The London School of Economics and Political Science

Essays on Sorting and Inequality

Lisa Verena Windsteiger

A thesis submitted to the Department of Economics of the London School of Economics for the degree of Doctor of Philosophy. London, July 2017

To my parents

Declaration

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of approximately 51,800 words.

Acknowledgements

I would like to thank my supervisor, Ronny Razin, for his support and guidance, for his rigorous attention and for all his time devoted to giving me feedback on my work. I am also grateful to many others at LSE: to Gilat Levy, Balazs Szentes, Michele Piccione and all other members of the theory group, as well as Ethan Ilzetzki, Camille Landais and Frank Cowell, for providing me with invaluable feedback on my work and for always being available for questions and advice. In particular, I want to thank Stephane Wolton, who was a great mentor to me and introduced me to the amazing PSPE group at LSE.

My gratitude extends to Lord Adair Turner, who I had the privilege to work for as a research assistant and who taught me about the history of economic thought, and to Thomas Piketty, who advised me on my survey in the context of the Piketty Masterclass.
I would also like to thank Christian Haefke and Franz X. Hof, who were important mentors to me during my studies in Austria - their guidance and encouragement helped me find my way into academia.

I've been very fortunate to share my PhD experience with amazing people: Sutanuka Roy, Ana McDowall, Dana Kassem, Stephan Maurer, Pily Lopez Uribe, Frank Pisch, Francesco Sannino, Roberto Sormani, Michel Azulai and many others have been fun and helpful colleagues and provided both intellectual and emotional support when needed.

I am eternally grateful to my family: To my wonderful father, who didn't live to see me finish my PhD but who gave me so much while he was there, to my siblings Laura and Lukas, whose presence at home enabled me to stay in London during tough times for my family, and to my amazing mother: Without her constant encouragement and help I wouldn't be where I am now. She is the strongest person in the world and my greatest idol, and she has taught me – already long before the "Wonder Woman" movie came out – that women can be superheroes.

I also want to thank my parents in law for always making me feel welcome in their home and for treating me like a family member from day one.

Finally, I want to thank my husband and best friend Stefan Rosenlechner for his immense support, understanding and patience, for giving me a home (and a health insurance) and for always bringing me back down to earth when I'm about to get lost in my ivory tower.

Abstract

This thesis consists of three papers that examine sorting and inequality.
In the first paper I present a model in which people sort into groups according to income and as a result become biased about the shape of the income distribution. Their biased beliefs in turn affect who they choose to interact with, and hence there is a two-way interaction between segregation and misperceptions about society. I show one possible application of this novel framework to the question of income inequality and the demand for redistribution. I demonstrate that under segregation an increase in income inequality can lead to a decline in perceived inequality and therefore to a fall in people's support for redistribution. I motivate my main assumptions with empirical evidence from a small survey that I conducted via Amazon Mechanical Turk.
In the second paper I develop a general model of how social segregation and beliefs interact. Sorting decisions will be affected by beliefs about society, but these beliefs about society are in turn influenced by social interactions. In my model, people sort into social groups according to income, but become biased about the income distribution once they interact only with their own social circle. I define "biased sorting equilibria", which are stable partitions in which people want to stay in their chosen group, despite their acquired misperceptions about the other groups. I introduce a refinement criterion – the consistency requirement – and find necessary and sufficient conditions for existence and uniqueness of biased sorting equilibria.
In the third paper I present a model in which a monopolist offers citizens the opportunity to segregate into groups according to income. I focus initially on the case of two groups and show that a monopolist with fixed costs of offering the sorting technology will see profits increase as income inequality increases. I then analyze how the monopolist's optimal group partition varies with inequality and show that for a broad field of income distributions, monopolist profits increase with inequality, while at the same time total welfare of sorting given the monopolist's optimal schedule decreases. In the last section I examine how these findings generalize if the monopolist doesn't face costs of offering the sorting technology and can therefore offer as many groups as she wants.

Contents

List of Figures

List of Tables

Chapter 1

The Redistributive Consequences of Segregation[1]

1.1 Introduction

Most industrialized countries have seen a remarkable increase in income and wealth inequality over the past 35 to 40 years (see e.g. Piketty (2014)). At the same time, support for redistributive policies hasn't exhibited a comparable trend in the majority of these countries. For instance, demand for redistribution as proxied by realized tax- and redistribution rates has remained relatively constant or even decreased over the last two decades in the US (see Piketty et al. (2014)). Of course there are many reasons - above all institutional ones - why realized tax rates need not reflect demand for redistribution well. However, also demand for redistribution as measured by household surveys has not evolved in the same way as (income) inequality (see Ashok et al. (2015) and Kenworthy and McCall (2008)). This is at odds with standard Political Economy models, which predict that high rates of income inequality trigger high demand for redistribution. For instance, in the baseline model of Meltzer and Richard (1981) the redistribution rate that is determined by majority voting is increasing in the difference between median and mean income.

Rising income and wealth inequality have frequently been accompanied by an increase in socio-economic segregation. Watson (2009) and Reardon and Bischoff (2011) demonstrate that both income inequality and income segregation have risen sharply in the US between 1970 and 2000, especially in metropolitan areas. Often, middle-income neighbourhoods have made way for both rich and poor communities, and segregation and the erosion of the middle class have gone hand in hand.

In the present paper, I want to combine these observations with the finding that people tend to misperceive the shape of the income distribution. Empirical studies in the US and Australia find that people underestimate income and wealth inequality and wage differentials (see e.g. Norton and Ariely (2011) and Kiatpongsan and Norton (2014)) and I detect similar types of misperception in my own survey conducted in the US via Amazon Mechanical Turk (see Section 1.5).

Connecting all these pieces, I build a model that explains why the relationship between income and wealth inequality and support for redistributive policies could be non-monotone in general: In my model people are segregated according to income, and therefore interact mainly with others who have similar incomes to themselves. As a result, they lose sight of the overall income

[1]I thank Ronny Razin, Stephane Wolton, Thomas Piketty, Mike Savage, Matt Levy, Camille Landais, Gilat Levy, Daniel Laurison, Dominik Hangartner, Stephan Maurer, Frank Cowell and the participants of the LSE Microeconomic Theory Work in Progress Seminar for their helpful suggestions and comments.

distribution and become biased about the shape of the income distribution. Specifically, they underestimate how different others are to themselves and therefore underestimate income inequality.

This has an effect on their support for redistributive policies: People in my model will in general demand less redistribution than in a model without misperceptions. Furthermore, I show that an increase in inequality will, in the presence of segregation and misperceptions, always lead to a smaller increase in demand for redistribution than in a model where people are unbiased, and that it can in certain circumstances even lead to a decrease in demand for redistribution.

At the end of the paper I support my assumptions about misperceptions of the income distribution and segregation by presenting evidence from a survey that I conducted via Amazon Mechanical Turk.

The rest of this paper is organized as follows: Section 1.2 discusses related literature. Section 1.3 presents a theoretical model of economic sorting with misperceptions where people underestimate inequality and Section 1.4 applies this model to the issue of voting for redistribution. Section 1.5 presents suggestive empirical evidence on misperceptions about the shape of the income distribution and on how socio-economic segregation and misperceptions of the income distribution are related. Section 1.6 concludes.

1.2 Relation to existing literature

In Windsteiger (2017b) I present a general model in which beliefs about society and segregation decisions interact to create an endogenous system of beliefs and social groups. For related literature on segregation and belief formation see Windsteiger (2017b).

In the present paper I apply this general model to the situation of sorting according to income and support for redistributive policies. Standard political economy models (see e.g. Meltzer and Richard (1981)) predict that the demand for redistribution should be higher, the poorer the median earner is relative to average income in society. However, studies comparing pre-tax income inequality to redistribution rates in democracies, and hence trying to confirm the Meltzer-Richard Model empirically, deliver mixed results. Some papers do indeed find a positive link between inequality and redistribution (see e.g. Borge and Rattsoe (2004), Meltzer and Richard (1983) and Milanovic (2000)). However, others detect a negative relationship (e.g. Georgiadis and Manning (2012) and Rodriguez (1999)) or no significant link at all (e.g. Kenworthy and McCall (2008) and Scervini (2012)).

There are many explanations for why a high degree of inequality might not be reflected in high realized redistribution rates in an economy: Bartels (2009) argues that the views of the majority might be disregarded by political leaders due to successful lobbying of the financially powerful. Moreover, poor people might participate in the political process to a lesser degree than rich people, which might shift the identity of the median - decisive - voter (see e.g. Larcinese (2005)). Finally, and importantly, people rarely get to vote directly on redistribution rates. Instead, policital candidates offer platforms that take a position on a variety of issues, and people might vote against their interest on the subject of redistribution if they consider other issues to be more important (see Matakos and Xefteris (2016)).[2] However, apart from these factors, which affect *realized* redistribution rates, it seems to be the case that even the pure redistributive preferences of the population are not in line with what we might call the "Meltzer-Richard-Hypothesis": that pre-tax inequality and the demand for redistribution should be positively correlated, both across countries and over time (see e.g. Ashok et al.

[2] For a concise overview see Bonica et al. (2013).

(2015)).
In the Meltzer-Richard Model, people aim to maximize their own after-tax income and hence their sole concern is their relative position in the income distribution as a direct predictor of how much they would benefit or lose from redistribution. More detailed models allow for people's preferences for redistribution to be influenced also by other factors, such as social mobility, the overall degree of inequality in society or social status concerns (see e.g. Piketty (1995), Benabou and Ok (2001), Alesina and Angeletos (2005) and Corneo and Gruener (2000)). This can explain why the median voter's relative position in the income distribution is not necessarily a good predictor of a society's demand for redistribution. However, also in these more elaborate models it will be the case that if inequality increases (ceteris paribus), demand for redistribution increases.[3] Nevertheless, empirically we find that periods of increasing inequality can be accompanied by stagnant or declining demand for redistribution.
The main contribution of my paper is that I show how my model of endogenous segregation and belief formation can be used to explain low support for redistribution in societies where inequality is high: As people interact only with people who have similar income to their own, they misperceive the shape of the whole income distribution, and poor people (including the median voter) underestimate how much they could gain from redistribution. Moreover, I demonstrate that with endogenous segregation and beliefs, the relationship between redistributive demand and inequality can be non-monotone - an increase in inequality can lead to a decline in the demand for redistribution, because people, if they see only a select group of society, might perceive that inequality has gone down due to the change in the income distribution.
There is a growing empirical literature on people's misperceptions of the income distribution. Cruces et al. (2013) find that poor people in Buenos Aires overestimate their relative position in the income distribution, while rich people underestimate it. They also show that this lowers poor people's demand for redistribution: when their biases are corrected, poor people's demand for redistribution increases. Importantly, they additionally show that (social resp. economic) segregation affects people's misperceptions. Karadja et al. (2015) conduct a similar study for Sweden and find that a majority of people there tend to underestimate their relative position. Norton and Ariely (2011) and Norton et al. (2014) find that people in the US and Australia tend to underestimate income and wealth inequality and Kiatpongsan and Norton (2014) find that people underestimate pay differences between different professions.
Kuziemko et al. (2015) perform a series of online experiments to analzye how information about inequality and its evolution over time effects people's demand for redistribution. They find that information has large effects on whether people see inequality as a problem, but it doesn't move redistributive preferences a lot. The only exception is the estates tax: informing people about the tiny share of inheritants who are subject to it drastically increases support for it. The latter result seems to be due to a huge degree of ex-ante misinformation about the estates tax and its incidence. They hypothesize that the relatively small effect of information on all other policy preferences might be due to the fact that becoming aware of the true extent of inequality and its increase makes people less confident that the government is capable of dealing with this issue, which is why respondents do not think redistributive policies can solve the problem.
Concerning the theoretical model of sorting according to income outlined in Section 1.3, my

[3] A notable exception here is Corneo and Gruener (2000), where an increase in economic inequality can lead to a decrease in the preferred tax rate of the middle class due to status concerns - the signalling power of wealth decreases more rapidly with the tax rate if income inequality is high and the middle class want to avoid mixing with the lower class. Note however that this depends crucially on the assumption that social and economic inequality move independently and that the middle class has a higher than average social status and a lower than average economic status.

paper is closely related to Levy and Razin (2015). They analyze preferences for redistribution in the presence of costly income sorting. They identify simple conditions on the shape of the income distribution such that a majority of the population (even people with income above average) respectively a benevolent social planner prefer full redistribution (or no sorting) to costly income sorting and they show that in both cases these conditions are satisfied for relatively equal income distributions. Hence, one implication of their model is that an increase in income inequality can make sorting more desirable from a welfare perspective.

1.3 Sorting with misperceptions

In the following section I will introduce a theoretical model of sorting with misperceptions. With the help of this framework I can then predict how groups in society will look like in equilibrium and - because social interactions affect beliefs - also what kind of misperceptions people will have about the overall income distribution. The model below is a simplified version of a more general model presented in Windsteiger (2017b).

Let income y in an economy be distributed according to an income distribution $F(y)$, on the interval $Y = [0, y_{\max}]$ where $y_{\max} < \infty$. Assume furthermore that $F(y)$ is continuous and strictly monotonic. As $F(y)$ is an income distribution, I will also assume that $F(y)$ is positively skewed (meaning that the median income is smaller than the average income).

Suppose that an agent's utility is increasing not only in her own income but also in the average income of the people that she interacts with, which I will henceforth call her "reference group". Specifically, a person with income y_j gets utility $U_j = y_j E(y|y \in S_i)$, where S_i is individual j's reference group. If there is no economic segregation, everybody's reference group is a representative sample of the whole population, such that $U_j = y_j E(y)$. However, a person with income y_j can pay a fee $b > 0$ to join group S_b and get utility

$$y_j E[y|y \in S_b] - b$$

or refrain from paying b and get

$$y_j E[y|y \in S_0]$$

where S_b is the set of incomes y of people who have paid b and S_0 is the set of incomes y of people who haven't paid b. If people are unbiased about the overall income distribution, a partition $\{S_0, S_b\}$ of Y and a sorting fee b constitute a **sorting equilibrium** iff

$$yE[y|y \in S_b] - b \leq yE[y|y \in S_0] \quad \forall y \in S_0 \tag{1.1}$$

$$yE[y|y \in S_b] - b \geq yE[y|y \in S_0] \quad \forall y \in S_b \tag{1.2}$$

In a sorting equilibrium people stay in the group that gives them the highest utility.

Suppose that people, once they are sorted into their group, become biased about average income in the other group and hence about the overall income distribution. I will model a group's belief about the other group as resulting from a group belief "technology". Specifically, I will assume that people's biased perception of the other group's average income can be characterized by the continuous belief function

$$B : \mathbf{P} \to Y^4$$

where $\mathbf{P}$ is the space of all monotone partitions $P = [S_0, S_b]$ of Y. For the following analysis, I will restrict my attention to monotone partitions, i.e. partitions $P = [S_0, S_b]$ of Y that can be uniquely characterized by a cutoff $\hat{y} \in Y$ (with the convention that $S_0 = [0, \hat{y})$ and

$S_b = [\hat{y}, y_{\max}]$), and I will henceforth call the people in S_0 "the poor" and the people in S_b "the rich". Without further assumptions, also non-monotone equilibria are possible if people have misperceptions. In Appendix 1.7.1 I show that restricting the analysis to monotone partitions is without loss of generality for the analysis that I conduct in this paper.[4]

I will assume that people are correct about average income in their own group. Furthermore, I require misperceptions to be constant within groups, i.e. people who are in the same group have the *same* misperception about the other group's average (and thus misperceptions do not depend on one's own income directly, but on group membership).

