strictly prefer (x^b, q^b). By revealed preference:

$$v(\beta^b, x^b) - z(\alpha^b, q^b) \geq v(\alpha^b, q^a) - z(\alpha^b, q^a) \quad \Longrightarrow \quad v(\beta^b, x^b) - v(\beta^b, x^a) \geq z(\alpha^b, q^b) - z(\alpha^b, q^a).$$

However, by strictly increasing differences, this gives:

$$v(\beta^a, x^b) - v(\beta^a, x^a) > z(\alpha^a, q^b) - z(\alpha^a, q^a) \quad \Longrightarrow \quad v(\beta^a, x^b) - z(\alpha^a, q^b) > v(\beta^a, x^a) - z(\alpha^a, q^a).$$

That is, (α^a, β^a) prefers (x^b, q^b) to (x^a, q^a); contradiction.

Now, suppose (α^a, β^a) described in the statement of the Fact does not exist; I will show the allocation must then be constant. Then for every $(\alpha^a, \beta^a) \in \Theta$ there exists some $\epsilon > 0$ such that $x^a_\delta = x^a$ for $\delta \in [-\epsilon, \epsilon]$. However, the above observation then tells us that $x(\alpha, \beta) = x^a$ for all $(\alpha, \beta) \in [\alpha^a - \epsilon, \alpha^a + \epsilon] \times [\beta^a - \epsilon, \beta^a + \epsilon]$. Consequently, for any $(\alpha, \beta) \in \Theta$ there exists a neighborhood around it in which the allocation is constant. Now, take any $(\alpha^b, \beta^b), (\alpha^c, \beta^c) \in \Theta$. Since Θ is connected, there exists a continuous path between them. Every point along this path has the same allocation as the points within its neighborhood. Therefore, the allocation has to be constant along the whole path, including the end-points. $\square$

Fix (α^a, β^a) from Fact 2 and assume a) holds. The argument for b) is analogous.

Fact 3. *If x^a_δ is right-continuous at $\delta = 0$, then for every (α^b, β^b) s.t. $x(\alpha^b, \beta^b) = x^a$, we have:*

$$\frac{v_x(\beta^b, x^a)}{z_q(\alpha^b, q^a)} \leq \frac{v_x(\beta^a, x^a)}{z_q(\alpha^a, q^a)}.$$

If x^a_δ is not right-continuous at $\delta = 0$, then for every (α^b, β^b) s.t. $x(\alpha^b, \beta^b) = x^a$, we have:

$$\frac{v(\beta^b, x^a_+) - v(\beta^b, x^a)}{z(\alpha^b, q^a_+) - z(\alpha^b, q^a)} \leq \frac{v(\beta^a, x^a_+) - v(\beta^a, x^a)}{z(\alpha^a, q^a_+) - z(\alpha^a, q^a)}.$$

Proof. Suppose x^a_δ is right-continuous at $\delta = 0$ and take a decreasing sequence $\{e_i\}_i$ such that $e_i \to 0$. Then $x^a_{e_i} \to x^a$ and, by case a) of Fact 2, $x^a < x^a_{e_i}$ for all i. Since x^a_δ is right-continuous at $\delta = 0$, there exists a subsequence $\{f_i\}_i$ of $\{e_i\}_i$ such that x_{f_i} is decreasing strictly. We can therefore apply Lemma 2 which completes the proof of this case.

Now suppose x^a_δ is not right-continuous at $\delta = 0$ and take a decreasing sequence $\{\delta_i\}_i$ such that $\delta_i \to 0$. Indirect utility has to be continuous at (α^a, β^a) (Milgrom and Segal (2002)), so:

$$v(\beta^a, x^a) - z(\alpha^a, q^a) = \lim_{i\to\infty} \left\{ v(\beta^a_{\delta_i}, x^a_{\delta_i}) - z(\alpha^a_{\delta_i}, q^a_{\delta_i}) \right\}.$$

Note $\lim_{i\to\infty} v(\beta^a_{\delta_i}, x^a_{\delta_i}) = v(\beta^a, x^a_+)$, where $x^a_+ > x^a$. Since $\alpha^a_{\delta_i} \to \alpha^a$ and z is continuous and strictly increasing in the latter argument, we have that $q^a_{\delta_i}$ converges to some $q^a_+ > q^a$. Hence:

$$v(\beta^a, x^a) - z(\alpha^a, q^a) = v(\beta^a, x^a_+) - z(\alpha^a, q^a_+) \implies \frac{v(\beta^a, x^a_+) - v(\beta^a, x^a)}{z(\alpha^a, q^a_+) - z(\alpha^a, q^a)} = 1. \tag{14}$$

By revealed preference, for every i we have:

$$v(\beta^b, x^b) - z(\alpha^b, q^b) \geq v(\beta^b, x^b_{\delta_i}) - z(\alpha^b, q^b_{\delta_i}) \implies \frac{v(\beta^b, x^b_{\delta_i}) - v(\beta^b, x^b)}{z(\alpha^b, q^b_{\delta_i}) - z(\alpha^b, q^b)} \leq 1.$$