The belief function B is thus a continuous function that maps all monotone partitions of Y (and note that any monotone partition can be uniquely characterized by the cutoff $\hat{y}$) into a four-dimensional vector of beliefs

$$B(\hat{y}) = (\underline{E}(\hat{y}), \bar{E}_p(\hat{y}), \underline{E}_r(\hat{y}), \bar{E}(\hat{y}))$$

where the first two entries denote the poor group's belief about average income in the poor and the rich group respectively and the last two entries denote the rich group's belief about average income in the poor and the rich group. $\underline{E}(\hat{y})$ is the true average income in the poor group, i.e. $\underline{E}(\hat{y}) = E[y|y < \hat{y}]$ and $\bar{E}(\hat{y})$ is the correct average income in the rich group, $\bar{E}(\hat{y}) = E[y|y \geq \hat{y}]$. The poor's belief about average income in the rich group is $\bar{E}_p(\hat{y})$ and the rich's belief about average income in the poor group is $\underline{E}_r(\hat{y})$.

Given the belief function B, I can define the following:

Definition 1.1 *A monotone partition of Y (characterized by an equilibrium cutoff $\hat{y}^*$) and a sorting fee $b > 0$ constitute a **biased sorting equilibrium** iff*

$$\begin{aligned} y\bar{E}_p(\hat{y}^*) - b &\leq y\underline{E}(\hat{y}^*) \quad \forall y \in [0, \hat{y}^*) && \text{(IC1)} \\ y\bar{E}(\hat{y}^*) - b &\geq y\underline{E}_r(\hat{y}^*) \quad \forall y \in [\hat{y}^*, y_{\max}] && \text{(IC2)} \end{aligned}$$

A biased sorting equilibrium is therefore a partition of Y that is "stable" given people's misperceptions about the other group. People compare the utility they obtain in their own group to the utility they *think* they could obtain in the other group, given their misperceptions about average income in the other group. In a biased sorting equilibrium people think that they reach the highest possible level of utility in their own group and therefore they do not want to move to the other group.

Assuming that people have misperceptions about average income in the other group creates consistency issues: In a biased sorting equilibrium, people's beliefs about the other group can be inconsistent with what they see. A person in the poor group might wonder why a person in the rich group finds it worthwhile to pay b, given the poor person's belief about average income in the rich group. Similarly, a person in the rich group might - given the rich group's misperception about average income in the poor group - wonder why a certain person in the poor group doesn't want to join the rich group.

However, this inconsistency vanishes if I introduce what I call the *consistency requirement.* A partition of society satisfies consistency if people's beliefs about the other group are in line with what they observe: People who are in the poor group think that the people who are in the rich group are correct in doing so and vice versa. In Windsteiger (2017b), I explain this requirement in detail.[5] Formally, the consistency requirement translates to

[4] I show that the refinement that I introduce in this section (the consistency requirement), implies monotonicity.

[5] If society is divided into more than two groups, the requirement can be stated in a *global* and a *local* form. In the case of two groups, the two notions coincide, which is why I will talk only about "consistency" in the present paper, without specifying whether it is local or global.

Definition 1.2 *A monotone partition of Y (characterized by a cutoff $\hat{y}$) and a sorting fee b satisfy* ***consistency*** *iff*

$$y\bar{E}(\hat{y})) - b \leq y\underline{E}_r(\hat{y}) \quad \forall y \in [0, \hat{y}) \tag{CR1}$$
$$y\bar{E}_p(\hat{y}) - b \geq y\underline{E}(\hat{y})) \quad \forall y \in [\hat{y}, y_{\max}] \tag{CR2}$$

In words, condition (CR1) requires that a person in the rich group who looks at any person with income y in the poor group thinks that this person cannot achieve higher utility by switching to the rich group (and note that the person from the rich group evaluates person y's utility in the poor group given her own biased perception of average income in the poor group, $\underline{E}_r(\hat{y})$). Condition (CR2) does the same for poor people's belief about the rich group. Without misperceptions, consistency is implicit in any sorting equilibrium. Because everybody has the same (correct) understanding of average incomes in both groups, people cannot be "puzzled" by other people's choices - everybody evaluates everybody else's utility in the same way. It is only when people have incorrect perceptions of the other group that consistency becomes a separate issue and is not implicit in the equilibrium definition. People can be happy with their own choices (which means the partition constitutes a biased sorting equilibrium), while at the same time not understanding other people's choices (which means that consistency is violated). Hence, it makes sense - as a refinement to biased sorting equilibria - to define biased sorting equilibria which additionally satisfy consistency:

Definition 1.3 *A monotone partition of Y (characterized by an equilibrium cutoff $\hat{y}^*$) and a sorting fee $b > 0$ constitute a* ***biased sorting equilibrium with consistency*** *iff*

$$y\bar{E}_p(\hat{y}^*) - b \leq y\underline{E}(\hat{y}^*) \quad \forall y \in [0, \hat{y}^*) \tag{IC1}$$
$$y\bar{E}(\hat{y}^*) - b \geq y\underline{E}_r(\hat{y}^*) \quad \forall y \in [\hat{y}^*, y_{\max}] \tag{IC2}$$

and

$$y\bar{E}(\hat{y}^*)) - b \leq y\underline{E}_r(\hat{y}^*)) \quad \forall y \in [0, \hat{y}^*) \tag{CR1}$$
$$y\bar{E}_p(\hat{y}^*) - b \geq y\underline{E}(\hat{y}^*) \quad \forall y \in [\hat{y}^*, y_{\max}] \tag{CR2}$$

In Windsteiger (2017b) I show the following:

Corollary 1.1 *A monotone partition of Y (characterized by a cutoff $\hat{y}^*$) and a sorting fee $b > 0$ constitute a biased sorting equilibrium with consistency iff*

$$\hat{y}\bar{E}_p(\hat{y}^*) - \hat{y}\underline{E}(\hat{y}^*) \tag{1.3}$$
$$= \hat{y}^*\bar{E}(\hat{y}^*) - \hat{y}^*\underline{E}_r(\hat{y}^*)$$
$$= b$$

A biased sorting equilibrium with consistency is thus a partition where the perceived benefit of being in the rich group rather than the poor group (in terms of utility) of the person with income at the equilibrium cutoff $\hat{y}^*$ is regarded to be equally high by both groups. Note that for a given equilibrium cutoff $\hat{y}^*$ that satisfies (1.3), the corresponding sorting fee b is unique. The equilibrium condition (1.3) restricts the set of belief functions which imply equilibrium existence. In Windsteiger (2017b), I derive conditions on this function such that equilibrium exists and is unique. For the remainder of this paper I want to focus on a particular type of belief function: One where the poor underestimate average income in the rich group and the rich overestimate average income in the poor group, and therefore both groups underestimate income inequality. As I argue in the introduction, this is what empirical evidence shows. I will present suggestive evidence for such misperceptions and how they are connected to segregation in Section 1.5, where I explain a survey that I conducted myself via Amazon Mechanical Turk. In Appendix 1.7.10 I examine the implications for the model if people have misperceptions of the opposite type, where both groups overestimate inequality, and I compare the two types of misperceptions in terms of welfare and profit of a monopolist who offers the sorting technology in Appendix 1.7.11 and 1.7.12, respectively.

1.3.1 Underestimating Inequality

Suppose the belief function B is such that the people in the poor group think that average income in the rich group is

$$\bar{E}_p(\hat{y}) = \beta(1 - F(\hat{y}))\hat{y} + (1 - \beta(1 - F(\hat{y}))\bar{E} \tag{1.4}$$

and the people in the rich group think that average income in the poor group is

$$\underline{E}_r(\hat{y}) = \gamma F(\hat{y})\hat{y} + (1 - \gamma F(\hat{y}))\underline{E}. \tag{1.5}$$

$\beta \in [0,1]$ and $\gamma \in [0,1]$ parameterize the "naivity" of agents and if β resp. γ is 0 agents have no misperceptions. It is straightforward to see that $\bar{E}_p(\hat{y}) < \bar{E}(\hat{y})$ and $\underline{E}_r(\hat{y}) > \underline{E}(\hat{y})$ for all $y \in (0, y_{\max})$, i.e. the poor underestimate average income of the rich and the rich overestimate average income of the poor for any interior cutoff. The functional form of $\bar{E}_p(\hat{y})$ and $\underline{E}_r(\hat{y})$ implies that the misperceptions are more severe, the smaller the part of the distribution that they can fully observe (which is $F(\hat{y})$ for the poor group and $1 - F(\hat{y})$ for the rich group). Specifically, we have that

$$\frac{d(\bar{E}(\hat{y}) - \bar{E}_p(\hat{y}))}{d\hat{y}} = -\beta(1 - F(\hat{y})) < 0 \quad \forall \hat{y} \in (0, y_{\max})$$

and

$$\frac{d(\underline{E}_r(\hat{y}) - \underline{E}(\hat{y}))}{d\hat{y}} = \gamma F(\hat{y}) > 0 \quad \forall \hat{y} \in (0, y_{\max})$$

and therefore the misperceptions converge to the truth monotonically as $\hat{y}$ goes to 0 resp. $y_{\max}$.[6]

Misperceptions of this type could arise in the following way: As people live in their segregated communities, they see mostly people who have income similar to their own (i.e. people from their own group). They do meet people from the other group, but they are not aware that most of the time they do not meet a representative sample of the other group (because they are more likely to meet people from the other group who are close to the cutoff). They see the average income in their own group, but what matters for their sorting decision is also the average income in the other group, which they do not see. Because they know $\hat{y}$ and the overall range of y (i.e. that y ranges from 0 to $y_{\max}$), they know that the average income of the other group lies somewhere between the cutoff $\hat{y}$ and 0 resp. $y_{\max}$. However, as they neglect the fact that they often do not meet a representative sample of the other group and are rather more likely to meet people very close to the cutoff, the poor think that the average in the rich

[6] For the following analysis it is not necessary that the misperceptions are of exactly of the form (1.4) and (1.5) . For the results of the next section to hold, I need the misperceptions to be such that a binary biased sorting equilibirum exists and is (ideally) unique. Sufficient conditions for this are stated in Windsteiger (2017). Furthermore, the equilibrium cutoff needs to be located above median income. In Appendix 1.7.2, I specify sufficient conditions on the belief function to guarantee that there is a unique interior equilibrium cutoff above the median.

group is closer to their own average than it actually is, and the same holds for the rich when thinking about the poor group's average. In short, people below the cutoff **underestimate** average income in the rich group and people above the cutoff **overestimate** average income in the poor group. This will lead both groups to underestimate the benefits of sorting: The rich because they think the poor are less poor than they actually are, and the poor because they think the rich are not as rich as they actually are.[7]

The functional form of the misperceptions as given by (1.4) and (1.5) is such that the sufficient conditions for existence of a biased sorting equilibrium with consistency are satisfied (see Windsteiger (2017b)): $\bar{E}_p(\hat{y})$ and $\underline{E}_r(\hat{y})$ are continuous functions and each group is correct at one of the endpoints, whereas the other group is maximally biased at that respective endpoint. Furthermore, the misperceptions converge to the truth monotonically, and therefore there exists a unique interior equilibrium cutoff if the utility function is linear. However, it is not necessary to invoke these general conditions here, as existence and uniqueness can be proved easily for the specific belief function that I use in the present paper:

The equilibrium cutoff with consistency can be calculated via the equilibrium condition

$$\hat{y}^*[\bar{E}(\hat{y}^*) - \underline{E}_r(\hat{y}^*)] = \hat{y}^*[\bar{E}_p(\hat{y}^*) - \underline{E}(\hat{y}^*)] \tag{1.6}$$

and note that the expressions on both sides also need to be equal to some $b > 0$, which rules out $\hat{y} = 0$ as an equilibrium cutoff. Hence, any equilibrium cutoff $\hat{y}^*$ must satisfy

$$\bar{E}(\hat{y}^*) - \underline{E}_r(\hat{y}^*) = \bar{E}_p(\hat{y}^*) - \underline{E}(\hat{y}^*). \tag{1.7}$$

Plugging in the functional form of the misperceptions, (1.4) and (1.5), and rearranging gives

$$\beta(1 - F(\hat{y}^*))(\bar{E}(\hat{y}^*) - \hat{y}^*) = \gamma F(\hat{y}^*)(\hat{y}^* - \underline{E}(\hat{y}^*))$$

and thus

$$\hat{y}^* = \frac{\beta(1 - F(\hat{y}^*))\bar{E}(\hat{y}^*) + \gamma F(\hat{y}^*)\underline{E}(\hat{y}^*)}{\beta(1 - F(\hat{y}^*)) + \gamma F(\hat{y}^*)}$$

which can be rewritten as

$$\hat{y}^* = \frac{a(1 - F(\hat{y}^*))\bar{E}(\hat{y}^*) + F(\hat{y}^*)\underline{E}(\hat{y}^*)}{a(1 - F(\hat{y}^*)) + F(\hat{y}^*)} \tag{1.8}$$

where $a = \beta/\gamma$.[8] An equiilibrium cutoff $\hat{y}^*$ must thus be a fixed point of the function

$$h(\hat{y}) = \frac{a(1 - F(\hat{y}^*))\bar{E}(\hat{y}^*) + F(\hat{y}^*)\underline{E}(\hat{y}^*)}{a(1 - F(\hat{y}^*)) + F(\hat{y}^*)}.$$

In Appendix 1.7.3 I prove that the function $h(\hat{y})$ has a unique fixed point and hence that

[7] The specific form of misperception that I use in this paper can be microfounded in the following way: People in the poor group only sometimes encounter a representative sample of the rich (e.g. if they go to the opera, watch a royal wedding or shop in a fancy store) and the rest of the time encounter only rich people who are very close to the cutoff (basically at $\hat{y}$), maybe because they are parents of their kids' school friends (upper-middle class families sometimes prefer to send their kids to state schools). However, people are not aware of this and therefore estimate average income as if they were observing a representative sample of the other group. The particular functional form of the bias can arise if the frequency of meeting a representative sample of the other group depends on the size of the own group, $F(\hat{y})$. This could be because "meeting a representative sample" does not actually require personal encounter but also comprises accounts from other people who are in one's own group. Then if people from different groups meet each other at a certain rate, the group with the bigger mass has a better understanding of the other group because people learn from others in their own group.

[8] $a > 0$ if both types are assumed to be naive to some degree, i.e. $\beta > 0$ and $\gamma > 0$. If one of the groups would be fully sophisticated, e.g. $\gamma = 0$, while the other group is naive, then consistency couldn't be satisfied for any (interior) cutoff. If both groups are fully sophisticated, i.e. $\beta = \gamma = 0$, the model turns into a standard model of unbiased sorting.

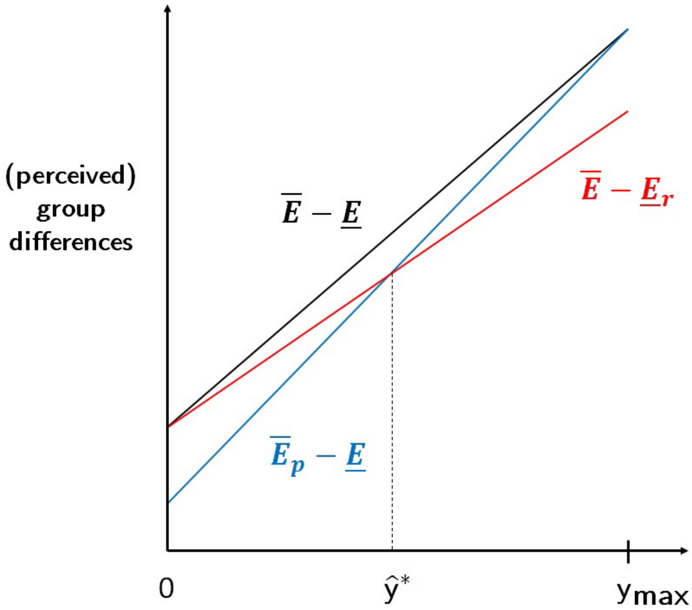

Figure 1.1: Perceived benefits of sorting for the rich (red) and poor (blue) and correct benefits of sorting as a function of the cutoff $\hat{y}$ (for a truncated lognormal distribution)

there always exists a unique biased sorting equilibrium cutoff $\hat{y}^*$. If $a = 1$ (and thus $\beta = \gamma$), (1.8) simplifies to $\hat{y}^* = E$ and the unique biased sorting equilibrium is such that the cutoff is exactly at the mean.

Proposition 1.1 *If $\underline{E}_r(\hat{y})$ and $\bar{E}_p(\hat{y})$ are defined according to (1.4) and (1.5), there exists a unique interior biased sorting equilibrium with consistency, and the unique equilibrium cutoff $\hat{y}^*$ is the fixed point of $h(\hat{y}) = \frac{a(1-F)\bar{E}+F\underline{E}}{a(1-F)+F}$ where $a = \beta/\gamma$. If $a = 1$, the unique cutoff is at $\hat{y}^* = E$.*

In Appendix 1.7.4 I analyze the relationship between naivety of the poor relative to the rich, a, and the equilibrium cutoff and show that the equilibrium cutoff $\hat{y}^*$ is increasing in a.

1.3.2 The consistency requirement

At this point it is instructive to look at the role of the consistency requirement in the model. The equilibrium condition in this specific example boils down to (1.6) and therefore the unique equilibrium cutoff $\hat{y}^*$ needs to satisfy

$$\bar{E}(\hat{y}^*) - \underline{E}_r(\hat{y}^*) = \bar{E}_p(\hat{y}^*) - \underline{E}(\hat{y}^*),$$

i.e. the perceived difference in group average incomes needs to be the same for both the rich and the poor group in equilibrium. Figure 1.1 depicts the perceived group differences (in terms of average income) of the poor group (blue) and the rich group (red) as well as the correct benefits of sorting of the person at the cutoff (black) as a function of the cutoff $\hat{y}$ (for a truncated log-normal income distribution). For small $\hat{y}$, the rich perceive the difference between the two groups almost correctly, while the poor underestimate it a lot. This is because of the assumption I make on the bias: the larger the part of the income distribution that a group sees, the less biased they are about the other group. This also implies that as $\hat{y}$ increases, the rich become more and more biased and the poor become more and more correct about the group difference. The blue and the red line cross at $\hat{y}^*$, the unique binary biased sorting equilibrium with consistency, where both groups have the same perceived benefits of sorting. As $\hat{y}$ increases beyond this point, the poor group starts to value sorting more than the rich group.

For a sorting equilibrium *without* consistency, the only condition that needs to be satisfied is that the cutoff is such that everybody in the rich group prefers being in the rich group to being in the poor group, while everybody in the poor group wants to stay in the poor group for some sorting fee $b > 0$. In Figure 1.1, all cutoffs $\hat{y}$ below $\hat{y}^*$ would satisfy this condition - if $\hat{y} \in (0, \hat{y}^*)$, the marginal person in the rich group values being in the rich group more than the marginal person in the poor group, and therefore we would be able to find a sorting fee $b > 0$ that the rich are willing to pay, while it doesn't seem worthwhile for the poor to do so. Hence, all $\hat{y} \in (0, \hat{y}^*]$ are binary biased sorting equilibria. Meanwhile, none of the $\hat{y}$ above $\hat{y}^*$ can be biased sorting equilibria, because the marginal person in the poor group would always be willing to pay more to join the rich group than the marginal person in the rich group, and thus no $b > 0$ could be found that separates the rich from the poor. Note however, that all $\hat{y} \in (0, \hat{y}^*)$, while constituting biased sorting equilibrium cutoffs, fail to satisfy the consistency requirement: Depending on the sorting fee (and note that the sorting fee is not unique if $\hat{y} \in (0, \hat{y}^*)$, any b between $\hat{y}(\bar{E}_p - \underline{E})$ and $\hat{y}(\bar{E} - \underline{E}_r)$ would work), either the people in the poor group would not understand why people at the bottom of the rich group want to pay b to be part of the rich group (because for the poor, being in the rich group is worth less), or the rich would wonder why people at the top of the poor group don't want to join their group, or both happens at the same time (if b is neither $\hat{y}(\bar{E}_p - \underline{E})$ nor $\hat{y}(\bar{E} - \underline{E}_r)$ but somewhere in between).

In the specific case analyzed here, the consistency requirement selects a unique equilibrium out of the range of sorting equilibria. This is because the misperceptions converge to the truth monotonically, which implies that the blue line approaches the black line monotonically as the cutoff increases, while the red line approaches the black line monotonically as $\hat{y}$ decreases. Therefore, the two lines can only cut once. If the misperceptions were not monotone, the distance between the black line and the blue resp. red line could be non-monotone, and therefore the blue and the red line could intersect several times. Each of those intersections would then constitute a biased sorting equilibrium with consistency. Consistency alone is not enough to guarantee uniqueness. Consistency and monotonicity of the misperceptions together do the job.

Another way to interpret the consistency requirement is a refinement to "no-learning partitions". If a partition satisfies consistency, then people never come across anything that goes against their beliefs and surprises them, therefore they have no impulse to modify their beliefs or their actions in any way.

I do *not* model any form of learning in this paper. I also do not make any assumptions about what happens if people encounter other people, whose choices they do not understand. One possibility is that people just assume that the others are wrong if they are puzzled by their choices, and do not modify their own beliefs or actions. Another possibility is that they start to question their own beliefs about the other groups and maybe try to update them, based on choices of other people that they observe. Alternatively, they might even experiment and join another group to learn about average income in that group. The consistency requirement restricts the set of biased sorting equilibria to those partitions where neither of the above happens, because people are simply not puzzled by anybody else's choices. In that sense, the consistency requirement can be viewed as a stability refinement: consistent equilibrium partitions are stable with respect to learning, experimenting or updating. Because what they see is consistent with their beliefs about the world, people have no incentives to question or change their beliefs, and thus the partition is stable irrespective of what they would do if they would encounter anything that is at odds with their beliefs.

1.4 Voting for Redistribution

Economic segregation can exacerbate inequalities in various ways. Schooling is one prominent example: If children living in affluent areas get better education than children from poor neighbourhoods because their local schools are of a better standard due to high local investment, income inequality in the next generation will be amplified. This effect is specifically pronounced in the United States, where school choice is linked to neighbourhood (see e.g. Chetty et al. (2014)). Moreover, having class mates from rich and influential families might not only have the direct effect on education via better quality of schooling, but might also yield benefits later in life through social connections that lead to better jobs and opportunities (see e.g. Savage (2015)).

In this section, I demonstrate that there might be another channel through which segregation can affect economic inequality: Economic segregation, if it leads to misperceptions of the income distribution, can have significant consequences for support for redistribution in society, and hence for (post-tax and post-redistribution) income inequality. I show that segregation leads poor people to underestimate what they can gain from redistribution and therefore to show less support for redistribution than if they would have perfect knowledge of the income distribution. Moreover, an increase in inequality (in the form of a mean-preserving spread of the income distribution) always leads to a smaller increase in perceived inequality and therefore in the demand for redistribution than if people were unbiased. The reason for this is that people with income below average fully observe the fall of low incomes, but do not fully see the offsetting increase of high incomes. Therefore, they think that average income has decreased. But because people's gains from redistribution depend positively on the difference between their own income and (perceived) average income, and both decrease if people are biased, demand for redistribution increases less than if people are unbiased and know that average income hasn't changed. I show that the increase in inequality can even be such that perceived inequality declines and therefore people's support for redistribution falls.

In the following analysis, I continue to use the functional forms of $\bar{E}_p(\hat{y})$ and $\underline{E}_r(\hat{y})$ as specified in (1.4) and (1.5), because this enables me to derive precise results. However, the general flavour of those results would not change if more general specifications of $\bar{E}_p(\hat{y})$ and $\underline{E}_r(\hat{y})$ were used that satisfy the conditions for existence of a unique equilibrium above the median, given in Appendix 1.7.2.

1.4.1 Inequality and the demand for redistribution

Suppose that everybody in the economy has to pay a proportional tax t and the government redistributes the proceeds equally among all its citizens afterwards. Hence, a person with pre-tax income of y_i has after-tax and after-redistribution income

$$(1-t)y_i + \tau(t)E,$$

where the function $\tau(t) \leq t$ accounts for the fact that there is a deadweight loss of taxation. (And let $\tau(.)$ be such that $\tau(t) > 0\ \forall t \in (0,1)$, $\tau(0) = 0$, $\tau''(t) \leq 0$, $\tau(1) = 0$, $\tau'''(t) \geq 0$ [this guarantees that $\tau'(t)$ is convex and hence also τ'^{-1} is convex, given that τ' is decreasing]).

Suppose furthermore that people vote to decide on the tax rate, and suppose that they care only about their own post-tax income.

Meltzer and Richard (1981) have examined the relationship between inequality and the demand for redistribution in this model: If people are unbiased about the income distribution, when voting for the redistribution rate a person with income y_i will simply choose the tax rate t that maximizes her post-tax income

$$(1-t)y_i + \tau(t)E.$$

As preferences are single-peaked in this case, the tax rate determined by majority voting will be the median earner's optimal tax rate given by

$$\tau'(t^*) = \frac{y^M}{E}$$

if $\frac{y^M}{E} \leq 1$ and $t^* = 0$ otherwise. As $\tau'(t)$ is decreasing in t, the decisive voter's optimal tax rate t^* is decreasing in the ratio between median and average income.

The ratio $\frac{y^M}{E}$ can be regarded as an, albeit rudimentary, measure of the degree of income equality in society. If the ratio is small, this means the difference between median and mean income is large and the income distribution has a large positive skew with a majority of people earning income below average and a few very rich people. Therefore, income equality is low and the demand for redistribution will be high in that case. If, on the other hand, the income distribution is almost symmetric, with most people being middle-class and only a few at the bottom and the top of the distribution, the equality ratio $\frac{y^M}{E}$ will be large (i.e. close to 1), and demand for redistribution will be low.

To analyze people's preferences for redistribution if they are biased, I need to establish what their perception of average income is: If people would correctly perceive both average income in their group and average income in the other group, they could simply calculate overall average income via the formula

$$E = F(\hat{y})\underline{E}(\hat{y}) + (1 - F(\hat{y}))\bar{E}(\hat{y})$$

for any cutoff $\hat{y}$.[9] However, if there is economic segregation and people are biased, then people misperceive average income in the other group, and hence they mis-estimate overall average income. Specifically, poor people think that average income is

$$E_p(\hat{y}) = F(\hat{y})\underline{E}(\hat{y}) + (1 - F(\hat{y}))\bar{E}_p(\hat{y}) < E.$$

Because they underestimate average income in the rich group,

$$\bar{E}_p(\hat{y}) < \bar{E}(\hat{y}),$$

they end up underestimating overall average income. Analogously, rich people overestimate average income,

$$E_r(\hat{y}) = F(\hat{y})\underline{E}_r(\hat{y}) + (1 - F(\hat{y}))\bar{E}(\hat{y}) > E.$$

Let me for simplicity of exposition assume henceforth that rich and poor people are equally naive, i.e. $\beta = \gamma$,[10] and remember that in this case the equilibrium cutoff will always be at average income E. This implies that the median earner is in the poor group (because the

[9] Note that I assume that people know the relative size of their respective group, i.e. they know $F(\hat{y})$ and $1 - F(\hat{y})$. They also know the range of the distribution and where the cutoff lies. They only misperceive the shape of the distribution function in the other group. With the type of bias that I examine here, their perceived income distribution in the other group is more skewed towards $\hat{y}$ compared to the actual distribution.

[10] The analysis can be done in a similar way for the general case of $\beta \neq \gamma$.

income distribution is positively skewed) and her preferred tax rate is given by

$$\tau'(\tilde{t}^*) = \frac{y^M}{E_p(E)} \text{ (or } \tilde{t}^* = 0 \text{ if } E_p(E) < y^M).$$

E_p is smaller than E, hence the median earner's perceived degree of equality as measured by $\frac{y^M}{E_p(E)}$ is higher than without segregation. Therefore, her optimal tax rate is lower in the presence of economic segregation.

Lemma 1.1 *In the model with segregation and misperceptions the median earner's preferred tax rate is lower compared to the model without misperceptions.*

For the remainder of this paper, I will assume that the following condition on the income distribution and people's naivity holds:

Assumption 1.1 *The distance between median and mean income is sufficiently high, such that*

$$\frac{E}{E_r(E)} \geq \frac{y^M}{E_p(E)}.$$

Remark 1.1 *In Appendix 1.7.6 I show that* $\frac{E}{E_r(E)} \geq \frac{y^M}{E_p(E)}$ *is guaranteed for misperceptions (1.4) and (1.5) if*

$$E - y^M \geq \beta \frac{\bar{E}(E) - \underline{E}(E)}{4}.$$

This condition holds if $E - y^M$ *is large enough compared to* $\bar{E}(E) - \underline{E}(E)$*, i.e. if the distribution is positively skewed but there is not too much mass at the tails of the distribution, and if* β *is small, i.e. people are not too biased.*[11]

Lemma 1.2 *If Assumption 1.7.6 holds, the median earner is the decisive voter.*

The preferred tax rate of the poorest person in the rich group (i.e. the person earning average income E) is given by

$$\tau'(t) = \frac{E}{E_r(E)}.$$

If the distance between median and mean income is sufficiently high, such that Assumption 1.1 holds, then this person will demand a lower tax rate than the median earner, and hence the median earner will be the decisive voter. As the median earner wants less redistribution than in the unbiased case, the tax rate selected by majority voting will be lower and therefore demand for redistribution in this segregated society will be lower than in a society without segregation and misperceptions.

Proposition 1.2 *The tax rate selected by majority voting in a segregated society where people misperceive the shape of the income distribution as described above is lower than in a society without segregation and misperception of the income distribution.*

Proof. See above. ■

1.4.2 The effect of changing inequality on demand for redistribution

In the following section I analyze what happens to people's (mis)perceptions and the support for redistribution in a segregated society if income inequality increases and how the effects

[11] Assumption 1.1 holds for positively skewed income distributions that look like actual income distributions that we observe in the real world, for example it holds for a truncated lognormal (on $(0, 10^8)$) with $\mu = 10.85$ and $\sigma = 0.85$ (the US household income distribution can be approximated by this function), and equally for a scaled down version of it, a truncated lognormal on $(0, 10)$ with $\mu = 0$ and $\sigma = 0.85$ (both times $\beta = 0.1$).

differ compared to a society without segregation. When analyzing the effect of an increase in inequality, it is important to clearly specify the exact form of this increase in inequality. Some changes in the shape of the income distribution are such that it cannot even be unequivocally decided whether they lead to an increase or decrease in inequality - different measures of inequality might yield different results. However, any mean-preserving spread of the income distribution always implies an increase in inequality, irrespective of the measure that is used, because it can be decomposed into (potentially infinitely many) transfers between rich and poor where money is transferred from a relatively poor to a relatively rich person. It therefore increases all measures of inequality that respect the *principle of transfers*, such as the Gini coefficient or the Theil index (see also Cowell (2000) and Dalton (1920)).[12] Hence, I will focus on the effect of a mean-preserving spread of the income distribution on group formation and demand for redistribution.