Taking $i \to \infty$ gives:

$$\frac{v(\beta^b, x^b_+) - v(\beta^b, x^b)}{z(\alpha^b, q^b_+) - z(\alpha^b, q^b)} \leq 1. \tag{15}$$

Combining the latter inequality in (14) with (15) completes the proof. $\square$

In the respective cases of Fact 3, Assumption 1 ensures that the sets of $(\alpha^b, \beta^b) \in \Theta$ for which

$$\frac{v_x(\beta^b, x^a)}{z_q(\alpha^b, q^a)} = \frac{v_x(\beta^a, x^a)}{z_q(\alpha^a, q^a)} \quad \text{and} \quad \frac{v(\beta^b, x^a_+) - v(\beta^b, x^a)}{z_+(\alpha^b, q^a_+) - z(\alpha^b, q^a)} = \frac{v(\beta^a, x^a_+) - v(\beta^a, x^a)}{z(\alpha^a, q^a_+) - z(\alpha^a, q^a)}, \tag{16}$$

are downwards-sloping and differentiable curves. I will refer to them as *iso-MRS* and *iso-difference* curves at (α^a, β^a). Since the LHSs of (16) are increasing in β^b and α^b, Fact 3 tells us that all types allocated x^a must lie weakly below the iso-MRS curve at (α^a, β^a) in the first case, and weakly below the iso-difference curve at (α^a, β^a) in the second case. Hence, the only possible element of $D(\alpha^a, \beta^a)$ is the slope of the iso-MRS curve at (α^a, β^a) in the first case, and the slope of the iso-difference curve at (α^a, β^a) in the second case:

$$s^q_{MRS}(\alpha^a, \beta^a) := \frac{v_x(\alpha^a, x^a)}{z_q(\alpha^a, q^a)} \frac{z_{\alpha q}(\alpha^a, q^a)}{v_{\beta x}(\alpha^a, x^a)},$$

$$s^q_{diff}(\alpha^a, \beta^a) := \frac{v(\beta^a, x^a_+) - v(\beta^a, x^a)}{z(\alpha^a, q^a_+) - z(\alpha^a, q^a)} \frac{z_\alpha(\alpha^a, q^a_+) - z_\alpha(\alpha^a, q^a)}{v_\beta(\beta^a, x^a_+) - v_\beta(\beta^a, x^a)}.$$

By Assumption 1, we have $s^q_{MRS}(\alpha^a, \beta^a),\ s^q_{diff}(\alpha^a, \beta^a) < 0$.

An analogous argument produces corresponding expressions for the case where we screen only with payments. Then, however, $s^p_{MRS}(\alpha^a,\beta^a)$, $s^p_{diff}(\alpha^a,\beta^a) > 0$ for every $(\alpha^a,\beta^a)\in\Theta$. Now, note that the slope of the iso-merit curve at (α^a,β^a) equals to $-(\eta_\alpha(\alpha^a,\beta^a))/(\eta_\beta(\alpha^a,\beta^a))$ and is negative, since $\eta_\alpha,\eta_\beta > 0$. Therefore, we have the following lower bound on equity violations that holds across all x_p screening only with payments:

$$L(x_p) > \left|\frac{\pi}{2} - \arctan\left(-\frac{\eta_\alpha(\alpha,\beta)}{\eta_\beta(\alpha,\beta)}\right)\right|. \tag{17}$$

Now, notice:

$$\inf_{(\alpha,\beta)\in\Theta,\ x\in[\underline{x},\bar{x}]} v_x(\beta,x) \le v_x(\beta^b,x^a), \quad \frac{v(\beta^a,x^a_+)-v(\beta^a,x^a)}{x^a_+-x^a} \le \sup_{(\alpha,\beta)\in\Theta,\ x\in[\underline{x},\bar{x}]} v_x(\beta,x),$$

$$\inf_{(\alpha,\beta)\in\Theta,\ q\in[\underline{q},\bar{q}]} z_q(\alpha,q) \le z_q(\alpha^a,q^a), \quad \frac{z(\alpha^a,q^a_+)-z(\alpha^a,q^a)}{q^a_+-q^a} \le \sup_{(\alpha,\beta)\in\Theta,\ q\in[\underline{q},\bar{q}]} z_q(\alpha,q),$$

$$\inf_{(\alpha,\beta)\in\Theta,\ x\in[\underline{x},\bar{x}]} v_{\beta x}(\beta,x) \le v_{\beta x}(\beta^a,x^a), \quad \frac{v_\beta(\beta^a,x^a_+)-v_\beta(\beta^a,x^a)}{x^a_+-x^a} \le \sup_{(\alpha,\beta)\in\Theta,\ x\in[\underline{x},\bar{x}]} v_{\beta x}(\beta,x),$$

$$\inf_{(\alpha,\beta)\in\Theta,\ q\in[\underline{q},\bar{q}]} z_{\alpha q}(\alpha,q) \le z_{\alpha q}(\alpha^a,q^a), \quad \frac{z_\alpha(\alpha^a,q^a_+)-z_\alpha(\alpha^a,q^a)}{q^a_+-q^a} \le \sup_{(\alpha,\beta)\in\Theta,\ q\in[\underline{q},\bar{q}]} z_{\alpha q}(\alpha,q).$$