For simplicity, I require the mean-preserving spread to be such that the mass of people below and above the mean remain the same, but mass shifts from the middle towards the endpoints of the distribution, such that median income declines.[13] Specifically, I will analyze the effect of what I call a *monotone* mean-preserving spread of the income distribution, which is such that $\bar{E}(\hat{y})$ increases and $\underline{E}(\hat{y})$ decreases for any cutoff $\hat{y}$ (see Windsteiger (2017c)).[14] I will also require that the mean-preserving spread is such that $F(E)$ remains unchanged, and I require Assumption 1.1 to hold before and after the change in inequality. As this implies that the median earner is always the median voter, I will use these two expressions interchangeably.

In the absence of segregation and misperceptions, the median voter's support for redistribution increases due to a mean-preserving spread of the above described form, because median income declines relative to average income and hence the equality ratio $\frac{y^M}{E}$ decreases,

$$\Delta\left(\frac{y^M}{E}\right) = \frac{\Delta y^M}{E} = \frac{\Delta y^M}{y^M}\frac{y^M}{E},$$

i.e. the percentage change in $\frac{y^M}{E}$ is $\frac{\Delta y^M}{y^M}$ (where $\Delta y^M < 0$). This means that demand for redistribution, given by

$$\tau'(t^*) = \left(\frac{y^M}{E}\right),$$

increases. The increase in the median voter's optimal tax rate t^* is

$$\Delta t^* = \tau'^{-1}\left(\frac{y^M + \Delta y^M}{E}\right) - \tau'^{-1}\left(\frac{y^M}{E}\right).$$

In a segregated society, where people misperceive the shape of the income distribution, the effect of an increase in inequality on the support for redistribution depends on its impact on the location of the equilibrium cutoff $\hat{y}^*$, because this determines people's beliefs about the other group's average income. Recall that the equilibrium cutoff $\hat{y}^*$ is the fixed point of the function

$$h(\hat{y}) = \frac{a(1 - F(\hat{y}^*))\bar{E}(\hat{y}^*) + F(\hat{y}^*)\underline{E}(\hat{y}^*)}{a(1 - F(\hat{y}^*)) + F(\hat{y}^*)}.$$

As described in Section 1.3.1, $h(\hat{y})$ has a unique fixed point, which is at average income E if $a = 1$. Hence, the position of the equilibrium cutoff does not change due to a mean-preserving spread if $a = 1$.

[12] In the income and wealth inequality literature, an inequality measure is generally required to satisfy four properties: anonymity, scale independence, population independence and the principle of transfers. For an extensive discussion of different inequality measures see Cowell (2000).

[13] This implies that the distance between mean and median income increases.

[14] Such a mean-preserving spread can always be constructed if the initial distribution is strictly monotonic. The easiest way is to take mass from the middle of the distribution and add it to the endpoints 0 and $y_{\max}$ (in such a way that average income doesn't change).

What happens to perceived inequality and the demand for redistribution? As I explained in the previous section, if people are biased due to segregation, the median voter's optimal tax rate $\tilde{t}^*$ is characterized by the equation

$$\tau'(\tilde{t}^*) = \left(\frac{y^M}{E_p(E)}\right)$$

where $\tilde{t}^* < t^*$ (because $E_p < E$) - the median earner's preferred tax rate is lower under segregation because perceived equality $\frac{y^M}{E_p}$ is higher. While average income E does not change due to a mean-preserving spread, I show in Appendix 1.7.7 that average perceived income of the poor, E_p, declines. The poor feel that average income declines because they experience the decline of average income in their own group fully, but only partially take note of the compensating increase in average income among the rich. Hence, they think that society as a whole has become poorer. As a result, the change in the perceived equality ratio $\frac{y^M}{E_p}$ amounts to

$$\Delta\left(\frac{y^M}{E_p}\right) = \frac{\Delta y^M E_p - y^M \Delta E_p}{(E_p)^2} = \left(\frac{\Delta y^M}{y^M} - \frac{\Delta E_p}{E_p}\right)\frac{y^M}{E_p}$$

and thus the percentage decrease in $\frac{y^M}{E_p}$ is $\frac{\Delta y^M}{y^M} - \frac{\Delta E_p}{E_p}$, which is smaller (in absolute terms) than the percentage decrease of $\frac{y^M}{E}$ in the unbiased case, because $\frac{\Delta E_p}{E_p} < 0$.

Proposition 1.3 *If society is segregated, an increase in inequality (in the form of a monotone mean-preserving spread that keeps $F(E)$ constant) always leads to a smaller percentage increase in the median voter's perceived inequality than in the absence of segregation and misperception.*

Moreover, in Appendix 1.7.7 I demonstrate that one can always construct a mean-preserving spread that leads the median voter to believe that society has become more rather than less equal, i.e. that inequality has decreased rather than increased.

Proposition 1.4 *There exists an increase in inequality that causes a decrease of the median earner's perceived degree of inequality under segregation.*

The intuition for Proposition 1.4 is that, unlike in the non-segregated case, the median voter's perceived equality ratio $\frac{y^M}{E_p}$ can increase due to a mean preserving spread if people are biased, because both y^M and E_p decline. If the mean-preserving spread is such that the median voter's perceived degree of inequality decreases, as in Proposition 1.4, then also the median voter's demand for redistribution (i.e. her preferred tax rate) must necessarily decrease.

Corollary 1.2 *There always exists an increase in inequality such that the tax rate determined by majority voting decreases under segregation.*

In Appendix 1.7.7, I derive the condition on the mean-preserving spread that guarantees Proposition 1.4. As I explain above, this condition must ensure that the decline in E_p is larger than the decline in y^M. I also derive a weaker condition on the mean-preserving spread that guarantees that even if perceived inequality does not decrease, demand for redistribution increases less under segregation than without segregation. The step-by-step calculations in Appendix 1.7.7 can be summarized as follows: If perceived equality decreases due to a mean-preserving spread under segregation, the fact that the percentage decrease in perceived equality is smaller if society is segregated is not enough to guarantee that also the increase in demand for redistribution will be smaller than without segregation. There are two reasons for this: First, as perceived equality is higher to start with under segregation, a smaller percentage decrease does not automatically imply a smaller absolute decrease than in the absence of segregation. Second, even if the decrease in perceived equality is lower also in absolute terms, it is not clear whether the increase in demand for redistribution will be lower as well: this depends on the shape of the deadweight loss function $\tau(.)$. However, it turns out that the assumption that τ' is decreasing and convex is sufficient to ensure that demand for redistribution increases less under segregation if the absolute decrease in perceived equality is smaller than in the absence of segregation. The condition on the mean-preserving spread that guarantees that demand for redistribution under segregation increases by less if inequality increases compared to a situation without segregation is weaker than the condition that is needed for Proposition 1.4. In Appendix 1.7.9, I describe how more general changes in the shape of the income distribution affect demand for redistribution if society is segregated.

1.4.3 Inequality and the supply side of sorting

An alternative way to model the decline in perceived inequality after an increase in inequality is to assume that there is no segregation in place before the change (because whoever offers the sorting technology doesn't find it worthwhile) but then as inequality increases, offering the sorting technology becomes profitable and therefore society becomes segregated (and people become biased). I examine this in the following section for the case of a profit-maximizing monopolist.

Suppose a profit-maximizing monopolist, who has a fixed cost $c > 0$ of offering the sorting technology, can decide whether or not to become active.[15] Her profits from offering sorting are

$$\Pi(\hat{y}^*) = \hat{y}^*(\bar{E}(\hat{y}^*) - \underline{E}_r(\hat{y}^*))(1 - F(\hat{y}^*)) - c$$

Given that the equilibrium cutoff is at E and substituting for $\underline{E}_r$, this can be rewritten as

$$\Pi(E) = E(E - \underline{E}(E))[1 - \gamma F(E)(1 - F(E))] - c \tag{1.9}$$

Suppose that initially the income distribution is such that

$$E(E - \underline{E}(E))[1 - \gamma F(E)(1 - F(E))] - c < 0$$

and hence the monopolist prefers to stay out of the market. If inequality increases (again in the sense of a monotone mean-preserving spread of the income distribution which leaves $F(E)$ constant), $E - \underline{E}$ increases. This means that if the increase in inequality is sufficiently large, the profits from offering the sorting technology will become positive and the society will become segregated. Thus, a large enough increase in inequality will lead to economic segregation.

Lemma 1.3 *Suppose that the income distribution is initially such that a profit maximizing monopolist with fixed costs $c > 0$ does not find it profitable to offer the sorting technology. Then for any $c > 0$ there exists a mean-preserving spread of the income distribution such that the monopolist's profits become positive.*

Hence, I can compare the effect of increasing inequality in the presence of segregation to its effect without taking into account segregation (and the resulting misperception). As in the previous sections, I require Assumption 1.1 to be satisfied after the increase in inequality, to ensure that the median earner is the decisive voter.

If inequality increases and there is no segregation and people are unbiased, the median voter will demand more redistribution than before the change, because median income y^M is smaller

[15] In Appendix 1.7.11, I show that the argument works in the same way if a welfare-maximizing social planner decides about offering the sorting technology.

as a result of the mean-preserving spread, and hence also $\frac{y^M}{E}$ decreases:

$$\Delta\left(\frac{y^M}{E}\right) = \frac{\Delta y^M}{E} < 0$$

Therefore, the median earner's demand for redistribution increases from

$$\tau'^{-1}\left(\frac{y^M}{E}\right)$$

to

$$\tau'^{-1}\left(\frac{\dot{y}^M}{E}\right),$$

where $\dot{y}^M = y^M + \Delta y^M < y^M$ is median income after the increase in inequality.
If the increase in inequality leads to economic segregation and hence causes people to be biased, then the median voter's demand for redistribution changes from

$$\tau'^{-1}\left(\frac{y^M}{E}\right)$$

to

$$\tau'^{-1}\left(\frac{\dot{y}^M}{E_p(E)}\right),$$

where

$$E_p(E) = E - \beta(1 - F(E))^2(\bar{E}(E) + \Delta\bar{E}(E) - E).$$

As $E_p < E$, the increase in the median voter's demand for redistribution will be smaller than in the absence of economic segregation.

Proposition 1.5 *If an increase in inequality leads to economic segregation, the median voter's demand for redistribution will increase less than in the absence of segregation.*

In Appendix 1.7.8 I show that I can always construct a mean preserving spread of the income distribution such that demand for redistribution decreases under segregation.

Proposition 1.6 *There exists an increase in inequality that causes economic segregation and leads to a decline in the tax rate determined by majority voting.*

Apart from the mean-preserving spread described above there are also other types of increases in inequality that would make it profitable for the monopolist to offer one cutoff. I demonstrate in Appendix 1.7.5 that for the lognormal distribution an increase in the log-variance σ (which corresponds to an increase in the Gini-coefficient but is a median-preserving instead of a mean-preserving spread) also leads to an increase in the monopolist's profits (1.9).

1.5 Empirical Evidence

In February 2016, I conducted an online survey on 600 US citizens above the age of 18. The survey was distributed via Amazon Mechanical Turk and the original questionnaire can be accessed at `https://lse.ut1.qualtrics.com/jfe/form/SV_eDLNkeGfQg2ycM5`. A description of the sample (i.e. respondents' characteristics) can be found in Appendix 1.8.[16] The advantages and potential pitfalls of using Amazon Mechanical Turk in academic research have been discussed by Kuziemko et al. (2015) in their Online Appendix. I summarize some of their points and document my own experiences in Appendix 1.8.1.

[16] The data and all do-files are available upon request.

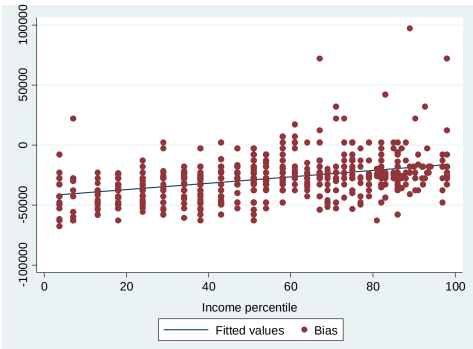

Figure 1.2: People's estimate of average income is increasing in their own income (Bias = correct average income - perceived average income)

By conducting this survey, I wanted to address two main questions:

1. Is there evidence that people misperceive the income distribution in the way I assume in the application of my theoretical model of sorting with misperceptions to the question of demand for redistribution? For example, do poor people underestimate overall average income and do rich people overestimate it?

2. Are people with a diverse social circle (i.e. people who are not very "segregated") less biased?

To tackle the first question, I asked people about their own household income and their estimate of average US household income. Figure 1.2 plots the relationship between the two: It turns out that, in general, both rich and poor people underestimate mean household income on average. However, people's estimate of average household income is increasing in their own income. This is roughly in line with my model, which would predict that poor people underestimate average income (because they know average income in their group and underestimate average income in the rich group) and rich people overestimate average income (because they know their own average income and overestimate the poor group's income).
The first attempt to identify a link between segregation and misperception is to look at the relationship between the degree of income segregation that a respondent lives in and (the absolute value of) her bias. For this purpose, I match the survey data with county-level income segregation data computed by Chetty et al. (2014). However, I do not find any relationship between county-level income segregation and a respondent's absolute level of bias. I suspect that county-level data is too coarse to be useful as a proxy for an individual's degree of segregation. Unfortunately, I cannot repeat the analysis with a more precise measure of income segregation because I have neither lower-level locational information about my respondents, nor data on lower-level income segregation in the US.
However, I also tried to elicit respondents' individual degrees of segregation by asking about the diversity of their social interactions. In particular, I asked them about their friends and colleagues, and how many of them have similar respectively different levels of household income and education. Then I employed a scale from 0 to 4 to classify respondents as more or less segregated (4 indicating the highest possible degree of segregation) concerning those social circles, depending on how similar their work colleagues respectively friends are to themselves. Subsequently, I used factor analysis to identify a common factor out of these categorical response variables (for detailed explanations see Appendix 1.8).
I find that the severity of misperception of average income is correlated with the degree of social segregation: poor people tend to underestimate average household income less and rich people

Table 1.1: Regression results for social segregation as measured by factor analysis

	Bias
Income percentile	0.004***
	(0.001)
(Income percentile) x (Social segregation)	0.002**
	(0.001)
Social segregation	-0.073
	(0.060)
Intercept	-0.598***
	(0.041)
N	592

p-values in parentheses
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

tend to overestimate it less if their social circle is more diverse. Table 1.1 shows the results of regressing people's bias about average income (in percentage terms, where a positive bias means average income is overestimated) on their own income percentile, the degree of social segregation as measured by common factor identified by factor analysis and the interaction between own income percentile and the factor: Misperceptions of average household income are less severe for respondents with more diverse social circles.
Furthermore, I asked the so-called "Lin position generator" question in the version of the "Great British Class calculator"[17], which is the short version of a similar question asked in the Great British Class Survey (see Savage (2015)).[18] This question tries to identify the diversity of the respondent's social circle by asking whether she socially knows people with certain occupations (eighteen different occupations), ranging from chief executive to cleaner. I measure diversity of the social circle by assigning to each of the occupations their status rank using the Cambridge Social Interaction and Stratification (CAMSIS) scale score (where low numbers correspond to high rank) and then calculating for each respondent the standard deviation of all the scores of occupations she knows: the higher this standard deviation, the more diverse can the respondent's social circle be assumed to be. Regressing the absolute value of people's misperception of average income in percentage terms (variable *Bias2*) on the standard deviation yields significant results and the coefficient has the expected sign: A more diverse social circle corresponds to less bias about average household income (see Table 1.2).