Since Θ is bounded and $v_x, z_q, v_{\beta x} > 0$, $z_{\alpha q} < 0$ on the closure of Θ, the bounds above are finite and there exists $M\in(-\infty,0)$ such that for any $(\alpha,\beta)\in\Theta$:

$$s^q_{MRS}(\alpha,\beta),\ s^q_{diff}(\alpha,\beta) < M.$$

That is, all iso-MRS and iso-difference curves for all allocations x_q screening only with waiting are steeper than M. Now, suppose all iso-merit curves are flatter than M:

$$\text{for all } (\alpha,\beta)\in\Theta, \quad -\frac{\eta_\alpha(\alpha,\beta)}{\eta_\beta(\alpha,\beta)} > M.$$

I will construct an allocation x_q whose equity violation is below the bound in (17). Fix any $(\alpha^b,\beta^b)\in\Theta$, $q^v > 0$ and $x^b\in[\underline{x},\bar{x}]$ such that $x^b > x_0$ and:

$$\frac{v(\beta^b,x^b)-v(\beta^b,x_0)}{w(\alpha^b,q^b)-w(\alpha^b,0)} = 1.$$

Note that for x^b sufficiently close to x_0, q^b will be close to 0 and hence belong to $[\underline{q},\bar{q}]$. Now, the following allocation satisfies (IR) and (IC): (α^b,β^b) and all types weakly above the iso-difference curve at (α^b,β^b) take (x^b,q^b); all types strictly below it take $(x_o,0)$.

Take any (α^c,β^c) strictly below or strictly above the iso-difference curve at (α^b,β^b). There is a neighborhood around (α^c,β^c) in which the allocation is constant, and so it is also constant along the slope of the iso-merit curve at (α^c,β^c) giving $l(\alpha^c,\beta^c) = 0$. Meanwhile, for any type (α^d,β^d) on the iso-difference curve, the allocation is locally constant along the slope

$s_{diff}(\alpha^d, \beta^d)$. Since the iso-merit curve is flatter than $s_{diff}(\alpha^d, \beta^d)$, we get:

$$L(x_q) \leq \left| \arctan(s^q_{diff}(\alpha^c, \beta^c)) - \arctan\left(-\frac{\eta_\alpha(\alpha^c, \beta^c)}{\eta_\beta(\alpha^c, \beta^c)}\right) \right| < \left| \frac{\pi}{2} - \arctan\left(-\frac{\eta_\alpha(\alpha^c, \beta^c)}{\eta_\beta(\alpha^c, \beta^c)}\right) \right| < L(x_p).$$

References

ABDULKADIROĞLU, A. AND T. SÖNMEZ (2003): "School Choice: A Mechanism Design Approach," *American Economic Review*, 93, 729–747.

AKBARPOUR, M., E. BUDISH, P. DWORCZAK, AND S. D. KOMINERS (2023): "An Economic Framework for Vaccine Prioritization*," *The Quarterly Journal of Economics*, qjad022.

AKBARPOUR, M., P. DWORCZAK, AND S. D. KOMINERS (2020): "Redistributive allocation mechanisms," *Available at SSRN 3609182*.

ALVES, G., F. BERNIER, M. COUCEIRO, K. MAKHLOUF, C. PALAMIDESSI, AND S. ZHIOUA (2023): "Survey on Fairness Notions and Related Tensions," .

CONDORELLI, D. (2013): "Market and non-market mechanisms for the optimal allocation of scarce resources," *Games and Economic Behavior*, 82, 582–591.

CRAXÌ, L., A. CASUCCIO, E. AMODIO, AND V. RESTIVO (2021): "Who should get COVID-19 vaccine first? A survey to evaluate hospital workers' opinion," *Vaccines*, 9, 189.

DESSEIN, W., A. FRANKEL, AND N. KARTIK (2023): "Test-Optional Admissions," *arXiv preprint arXiv:2304.07551*.

ELSTER, J. (1989): *Solomonic judgements: Studies in the limitation of rationality*, Cambridge University Press.

KASS, L. R. (1997): "The wisdom of repugnance: Why we should ban the cloning of humans," *Val. UL Rev.*, 32, 679.

MILGROM, P. AND I. SEGAL (2002): "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 70, 583–601.

ONUCHIC, P. (2022): "Recent contributions to theories of discrimination," *arXiv preprint arXiv:2205.05994*.

ROCHET, J.-C. (1987): "A necessary and sufficient condition for rationalizability in a quasilinear context," *Journal of Mathematical Economics*, 16, 191–200.

ROTH, A. E. (2007): "Repugnance as a Constraint on Markets," *Journal of Economic Perspectives*, 21, 37–58.

STONE, P. (2008): "What can lotteries do for education?" *Theory and Research in Education*, 6, 267–282.

WALZER, M. (1983): "Spheres of Justice. A defense of Pluralism and Equality. New York: Basis Book," *Inc., Publishers*.

WEITZMAN, M. L. (1977): "Is the Price System or Rationing More Effective in Getting a Commodity to Those Who Need it Most?" *The Bell Journal of Economics*, 8, 517–524.