1.6 Conclusion

In the present paper I have showed how the model of sorting in the presence of misperceptions that is analyzed in detail in Windsteiger (2017b) can be applied to the issue of income inequality and preferences for redistribution: If people are segregated according to income, there will be less demand for redistribution in society. Furthermore, an increase in inequality will lead to a smaller increase in support for redistribution than in the absence of segregation, and certain mean-preserving spreads of the income distribution can even lead to a decrease in demand for redistribution, because they result in a decline in perceived inequality.
Finally, I have reported some of my empirical findings on misperception of the shape of the income distribution and segregation: I have showed evidence that people's estimate of average household income is increasing in their own income, and that people's misperceptions are more

[17] see `http://www.bbc.co.uk/news/magazine-22000973`
[18] The question is named after the sociologist Nan Lin who developed it in the 1980s.

Table 1.2: Regression results for social diversity as measured by CAMSIS score standard deviation

	(1) Bias2	(2) Bias2
Social circle status diversity	-0.0107***	-0.00916**
	(0.005)	(0.015)
Income percentile		-0.00181***
		(0.000)
Intercept	0.483***	0.568***
	(0.000)	(0.000)
N	592	592

p-values in parentheses
$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

severe, the more socially segregated they are.
My approach shows that modelling segregation and belief formation simultaneously can yield interesting and unexpected results and offers new perspectives on issues such as income inequality and redistribution. In the present paper, I have used the model to examine the implications of segregation and biased beliefs on redistributive demand, but the general framework presented in Windsteiger (2017b) offers itself to a wide set of applications related to segregation, such as education policy and housing.

1.7 Appendix A: Theoretical Appendix

1.7.1 Consistency and monotonicity

Without imposing the consistency requirement, also non-monotone partitions can be biased sorting equilibria (if the belief function is of a certain form): Suppose that $y_1 \in S_b$ and $y_2 \in S_0$ with $y_1 < y_2$. In order for the partition $[S_0, S_b]$ to constitute a biased sorting equilibrium, it must be the case that

$$y_1 E_b[S_0] \leq y_1 E[S_b] - b$$

and

$$y_2 E[S_0] \geq y_2 E_0[S_b] - b.$$

(Notation: $E_i[S_j]$ is group S_i's belief about average income in S_j.) Combined, these two conditions give

$$y_2 E_0[S_b] - y_2 E[S_0] \leq b \leq y_1 E[S_b] - y_1 E_b[S_0].$$

It is immediate to see that whether this inequality can hold depends on the belief function, because even though $y_1 < y_2$, the misperceptions $E_0[S_b]$ and $E_b[S_0]$ could be defined in such a way that

$$y_1 E[S_b] - y_1 E_b[S_0] \geq y_2 E_0[S_b] - y_2 E[S_0].$$

However, the consistency requirement rules out non-monotone equilibrium partitions for any belief function.

Proposition 1.7 *All biased sorting equilibria with consistency satisfy monotonicity.*

Proof. Suppose a non-monotone equilibrium exists. Then it must be the case that there exist $y_1 \in S_b$ and $y_2 \in S_0$ with $y_1 < y_2$. Then the IC constraint for y_1 requires that

$$y_1 E_b[S_0] \leq y_1 E[S_b] - b$$

and note that this implies that $E[S_b] - E_b[S_0] > 0$. The consistency requirement additionally requires that

$$y_2 E_b[S_0] \geq y_2 E[S_b] - b.$$

But these two conditions combined give

$$y_1 E[S_b] - y_1 E_b[S_0] \geq y_2 E[S_b] - y_2 E_b[S_0],$$

which cannot hold for any belief function B if $y_1 < y_2$, because as noted above $E[S_b] - E_b[S_0] > 0$. ■

1.7.2 Conditions for a unique equilibrium above the median with linear utility

Proposition 1.8 *(Windsteiger (2017b)) If the belief function is such that the rich overestimate average income of the poor group, and the poor underestimate average income of the rich group, such that*

$$\underline{E}_r(\hat{y}) > \underline{E}(\hat{y}) \quad \forall \hat{y} \in [0, y_{\max}) \tag{1.10}$$

and

$$\bar{E}(\hat{y}) < \bar{E}_p(\hat{y}) \quad \forall \hat{y} \in (0, y_{\max}], \tag{1.11}$$

a binary biased sorting equilibrium with consistency always exists. If additionally the severity of the misperceptions is monotone in the cutoff, i.e.

$$\frac{d(\bar{E}(\hat{y}) - \bar{E}_p(\hat{y}))}{d\hat{y}} < 0 \quad \text{and} \quad \frac{d(\underline{E}_r(\hat{y}) - \underline{E}(\hat{y}))}{d\hat{y}} > 0 \quad \forall \hat{y} \in (0, y_{\max}) \tag{1.12}$$

the biased sorting equilibrium with consistency is unique.

Proof. Conditions (1.10) and (1.11) together with Assumption 1 and the fact that $\underline{E}_r(\hat{y})$, $\underline{E}(\hat{y})$, $\bar{E}_p(\hat{y})$ and $\bar{E}(\hat{y})$ are continuous ensure existence. Condition (1.12) implies that people's misperceptions converge to the truth monotonically as $\hat{y}$ goes to 0 resp. $y_{\max}$ and hence there will be a unique $\hat{y}^*$ for which both groups have the same belief about the difference in average incomes (and thus about the benefits of sorting). For more explanations see Windsteiger (2017b). ■

Proposition 1.9 *If both groups underestimate inequality, sufficient conditions for a unique equilibrium cutoff $\hat{y}^*$ above the median are conditions (1.10), (1.11) and (1.12) and additionally*

$$\bar{E}_p(y^M) + \underline{E}_r(y^M) < 2E.$$

Proof. The first three conditions guarantee existence and uniqueness (see above). Concerning the last condition, note that if $\bar{E} - \bar{E}_p$ is monotonically increasing and $\underline{E}_r - \underline{E}$ is monotonically decreasing in $\hat{y}$, then

$$\bar{E}_p(\hat{y}) - \underline{E}(\hat{y}) < \bar{E}(\hat{y}) - \underline{E}_r(\hat{y})$$

for all $\hat{y}$ below the unique equilibrium cutoff, and the inequality must hold in the other direction above the unique equilibrium cutoff. That implies

$$\bar{E}_p(\hat{y}) + \underline{E}_r(\hat{y}) < \bar{E}(\hat{y}) + \underline{E}(\hat{y})$$

for all $\hat{y}$ below the equilibrium cutoff, and

$$\bar{E}_p(\hat{y}) + \underline{E}_r(\hat{y}) > \bar{E}(\hat{y}) + \underline{E}(\hat{y})$$

for all $\hat{y}$ above the equilibrium cutoff. If the equilibrium should lie above the median, then at the median it must be the case that

$$\bar{E}_p(y^M) + \underline{E}_r(y^M) < \bar{E}(y^M) + \underline{E}(y^M),$$

because the median must be below the cutoff. The fact that

$$E = (1 - F(y^M))\bar{E}(y^M) + F(y^M)\underline{E}(y^M) = \frac{\bar{E}(y^M) + \underline{E}(y^M)}{2}$$

at the median proves the claim. ■

1.7.3 Analysis of the unique binary biased sorting equilibrium

As established in Section 1.3.1, any equilibrium cutoff is characterized by

$$\hat{y}^* = \frac{a(1 - F(\hat{y}^*))\bar{E}(\hat{y}^*) + F(\hat{y}^*)\underline{E}(\hat{y}^*)}{a(1 - F(\hat{y}^*)) + F(\hat{y}^*)} \tag{1.13}$$

and hence it is the fixed point of

$$h(\hat{y}) = \frac{a(1 - F(\hat{y}))\bar{E}(\hat{y}) + F(\hat{y})\underline{E}(\hat{y})}{a(1 - F(\hat{y})) + F(\hat{y})}$$

Therefore, the equilibrium cutoff is exactly where the 45 degree line cuts the function h. As $\hat{y}^*$ approaches 0, the left hand side of (1.13) becomes zero, while the right hand side becomes $h(0) = E$, and hence larger than the left hand side. As $\hat{y}^*$ approaches $y_{\max}$, the opposite happens: the left hand side becomes $y_{\max}$, and thus larger than the right hand side, which is again $h(y_{\max}) = E$. Hence, because the expressions on both sides are continuous in $\hat{y}$, we know that there must be a $\hat{y}$ in $(0, y_{\max})$ for which equality holds. This concludes the proof that an equilibrium cutoff always exists in my model.
To ensure that there can only be one such intersection point, I can calculate

$$h'(\hat{y}) = \frac{\left[\begin{array}{c} \left(-af(\hat{y})\bar{E}(\hat{y}) + a(1 - F(\hat{y}))\frac{\partial \bar{E}(\hat{y})}{\partial \hat{y}} + f(\hat{y})\underline{E}(\hat{y}) + F(\hat{y})\frac{\partial \underline{E}(\hat{y})}{\partial \hat{y}}\right)(a(1 - F(\hat{y})) + F(\hat{y})) \\ -\left(a(1 - F(\hat{y}))\bar{E}(\hat{y}) + F(\hat{y})\underline{E}(\hat{y})\right)(-af(\hat{y}) + f(\hat{y})) \end{array}\right]}{(a(1 - F(\hat{y})) + F(\hat{y}))^2}$$

which can be simplified to

$$h'(\hat{y}) = \frac{(1 - a)f(\hat{y})}{(a(1 - F(\hat{y})) + F(\hat{y}))^2}\left[a(1 - F(\hat{y}))(\hat{y} - \bar{E}(\hat{y})) + F(\hat{y})(\hat{y} - \underline{E}(\hat{y}))\right].$$

This implies that h has a local extremum or saddle point $\hat{y}^{**}$ characterized by

$$a(1 - F(\hat{y}^{**}))(\hat{y}^{**} - \bar{E}(\hat{y}^{**})) + F(\hat{y}^{**})(\hat{y}^{**} - \underline{E}(\hat{y}^{**})) = 0$$

or equivalently

$$\hat{y}^{**} = \frac{a(1-F(\hat{y}^{**}))\bar{E}(\hat{y}^{**}) + F(\hat{y}^{**})\underline{E}(\hat{y}^{**})}{a(1-F(\hat{y}^{**})) + F(\hat{y}^{**})} \tag{1.14}$$

This is exactly the equation that characterizes the equilibrium cutoff and the fixed point of h, i.e. we find that $\hat{y}^{**} = \hat{y}^*$. Whenever the 45 degree line cuts h it must therefore be where the slope of h is 0. This means that at any intersection, the 45 degree line cuts h from below, which implies that such an intersection can only happen once. It follows that h will have a unique fixed point and the equilibrium cutoff is unique.

The fixed point of h characterized by (1.14) (or equivalently (1.13)) is a local maximum if $a > 1$ and a local minimum if $a < 1$. This can be seen from noting that

$$h''(\hat{y}) = \frac{(1-a)f'(\hat{y})}{(a(1-F(\hat{y}))+F(\hat{y}))^2}\left[a(1-F(\hat{y}))(\hat{y}-\bar{E}(\hat{y})) + F(\hat{y})(\hat{y}-\underline{E}(\hat{y}))\right]$$
$$+\frac{(1-a)f(\hat{y})}{(a(1-F(\hat{y}))+F(\hat{y}))}$$
$$-\frac{2(1-a)^2 f^2(\hat{y})\left[a(1-F(\hat{y}))(\hat{y}-\bar{E}(\hat{y})) + F(\hat{y})(\hat{y}-\underline{E}(\hat{y}))\right]}{(a(1-F(\hat{y}))+F(\hat{y}))^3}.$$

At $\hat{y}^*$ we know that

$$a(1-F(\hat{y}^*))(\hat{y}^*-\bar{E}(\hat{y}^*)) + F(\hat{y}^*)(\hat{y}^*-\underline{E}(\hat{y}^*)) = 0$$

and thus the first and the third term drop out of the second derivative and we get

$$h''(\hat{y}^*) = \frac{(1-a)f(\hat{y}^*)}{(a(1-F(\hat{y}^*))+F(\hat{y}^*))}.$$

As this expression is negative for a larger than 1 and positive for a smaller than 1, $\hat{y}^*$ is a local maximum if $a > 1$ and a local minimum at $a < 1$. Figures 1.3 and 1.4 depict the intersection of h and the 45 degree line for $a < 1$ and $a > 1$ (where the underlying income distribution is a truncated lognormal distribution). If $a = 1$ the problem becomes very simple, as the expression for h reduces to

$$h(\hat{y}) = E,$$

i.e. h is just a horizontal straight line at E and the unique equilibrium cutoff is at E.

1.7.4 The relationship between naivety and the equilibrium cutoff $\hat{y}^*$

As noted in Section 1.3.1, the equilibrium cutoff depends on the naivety of the rich and the poor via a single parameter, $\frac{\beta}{\gamma} = a$, which describes the severity of the poor's naivety relative to the rich's. If $a = 1$ then both groups are "equally naive", if $a > 1$ then the poor are more naive than the rich. Using the equilibrium condition

$$\hat{y}^* = \frac{a(1-F(\hat{y}^*))\bar{E}(\hat{y}^*) + F(\hat{y}^*)\underline{E}(\hat{y}^*)}{a(1-F(\hat{y}^*)) + F(\hat{y}^*)}, \tag{1.15}$$

I can investigate how $\hat{y}^*$ changes with a:

$$(1-F(\hat{y}^*))\bar{E}(\hat{y}^*)da +$$
$$\left(-af(\hat{y}^*)\bar{E}(\hat{y}^*) + a(1-F(\hat{y}^*))\frac{\bar{E}(\hat{y}^*)-\hat{y}^*}{1-F(\hat{y}^*)}f(\hat{y}^*) + f(\hat{y}^*)\underline{E}(\hat{y}^*) + F(\hat{y}^*)\frac{(\hat{y}^*-\underline{E}(\hat{y}^*))}{F(\hat{y}^*)}f(\hat{y}^*)\right)d\hat{y}^*$$
$$= (a(1-F(\hat{y}^*)) + F(\hat{y}^*) + \hat{y}^*(-af(\hat{y}^*) + f(\hat{y}^*)))d\hat{y}^* + (1-F(\hat{y}^*))\hat{y}^* da$$

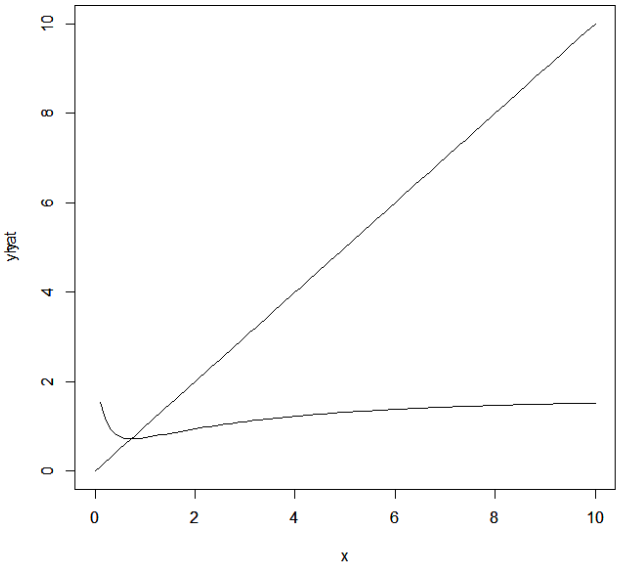

Figure 1.3: Equilibrium cutoff $\hat{y}^*$ if $a < 1$

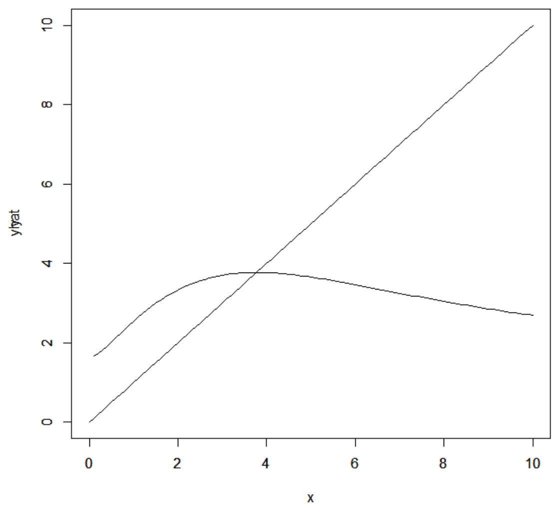

Figure 1.4: Equilibrium cutoff $\hat{y}^*$ if $a > 1$

$$\iff (1-F)(\bar{E}-\hat{y}^*)da = [af\bar{E} - a(\bar{E}-\hat{y}^*)f - \underline{E}f - (\hat{y}^* - \underline{E})f$$
$$+a(1-F) + F + \hat{y}^* f(1-a)]d\hat{y}^*$$
$$\iff \frac{d\hat{y}^*}{da} = \frac{(1-F(\hat{y}^*))(\bar{E}(\hat{y}^*)-\hat{y}^*)}{a(1-F(\hat{y}^*))+F(\hat{y}^*)} > 0 \tag{1.16}$$

The equilibrium cutoff $\hat{y}^*$ is increasing in the degree of naivety of the poor relative to the rich. The higher a, the more the poor tend to underestimate the benefits of sorting (relative to the rich) and hence the more they need to see of the whole distribution relative to the rich to have the same perceived benefits of sorting as the rich.

As naivety goes to zero, what happens to the equilibrium cutoff depends on the speed of convergence of β respectively γ. If β converges to zero faster than γ, a goes to zero and $\hat{y}^*$ goes to 0. If γ converges at a faster speed than β, a converges to infinity and the equilibrium cutoff goes to $y_{\max}$.[19]

1.7.5 A median-preserving spread of the lognormal distribution and monopolist profits

Recall that the monopolist's profits from offering one cutoff (which in equilibrium will be at E if $a = 1$) can be written as

$$E(E-\underline{E})[1-\gamma F(E)(1-F(E))] - c$$

For the lognormal distribution, this becomes

$$\Pi = E\left[E\left(1-\frac{\Phi\left(\frac{\ln(E)-\mu}{\sigma}-\sigma\right)}{\Phi\left(\frac{\ln(E)-\mu}{\sigma}\right)}\right)\left(1-\gamma\Phi\left(\frac{\ln(E)-\mu}{\sigma}\right)+\gamma\left[\Phi\left(\frac{\ln(E)-\mu}{\sigma}\right)\right]^2\right)\right]-c$$

which can be simplified to

$$\Pi = E\left[E\left(1-\frac{1-\Phi\left(\frac{\sigma}{2}\right)}{\Phi\left(\frac{\sigma}{2}\right)}\right)\left(1-\gamma\Phi\left(\frac{\sigma}{2}\right)+\gamma\left[\Phi\left(\frac{\sigma}{2}\right)\right]^2\right)\right]-c$$
$$= E^2\left(\frac{2\Phi\left(\frac{\sigma}{2}\right)-1}{\Phi\left(\frac{\sigma}{2}\right)}\right)\left(1-\gamma\Phi\left(\frac{\sigma}{2}\right)+\gamma\left[\Phi\left(\frac{\sigma}{2}\right)\right]^2\right)-c$$

because $\ln\hat{y} = \mu+\sigma^2$ if $\hat{y} = E$.

I find that

$$\frac{d\Pi}{d\sigma} = 2\sigma E^2\left(\frac{2\Phi\left(\frac{\sigma}{2}\right)-1}{\Phi\left(\frac{\sigma}{2}\right)}\right)\left(1-\gamma\Phi\left(\frac{\sigma}{2}\right)+\gamma\left[\Phi\left(\frac{\sigma}{2}\right)\right]^2\right)$$
$$+E^2\left[\frac{\phi(\frac{\sigma}{2})\frac{1}{2}}{\Phi\left(\frac{\sigma}{2}\right)}\right]\left(1-\gamma\Phi\left(\frac{\sigma}{2}\right)+\gamma\left[\Phi\left(\frac{\sigma}{2}\right)\right]^2\right)$$
$$+E^2\left(\frac{2\Phi\left(\frac{\sigma}{2}\right)-1}{\Phi\left(\frac{\sigma}{2}\right)}\right)\gamma\phi\left(\frac{\sigma}{2}\right)\left(\Phi\left(\frac{\sigma}{2}\right)-\frac{1}{2}\right)$$

As $\Phi\left(\frac{\sigma}{2}\right) > \frac{1}{2}$, all of the terms are positive and hence the monopolist's profit always increases if σ increases.

Proposition 1.10 *If income is lognormally distributed, an increase in inequality in the form of a median-preserving spread increases the monopolist's revenues from offering the sorting*

[19] The best way to see the latter is to introduce the auxiliary parameter $b = \frac{\gamma}{\beta}$ in this case and rewrite $h(\hat{y})$ in terms of b.

technology.

1.7.6 Sufficient conditions for Assumption 1.1

$$\frac{y^M}{E_p} \leq \frac{E}{E_r}$$
$$\iff y^M(F\underline{E}_r + (1-F)\bar{E}) \leq E(F\underline{E} + (1-F)\bar{E}_p)$$

If $\beta = \gamma$, this can be simplified to

$$\beta(y^M F^2(E - \underline{E}) + E(1-F)^2(\bar{E} - E)) \leq E(E - y^M)$$

Noting that
$$E - \underline{E} = (1-F)(\bar{E} - \underline{E})$$

and
$$\bar{E} - E = F(\bar{E} - \underline{E})$$

I can further simplify to

$$\beta F(1-F)\left(F\frac{y^M}{E} + (1-F)\right)(\bar{E} - \underline{E}) \leq (E - y^M)$$

Given that $F(1-F) < 0.25$ (because $y^M < E$) and $\frac{y^M}{E} < 1$, I have that

$$\beta F(1-F)\left(F\frac{y^M}{E} + (1-F)\right)(\bar{E} - \underline{E}) < \beta\frac{(\bar{E} - \underline{E})}{4} \tag{1.17}$$

and it follows that
$$\beta\frac{(\bar{E} - \underline{E})}{4} \leq E - y^M$$

is a sufficient condition for
$$\frac{y^M}{E_p} \leq \frac{E}{E_r}$$

(in fact it is even a sufficient condition for $\frac{y^M}{E_p} < \frac{E}{E_r}$, given that inequality (1.17) is strict).

1.7.7 Detailed calculations for Section 1.4.2

Average income E does not change due to a mean-preserving spread and hence[20]

$$\Delta E = F\Delta\underline{E} + (1-F)\Delta\bar{E} = 0, \tag{1.18}$$

Average perceived income of the poor, E_p, declines, because

$$\Delta E_p = F\Delta\underline{E} + (1-F)\Delta\bar{E}_p$$

and
$$\bar{E}_p(\hat{y}) = \beta(1-F)\hat{y} + (1-\beta(1-F))\bar{E}$$

which implies
$$\Delta\bar{E}_p(E) = (1-\beta(1-F))\Delta\bar{E} < \Delta\bar{E} \tag{1.19}$$

[20] And note that I require the mean-preserving spread to be such that $F(\hat{y}^*) = F(E)$ doesn't change.

(as $\hat{y} = E$ doesn't change). The change in $\frac{y^M}{E_p}$ amounts to

$$\Delta\left(\frac{y^M}{E_p}\right) = \frac{\Delta y^M E_p - y^M \Delta E_p}{(E_p)^2} = \left(\frac{\Delta y^M}{y^M} - \frac{\Delta E_p}{E_p}\right)\frac{y^M}{E_p}$$

and thus the percentage change in $\frac{y^M}{E_p}$ is $\frac{\Delta y^M}{y^M} - \frac{\Delta E_p}{E_p}$, which is smaller (in absolute terms) than the percentage change of $\frac{y^M}{E}$ in the unbiased case, because $\frac{\Delta E_p}{E_p} < 0$. In the following I show that if $\left|\frac{\Delta E_p}{E_p}\right|$ is large enough relative to $\left|\frac{\Delta y^M}{y^M}\right|$, the median earner will even think that inequality has decreased, i.e. the percentage change in $\frac{y^M}{E_p}$ (and hence also the absolute change in $\frac{y^M}{E_p}$) can be positive:
From (1.18) and (1.19) it follows that

$$\Delta E_p(E) = -(1-F)\Delta\bar{E} + (1-F)\Delta\bar{E}_p(E) = -\beta(1-F)^2\Delta\bar{E}(E)$$

Furthermore,

$$E_p(E) = F\underline{E}(E) + (1-F)\bar{E}_p(E) = E - \beta(1-F)^2(\bar{E}(E) - E)$$

and therefore

$$\frac{\Delta E_p}{E_p} = \frac{-\beta(1-F)^2\Delta\bar{E}}{E - \beta(1-F)^2(\bar{E}-E)} = \frac{\beta(1-F)F\Delta\underline{E}}{E - \beta(1-F)^2(\bar{E}-E)}$$

(using (1.18) again). Hence, I get

$$\frac{\Delta y^M}{y^M} - \frac{\Delta E_p}{E_p} > 0 \iff \frac{\Delta y^M}{y^M} > \frac{\beta(1-F)F\Delta\underline{E}}{E - \beta(1-F)^2(\bar{E}-E)}$$

$$\iff \frac{\Delta y^M}{\Delta\underline{E}} < \frac{\beta y^M(1-F)F}{E - \beta(1-F)^2(\bar{E}-E)} \tag{1.20}$$

(where both sides are positive). For a given $\frac{\Delta y^M}{\Delta\underline{E}}$ this condition is more likely to be satisfied if β is large, because

$$\frac{\partial}{\partial\beta}\left(\frac{\beta}{E - \beta(1-F)^2(\bar{E}-E)}\right) = \frac{E}{\left[E - \beta(1-F)^2(\bar{E}-E)\right]^2} > 0$$

and hence the RHS is increasing in β. Furthermore, $(1-F(E))F(E)$ should not be too small, i.e. the income distribution cannot be too positively skewed, such that $F(E)$ is not too far above 0.5. Note however, that such a monotone mean-preserving spread can be constructed for any given income distribution, by ensuring that Δy^M and $\Delta\underline{E}$ are such that (1.20) holds. To see this, note that we need to ensure that Assumption 1.1 is satisfied before and after the mean-preserving spread. A sufficient condition for this is that

$$\beta\frac{(\bar{E}(E) - \underline{E}(E))}{4} \leq E - y^M \tag{1.21}$$

holds (see Section 1.7.6). If the mean-preserving spread would be such that $\Delta y^M = 0$ (i.e. all the mass shifts around below and above the median, but the median stays the same), then (1.20) would be satisfied. If this mean-preserving spread is such that $\Delta\underline{E}$ (and corresponding $\Delta\bar{E}$) are small enough (in absolute value) and therefore Assumption 1.1 still holds after the change[21], this mean-preserving spread would lead to a decrease in demand for redistribution

[21] By continuity, such a mean-preserving spread can always be found because if (1.21) is satisfied initially

(whereas in the unbiased case demand for redistribution would not change because neither y^M nor E has changed). Because all the expressions are continuous, it follows that we can analogously construct a mean-preserving spread that satisfies (1.21) and has $\Delta y^M > 0$ (and where thus demand for redistribution increases in the unbiased case, but decreases in the presence of misperceptions). Hence, I can conclude that

Lemma 1.4 *For any $\beta > 0$ there exists a mean-preserving spread of the income distribution such that an increase in inequality leads to a decrease in the median earner's perceived degree of inequality.*

Now let me examine the absolute change of $\frac{y^M}{E}$ and $\frac{y^M}{E_p}$: I want to derive sufficient conditions for the absolute decrease in perceived equality to be smaller under segregation, i.e.

$$\Delta\left(\frac{y^M}{E}\right) < \Delta\left(\frac{y^M}{E_p}\right) \tag{1.22}$$

(because both sides of this inequality are negative). Lemma 1.4 shows that I can always construct a mean-preserving spread satisfying (1.20) such that perceived equality $\frac{y^M}{E_p}$ increases under segregation (in which case inequality (1.22) trivially holds, because $\frac{y^M}{E}$ will always decrease). However, less strong conditions can be derived in order for (1.22) to hold without perceived inequality having to decrease:

$$\Delta\left(\frac{y^M}{E}\right) < \Delta\left(\frac{y^M}{E_p}\right) \iff$$

$$\frac{\Delta y^M}{E} < \frac{\Delta y^M E_p - y^M \Delta E_p}{(E_p)^2} = \frac{\Delta y^M}{E_p} - \frac{y^M \Delta E_p}{(E_p)^2}$$

$$\iff \frac{\Delta y^M}{E} < \frac{\Delta y^M}{E - \beta(1-F)^2(\bar{E}-E)} + \frac{y^M \beta(1-F)^2 \Delta\bar{E}}{\left(E - \beta(1-F)^2(\bar{E}-E)\right)^2}$$

$$\iff \Delta y^M \left(\frac{1}{E} - \frac{1}{E - \beta(1-F)^2(\bar{E}-E)}\right) < \frac{y^M \beta(1-F)^2 \Delta\bar{E}}{\left(E - \beta(1-F)^2(\bar{E}-E)\right)^2}$$

$$\iff \frac{\Delta y^M}{y^M} \left(\frac{-\beta(1-F)^2(\bar{E}-E)}{E(E - \beta(1-F)^2(\bar{E}-E))}\right) < \frac{\beta(1-F)^2 \Delta\bar{E}}{\left(E - \beta(1-F)^2(\bar{E}-E)\right)^2}$$

$$\iff \frac{\Delta y^M}{y^M} \left(\frac{-(\bar{E}-E)}{E}\right) < \frac{\Delta\bar{E}}{E - \beta(1-F)^2(\bar{E}-E)}$$

$$\iff \frac{\Delta y^M}{y^M} \left(\frac{-F(\bar{E}-\underline{E})}{E}\right) < \frac{\frac{-F\Delta\underline{E}}{1-F}}{E - \beta(1-F)^2(\bar{E}-E)}$$

$$\iff -\frac{\Delta y^M}{y^M} \left(\frac{(1-F)(\bar{E}-\underline{E})}{E}\right) < \frac{-\Delta\underline{E}}{E - \beta(1-F)^2(\bar{E}-E)}$$

$$\iff \frac{\Delta y^M}{\Delta\underline{E}} \left(\frac{E-\underline{E}}{E}\right) < \frac{y^M}{E - \beta(1-F)^2(\bar{E}-E)}$$

$$\iff \frac{\Delta y^M}{\Delta\underline{E}} < \frac{y^M E}{\left(E - \beta(1-F)^2(\bar{E}-E)\right)(E-\underline{E})} = \frac{y^M E}{E_p(E-\underline{E})}$$

For a given mean-preserving spread, this inequality is more likely to hold if β is large (such that E_p is small relative to E). Note however, that it is always possible to construct a mean-preserving spread that satisfies this inequality, by designing Δy^M and $\Delta\underline{E}$ accordingly.

---

then $\frac{y^M}{E_p}$ is strictly smaller than $\frac{E}{E_r}$ (see Section 1.7.6) and hence a small change in $\bar{E}$ and $\underline{E}$ will still leave $\frac{y^M}{E_p} \leq \frac{E}{E_r}$.

Lemma 1.5 *The (absolute) decrease in* $\frac{y^M}{E_p}$ *is smaller than the (absolute) decrease in* $\frac{y^M}{E}$ *iff the mean-preserving spread is such that*

$$\frac{\Delta y^M}{\Delta \underline{E}} < \frac{y^M E}{\left(E - \beta(1-F)^2(\bar{E} - E)\right)(E - \underline{E})}. \tag{1.23}$$

In the absence of segregation, the change in the median earner's preferred tax rate due to a mean-preserving spread is given by[22]

$$\Delta t^* = \tau'^{-1}\left(\frac{\dot{y}^M}{E}\right) - \tau'^{-1}\left(\frac{y^M}{E}\right).$$

If society is segregated, the change in the median earner's preferred tax rate amounts to[23]

$$\Delta \tilde{t}^* = \tau'^{-1}\left(\frac{\dot{y}^M}{\dot{E}_p}\right) - \tau'^{-1}\left(\frac{y^M}{E_p}\right).$$

If the conditions of Lemma (1.5) hold, the decrease in $\frac{y^M}{E_p}$ is smaller than the decrease in $\frac{y^M}{E}$. Furthermore, I know that $\frac{y^M}{E_p} > \frac{y^M}{E}$. Together with the fact that $\tau''(t) \leq 0$ and $\tau'''(t) \geq 0$, which implies that τ'^{-1} is decreasing and convex, this gives

$$\Delta \tilde{t}^* < \Delta t^*.$$

Lemma 1.6 *If the mean-preserving spread is such that (1.23) holds, the increase in the preferred tax rate is less in a segregated society than in the absence of segregation.*

1.7.8 Detailed calculations for Section 1.4.3

If a mean-preserving spread leads to economic segregation, the median earner's demand for redistribution declines if

$$\begin{aligned} & \frac{y^M}{E} < \frac{\dot{y}^M}{E_p} \\ \iff & y^M(E - \beta(1-F)^2(\bar{E} + \Delta\bar{E} - E)) < y^M E + E\Delta y^M \\ \iff & \frac{\beta(1-F)^2(\bar{E} + \Delta\bar{E} - E)}{E} > -\frac{\Delta y^M}{y^M} = \left|\frac{\Delta y^M}{y^M}\right| \end{aligned} \tag{1.24}$$

For a given mean-preserving spread this inequality holds if β is large enough (i.e. people are sufficiently naive) and the increase in average income in the rich group is large enough relative to the decline in median income. Again, it is immediate to see that a mean-preserving spread satisfying (1.24) can always be constructed by designing $\Delta\bar{E}$ and Δy^M accordingly (the proof is analogous to the proof of Lemma 1.4 in Section 1.7.7).

1.7.9 The effect of general changes in the shape of the income distribution on the demand for redistribution if society is segregated

What happens to people's preferred redistribution rate if inequality between groups changes when people are already segregated? First and foremost this depends on how this change

[22] Notation: $\dot{y}^M$ denotes median income after the mean-preserving spread.

[23] Notation: $\dot{y}^M$ denotes median income after the mean-preserving spread and $\dot{E}_p$ denotes the poor group's perception of average income after the mean-preserving spread.

affects the equilibrium cutoff $\hat{y}^*$. Recall that the equilibrium cutoff is given by

$$\hat{y}^* = \frac{a(1-F(\hat{y}^*))\bar{E}(\hat{y}^*) + F(\hat{y}^*)\underline{E}(\hat{y}^*)}{a(1-F(\hat{y}^*)) + F(\hat{y}^*)}, \tag{1.25}$$

i.e. $\hat{y}^*$ is the fixed point of the function

$$h(\hat{y}) = \frac{a(1-F(\hat{y}))\bar{E}(\hat{y}) + F(\hat{y})\underline{E}(\hat{y})}{a(1-F(\hat{y})) + F(\hat{y})}.$$

As described in Section 1.3.1 and Appendix 1.7.3, (1.25) has a unique fixed point. If $a = 1$, this fixed point is at average income E. For $a < 1$ the intersection between $h(\hat{y})$ and the 45 degree line looks like Figure 1.3, if $a > 1$ then it looks like Figure 1.4 (and if $a = 1$, $\hat{y}^*$ is where the 45 degree line intersects with the horizontal line at E). From these graphs it is immediate to see that the impact of an increase in inequality on the equilibrium cutoff depends on how this increase in inequality affects $h(\hat{y})$ (and thus the intersection of the 45 degree line with $h(\hat{y})$).

If $\underline{E}$ goes down while $F(\hat{y}^*)$ and $\bar{E}(\hat{y}^*)$ stay the same, $h(\hat{y})$ shifts down, and the intersection with the 45 degree line (= the equilibrium cutoff $\hat{y}^*$) goes down (both if $a > 1$ and if $a < 1$). Hence, the new equilibrium cutoff will be lower. The opposite happens if $\bar{E}$ goes up ceteris paribus, i.e. if the rich group gets richer on average: Then it is straightforward to see from (1.25) that the new equilibrium cutoff will be higher.

Suppose that both things happen, so $\bar{E}$ increases, while $\underline{E}$ decreases (while $F(\hat{y}^*)$ doesn't change). Then whether the new equilibrium cutoff is higher or lower than the old one depends on a and $F(\hat{y}^*)$ (resp. $1 - F(\hat{y}^*)$): if a is high, or $1 - F(\hat{y}^*)$ is high, such that $a(1-F)\Delta\bar{E} + F\Delta\underline{E} > 0$, then the new equilibrium cutoff will be higher, if a and/or $1 - F$ is low, then the new equilibrium cutoff will be lower. If $a = 1$ (meaning both groups are equally naive) then the cutoff is always E and hence will go down if E decreases due to this increase in inequality. E decreases if $F(E)$ is high and $(1 - F(E))$ is low, a feature that characterizes unequal distributions with positive skew.

If $\underline{E}$ decreases by $\Delta\underline{E}$ ceteris paribus, then as I have argued above, $\hat{y}^*$ will go down. What happens to preferences for redistribution depends on the position of y^M: If $a = 1$, the equilibrium cutoff is always at E, hence $y^M < \hat{y}^* = E$ before and after the decline in $\underline{E}$. If y^M is sufficiently below the cutoff, such that preferences for redistribution do not overlap (i.e. if Assumption 1.1 is satisfied at all times) the median earner is the decisive voter both before and after the change in $\underline{E}$.[24] Under these circumstances, a decrease in $\underline{E}$ and subsequently in $\hat{y}^*$ will mean that $\underline{E}$ decreases by $\Delta\underline{E} + \frac{\partial \underline{E}}{\partial \hat{y}}d\hat{y}^*$, $\bar{E}$ decreases by $\frac{\partial \bar{E}}{\partial \hat{y}}d\hat{y}^*$ and E decreases by $F\Delta\underline{E}$ (the decreases in $\underline{E}$ and $\bar{E}$ due to the decrease in $\hat{y}^*$ cancel out with changes in F and $1-F$ and do not affect E: clearly, where the cutoff is has no implications for average income). The decrease in $\underline{E}$ and subsequent fall in $\hat{y}^*$ will lead to a decrease in $\bar{E}_p$ for two reasons: because $\bar{E}$ decreases due to the decline in $\hat{y}^*$ and because as $\hat{y}^*$ decreases, the poor become more biased, i.e. $\bar{E} - \bar{E}_p$ increases. As $\bar{E}_p$ decreases, clearly also E_p decreases, and if the poor are sufficiently biased, then this can lead to a situation where the perceived equality ratio $\frac{y^M}{E_p}$ does not decrease (as the true equality ratio unambiguously will), but instead increases, because E_p decreases by more than y^M:

$$d\left(\frac{y^M}{E_p}\right) = \frac{\Delta y^M E_p - y^M dE_p}{(E_p)^2} = \left(\frac{\Delta y^M}{y^M} - \frac{dE_p}{E_p}\right)\frac{y^M}{E_p}$$

[24] By "preferences for redistribution do not overlap" I mean that the median earner should be sufficiently far away from the cutoff such that the person in the rich group with income just at the cutoff wants lower redistribution than the median earner.

It can be shown that

$$dE_p = F\Delta\underline{E} + (1-F)\beta[(1-F) + f(\bar{E} - \hat{y})]d\hat{y}^*$$

and hence

$$|dE_p| > |F\Delta\underline{E}| = |dE|\,.$$

This implies that

$$\left|\frac{dE_p}{E_p}\right| > \left|\frac{dE}{E}\right|$$

(because we also have that $E_p < E$) and therefore the percentage decline in the perceived equality ratio, $\frac{\Delta y^M}{y^M} - \frac{dE_p}{E_p}$, will always be smaller than the percentage decline in the true equality ratio, $\frac{\Delta y^M}{y^M} - \frac{dE}{E}$. Moreover, the decline in $\underline{E}$ can be such that

$$\frac{\Delta y^M}{y^M} - \frac{dE}{E} < 0$$

while

$$\frac{\Delta y^M}{y^M} - \frac{dE_p}{E_p} > 0$$

because E_p decreases by more than y^M. In fact it can be calculated that this will be the case if the decline in $\underline{E}$ is such that

$$\frac{\Delta y^M}{F\Delta\underline{E} + (1-F)\beta[(1-F) + f(\bar{E} - \hat{y})]d\hat{y}^*} < \frac{y^M}{E_p}.$$

Note that if $\underline{E}$ decreases ceteris paribus we have that

$$d\hat{y}^* = \frac{F\Delta\underline{E}}{a(1-F) + F}$$

This can be deduced from taking the total derivative of (1.25) which yields

$$d\hat{y}^* = h'(\hat{y}^*)d\hat{y}^* + \frac{F\Delta\underline{E}}{a(1-F) + F}$$

and noting that $h'(\hat{y}^*) = 0$ (see Appendix 1.7.3). Hence, if the decline in $\underline{E}$ is such that

$$\frac{\Delta y^M}{F\Delta\underline{E}\left(1 + \frac{(1-F)\beta[(1-F)+f(\bar{E}-\hat{y})]}{a(1-F)+F}\right)} < \frac{y^M}{E_p}$$

the true equality ratio $\frac{y^M}{E}$ decreases, while the perceived equality ratio $\frac{y^M}{E_p}$ increases - a change in inequality that leads to a decrease of $\frac{y^M}{E}$ if people are unbiased, will lead to an increase in $\frac{y^M}{E_p}$ in the biased case. Therefore, if $\underline{E}$ decreases, the new preferred tax rate after this increase in inequality can be lower than before. An increase in inequality can lead to a decrease in the demand for redistribution due to people's biased perception of the average income change in the other group and the change in group composition, which affects people's bias.
Suppose that instead of $\underline{E}$ decreasing, $\bar{E}$ increases by $\Delta\bar{E}$. Then the above analysis yields that $\hat{y}^*$ must increase - an initial increase in $\bar{E}$ by $\Delta\bar{E}$ means that the new equilibrium cutoff of the biased sorting equilibrium has to be higher. This implies that the total increase in $\bar{E}$ will be the sum of the shift $\Delta\bar{E}$ and the effect on $\bar{E}$ due to an increase in $\hat{y}^*$:

$$d\bar{E} = f\frac{(\bar{E} - \hat{y})}{(1-F)}d\hat{y}^* + \Delta\bar{E} \tag{1.26}$$

Furthermore, also $\underline{E}$ increases due to the change in the cutoff. Hence, I have

$$dE_p = \left[f\underline{E} + F\frac{\partial \underline{E}}{\partial \hat{y}}\right] d\hat{y}^* + (1-F)d\bar{E}_p - f\bar{E}_p d\hat{y}^*$$

Remember that $\bar{E}_p = E - \beta(1-F)(\bar{E} - \hat{y}^*)$ and therefore

$$d\bar{E}_p = d\bar{E} - \beta(1-F)d\bar{E} + \beta(1-F)(\bar{E} - \hat{y}^*)d\hat{y}^* \tag{1.27}$$

Using (1.26) and (1.27), I get

$$dE_p = (1-F)\Delta\bar{E} - \beta(1-F)^2\Delta\bar{E} + (1-F)\beta[(1-F) + f(\bar{E} - \hat{y})]d\hat{y}^* \tag{1.28}$$

and hence E_p increases in this case (both $\Delta\bar{E}$ and $d\hat{y}^*$ are positive here). Because $\hat{y}^*$ increases, the poor group is getting larger and therefore less biased, which means E_p gets closer to E and thus increases for two reasons: because E increases, and because the poor become less biased and underestimate average income by less.
The fact that both E and E_p increase implies that the perceived equality ratio $\frac{y^M}{E_p}$ will move in the same direction as the true equality ratio $\frac{y^M}{E}$, namely it will decrease due to an increase in the denominator (note that the numerator y^M doesn't change in this case because only the part of the income distribution that lies above $\hat{y}^*$ changes if $\bar{E}$ increases ceteris paribus). Under certain conditions, the percentage decrease in perceived equality can even be larger than the percentage decrease in true equality. The percentage decrease in perceived equality amounts to

$$\frac{d\left(\frac{y^M}{E_p}\right)}{\frac{y^M}{E_p}} = -\frac{dE_p}{E_p}$$

whereas the percentage decrease in true equality is

$$\frac{d\left(\frac{y^M}{E}\right)}{\frac{y^M}{E}} = -\frac{\Delta E}{E}.$$

In order for the percentage decrease in perceived inequality to be higher (in absolute value) we need

$$\frac{dE_p}{E_p} > \frac{\Delta E}{E} \tag{1.29}$$

which can be rewritten as

$$\frac{(1-F)\Delta\bar{E} - \beta(1-F)^2\Delta\bar{E} + (1-F)\beta[(1-F) + f(\bar{E} - \hat{y})]d\hat{y}^*}{E - \beta(1-F)^2(\bar{E} - \hat{y}^*)} > \frac{(1-F)\Delta\bar{E}}{E}$$

Using

$$d\hat{y}^* = h'(\hat{y}^*)d\hat{y}^* + \frac{a(1-F)\Delta\bar{E}}{a(1-F)+F} = \frac{a(1-F)\Delta\bar{E}}{a(1-F)+F}$$

(because $h'(\hat{y}^*) = 0$) this becomes, after simplifying,

$$\frac{f(\bar{E} - \hat{y}^*)a - F}{a(1-F)+F} > -\frac{(1-F)(\bar{E} - \hat{y}^*)}{E}.$$

Therefore, whether or not (1.29) holds depends on the parameters of the model and the distribution function. A sufficient condition for this to hold is that

$$\frac{f(\bar{E} - \hat{y}^*)a - F}{a(1-F)+F} > 0$$

which can be simplified to

$$af(\bar{E} - \hat{y}^*) > F. \tag{1.30}$$

For a given distribution function, this condition is more likely to be satisfied for large a. If (1.29) holds for $a = 1$ then (due to $\hat{y}^* = E$ in this case), median income y^M is below the equilibrium cutoff both before and after the change in $\bar{E}$. If Assumption 1.1 is satisfied and hence the median earner is the decisive voter, the demand for redistribution will increase by more than in the unbiased case if $\frac{y^M}{E_p}$ decreases by more than $\frac{y^M}{E}$.

If $\underline{E}$ decreases and $\bar{E}$ increases at the same time, the change in the equilibrium cutoff is given by

$$d\hat{y}^* = \frac{F\Delta\underline{E} + a(1-F)\Delta\bar{E}}{a(1-F)+F}$$

and hence whether $\hat{y}^*$ increases or decreases depends on the sign of $F\Delta\underline{E} + a(1-F)\Delta\bar{E}$. If $a(1-F)\bar{E} + F\underline{E}$ decreases[25], $\hat{y}^*$ goes down. The change in the perceived equality ratio $\frac{y^M}{E_p}$ amounts to

$$d\left(\frac{y^M}{E_p}\right) = \frac{dy^M E_p - y^M dE_p}{(E_p)^2} = \left(\frac{\Delta y^M}{y^M} - \frac{dE_p}{E_p}\right)\frac{y^M}{E_p}$$

where

$$E_p = F\underline{E} + (1-F)\bar{E}_p.$$

If both $\bar{E}$ and $\underline{E}$ change, then

$$dE_p = F\Delta\underline{E} + \left(f\underline{E} + F\frac{\partial\underline{E}}{\partial\hat{y}}\right)d\hat{y} + (1-F)d\bar{E}_p - f\bar{E}_p d\hat{y}$$

$$= F\Delta\underline{E} + (1-F)(1-\beta(1-F))\Delta\bar{E} + (1-F)\beta[(1-F) + f(\bar{E}-\hat{y})]d\hat{y}^* \tag{1.31}$$

Suppose that $F\Delta\underline{E} + a(1-F)\Delta\bar{E}$ is negative but $F\Delta\underline{E} + (1-F)\Delta\bar{E}$ is positive (implying that $a < 1$), such that $\hat{y}^*$ decreases due to an increase in $\bar{E}$ and a decrease in $\underline{E}$, and average income E increases. Suppose also that y^M decreases. Then the true equality ratio decreases, because the numerator decreases and the denominator increases:

$$d\left(\frac{y^M}{E}\right) = \left(\frac{\Delta y^M}{y^M} - \frac{\Delta E}{E}\right)\frac{y^M}{E} < 0$$

The change in the perceived equality ratio is given by

$$d\left(\frac{y^M}{E_p}\right) = \left(\frac{\Delta y^M}{y^M} - \frac{dE_p}{E_p}\right)\frac{y^M}{E_p}$$

(and assume that Assumption 1.1 holds before and after the change, such that the perceived equality ratio determines redistribution). (1.31) implies that E_p will increase by less than E. The percentage decline in the perceived equality ratio, $\frac{\Delta y^M}{y^M} - \frac{dE_p}{E_p}$ can therefore - for certain changes $F\Delta\underline{E} + a(1-F)\Delta\bar{E} < 0$ (but $F\Delta\underline{E} + (1-F)\Delta\bar{E} > 0$) - be smaller than the percentage decline in the true equality ratio, $\frac{dy^M}{y^M} - \frac{\Delta E}{E}$. Moreover, the change in the shape of the income distribution can be such that $d\left(\frac{y^M}{E_p}\right)$ is positive, and hence the demand for redistribution can go down as inequality increases.

If $a = 1$ then the equilibrium cutoff goes down if average income decreases and goes up if

[25] Note that if $a < 1$ (i.e. the poor are less naive than the rich) I can have that $a(1-F)\bar{E} + F\underline{E}$ decreases, while $E = (1-F)\bar{E} + F\underline{E}$ stays constant. An increase in inequality while E stays constant is probably the closest to reality that this model can get, as I have not modelled growth here. If I would have modelled growth, then this increase in inequality where $\underline{E}$ decreases and $\bar{E}$ increases while E stays constant would translate to $\underline{E}$ constant and $\bar{E}$ increasing while E increases, which is probably what has happened over the last 30 years in the US and Europe. I have refrained from modelling growth here, because this would just have complicated the analysis ($\hat{y}^*$ would have a time trend etc.) while not changing the results about existence, uniqueness etc.

average income increases due to the change in inequality. If $\Delta\bar{E} = -\Delta\underline{E}$ then E decreases iff $F(E) > 1 - F(E)$, i.e. if the income distribution is positively skewed. In that case, the percentage decrease in the perceived equality ratio is smaller (in absolute value) than the percentage decrease in the true equality ratio iff

$$\left(\frac{\Delta y^M}{y^M} - \frac{dE_p}{E_p}\right) > \left(\frac{\Delta y^M}{y^M} - \frac{\Delta E}{E}\right)$$

$$\iff \frac{dE_p}{E_p} < \frac{\Delta E}{E}$$

$$\iff \frac{F\Delta\underline{E} + (1-F)(1-\beta(1-F))\Delta\bar{E} + (1-F)\beta[(1-F) + f(\bar{E}-\hat{y})]\Delta E}{E - \beta(1-F)^2(\bar{E}-E)} < \frac{\Delta E}{E}$$

$$\iff \frac{\Delta E - \beta(1-F)^2\Delta\bar{E} + (1-F)\beta[(1-F) + f(\bar{E}-\hat{y})]\Delta E}{E - \beta(1-F)^2(\bar{E}-E)} < \frac{\Delta E}{E}$$

$$\iff \frac{1 - \beta(1-F)^2\frac{\Delta\bar{E}}{\Delta E} + (1-F)\beta[(1-F) + f(\bar{E}-\hat{y})]}{E - \beta(1-F)^2(\bar{E}-E)} > \frac{1}{E}$$

$$\iff E - \beta(1-F)^2\frac{\Delta\bar{E}}{\Delta E}E + (1-F)\beta[(1-F) + f(\bar{E}-\hat{y})]E > E - \beta(1-F)^2(\bar{E}-E)$$

$$\iff \frac{(\bar{E}-E)}{E} + [1 + \frac{f(\bar{E}-\hat{y})}{(1-F)}] > \frac{\Delta\bar{E}}{\Delta E}$$

This inequality always holds, because the fraction on the RHS is negative. Hence, this type of increase in inequality always leads to a smaller increase in demand for redistribution if people are biased, compared to the unbiased case.

Conclusion 1.1 *The effect of increasing inequality on support for redistribution if society is already segregated depends on the nature of the increase in inequality and on the rich and the poor's relative degree of naivety (resp. on a).*

- *If $a = 1$ and $\underline{E}(\hat{y}^*)$ decreases ceteris paribus, then the equilibrium cutoff will go down. This leads to a change in the composition of the two groups in society, and, because the poor group is getting smaller, to an increase in poor people's bias - $\frac{E_p}{E}$ will decrease. As described above, this means that even though people in the poor group have become poorer relative to the rich, because they misperceive average income more after the change in inequality, their perceived equality ratio might not have decreased by much, or might even have increased. Hence, whether support for redistribution increases or decreases in this case depends on the poor's degree of naivety and on how much the median income decreases due to the increase in inequality. In any case, even if the change in inequality is such that the demand for redistribution increases, the increase is smaller than what would be expected in the framework of the Meltzer-Richard Model.*

- *If $a = 1$ and $\bar{E}(\hat{y}^*)$ increases ceteris paribus, then the equilibrium cutoff will go up. This leads to a change in the composition of the two groups in society, and, because the poor group is getting larger, to a decrease in poor people's bias - $\frac{E_p}{E}$ will increase. However, if the income distribution is sufficiently unequal such that the median earner is the decisive voter, the median voter's preferred tax rate will still be smaller than in the absence of segregation and misperceptions. However, the observed increase in support for redistribution might be larger if people are biased, because as $\bar{E}(\hat{y}^*)$ increases demand for redistribution increases for two reasons: the median voter is getting poorer relative to the average, and the median voter is becoming less biased and hence more aware of the prevailing inequality. While the first effect is larger if people are unbiased, the second* *effect is only present if people are biased, and together, the two effects might lead to a larger increase than in the absence of a bias.*

- *If $a = 1$ and both $\underline{E}(\hat{y}^*)$ decreases and $\bar{E}(\hat{y}^*)$ increases, the change in support for redistribution depends on whether the equilibrium cutoff increases or decreases. If $\Delta\bar{E} = -\Delta\underline{E}$, the equilibrium cutoff decreases if the income distribution is positively skewed. In this case the increase in support for redistribution will again be smaller than in the absence of misperceptions and we might even observe a decrease in support for redistribution.*

Remark 1.2 *I do not have growth in my model, but my analysis would work in the same way if all variables would grow at a constant rate. In a model with growth, the case of $\bar{E}$ increasing and $\underline{E}$ decreasing would be translated into a situation where $\bar{E}$ increases a lot, while $\underline{E}$ stays constant (or increases only by a small rate), and we would see a decrease in the size of the poor group (corresponding to a decline in $\hat{y}^*$ with zero growth) if the distribution is sufficiently positively skewed. As Saez and Zucman (2016) point out, this constellation of high income growth of the rich accompanied by negligible growth rates of the bottom percentiles of the income distribution, is exactly what occurred during the past decades (at least in the US). Hence, my model can explain why, while inequality was increasing in the US over the past decades, people were, at least in the beginning, not demanding higher redistribution rates in response (if anything, then they were demanding lower redistribution rates, as documented by Kuziemko et al. (2015), who analyze the evolution of preferences for redistribution in the General Social Survey (GSS)).*

1.7.10 Overestimating Inequality: Existence and uniqueness of equilibrium

In the following section I will analyze misperceptions which are such that the poor people think average income in the rich group is higher than it actually is, while the rich people underestimate average income in the poor group, which implies that both groups overestimate inequality.[26] Let me specifically assume that the belief function is such that:

$$\bar{E}_p(\hat{y}) = \beta(1 - F(\hat{y}))y_{\max} + (1 - \beta(1 - F(\hat{y}))\bar{E}(\hat{y}) \tag{1.32}$$

and

$$\underline{E}_r(\hat{y}) = \gamma F(\hat{y})0 + (1 - \gamma F(\hat{y}))\underline{E}(\hat{y}). \tag{1.33}$$

Analogous to Section 1.3.1, β and γ parameterize the "naivity" of the poor and the rich respectively, and if β respectively γ is 0 agents have no misperceptions. The functional form of $\bar{E}_p$ and $\underline{E}_r$ implies that the misperceptions are more severe, the smaller the part of the distribution they can fully observe. It is straightforward to see that $\bar{E}_p(\hat{y}) \geq \bar{E}(\hat{y})$ and $\underline{E}_r(\hat{y}) \leq \underline{E}(\hat{y})$ $\forall \hat{y} \in Y$. The misperceptions converge to the truth monotonically, and therefore the sufficient conditions for existence and uniqueness of an interior sorting as stated in Windsteiger (2017b) are satisfied.

The equilibrium condition becomes

$$\hat{y}^* \left[\bar{E}(\hat{y}^*) - \underline{E}_r(\hat{y}^*)\right] = \hat{y}^*[\bar{E}_p(\hat{y}^*) - \underline{E}(\hat{y}^*)]$$

[26] Perhaps consumption of unrepresentative media could lead to such a bias: poor people watch "Celebrity Reality Shows" such as "Keeping up with the Kardashians" and conclude that rich people are very rich, while the rich read horror stories about deprivation in poor families and low standards of state schools.

and the unique interior equilibrium cutoff satisfies

$$\bar{E}(\hat{y}^*) - \underline{E}_r(\hat{y}^*) = \bar{E}_p(\hat{y}^*) - \underline{E}(\hat{y}^*).$$

Plugging in the functional form of the misperceptions, (1.32) and (1.33), yields

$$\beta(1 - F(\hat{y}^*))(y_{\max} - \bar{E}(\hat{y}^*)) = \gamma F(\hat{y}^*)\underline{E}(\hat{y}^*). \tag{1.34}$$

This equation indirectly characterizes $\hat{y}^*$.[27]
If poor and rich people are equally naive, then $\beta = \gamma$ and equation (1.34) simplifies to

$$1 - F(\hat{y}^*) = \frac{E}{y_{\max}}.$$

In this case, it is immediate to see that the equilibrium is unique, as the RHS does not vary with $\hat{y}$ and F is strictly increasing and hence there will be only one $\hat{y}^*$ that satisfies this equation. Moreover, the cutoff is decreasing in $\frac{E}{y_{\max}}$, i.e. it is higher the larger the difference is between maximum and average income. Furthermore, the equilibrium cutoff must lie above the median, because the income distribution is positively skewed and hence $\frac{E}{y_{\max}}$ will be smaller than $\frac{1}{2}$ and therefore $F(\hat{y}^*)$ must be larger than $\frac{1}{2}$. As in the case where people underestimate inequality, also here the equilibrium cutoff only depends on a, not on β and γ individually.

Application: Housing and education

Suppose that a city is segregated into two groups, rich and poor, and both groups overestimate inequality. This would imply that in rich neighborhoods, the average income of the poor (and hence the average benefit of mixing with them) is underestimated and hence the rich are willing to pay more to segregate from the poor than their actual benefit from sorting. (Equally, the poor are also willing to pay more to mix with the rich than in the unbiased model). This implies that for example housing prices in rich neighborhoods (if this is what we interpret the sorting fee b to correspond to) would be exaggeratedly high, or that fees for private schools are very high.[28] How exaggerated these prices are depends on the degree of naivety of the poor versus the rich and the shape of the income distribution, as these two factors determine the cutoff $\hat{y}^*$ and hence the sorting fee b and the severity of the misperceptions. The sorting fee b is given by

$$b = \hat{y}^* \left[\bar{E} - \underline{E}_r\right]$$

[27] Existence is confirmed by seeing that as $\hat{y} \to 0$ the LHS goes to $\beta(y_{\max} - E)$ whereas the LHS goes to 0, while at $\hat{y} \to y_{\max}$ the LHS (0) is smaller than the RHS (γE). As the expressions on both sides are continuous functions of $\hat{y}$, there must be a cutoff $\hat{y}^* \in (0, y_{\max})$ such that both sides are equal. Rewriting equation (1.34) using $a = \frac{\beta}{\gamma}$ I get

$$F\underline{E} + a(1 - F)\bar{E} = a(1 - F)y_{\max} \tag{1.35}$$

To confirm that there can only be one $\hat{y}^*$ satisfying this equation, I employ a single-crossing argument: Determine the slope of the LHS and the RHS by taking the derivative with respect to $\hat{y}$ on both sides. This yields

$$f\hat{y} - af\hat{y}$$

for the LHS and

$$-afy_{\max}$$

for the RHS. Clearly, for any $\hat{y}$ we have

$$f(\hat{y} - a\hat{y}) > -afy_{\max}$$

(because f is a pdf and therefore always positive). At any point $\hat{y}$ the slope of the LHS is larger than the slope of the RHS. This holds both when $a < 1$ (in which case the LHS is increasing in y, while the RHS is decreasing) and when $a > 1$ (in which case both sides are decreasing, but the RHS slope is steeper). Hence, the same must hold at any point where the two sides cross. This implies that the RHS must always cut the LHS from above, which means that the two can only cross once.

[28] In the US, housing and schooling are closely connected: Children have to attend local schools, and so basically house prices also reflect the quality of local schools, given that parents want their kids to attend the best schools possible.

or equivalently

$$b = \hat{y}^*[\bar{E}_p - \underline{E}]$$

and hence is increasing in the equilibrium cutoff and in the perceived benefits of sorting, $\bar{E} - \underline{E}_r$ respectively $\bar{E}_p - \underline{E}$. As

$$\bar{E} - \underline{E}_r = \bar{E} - (1 - \gamma F)\underline{E},$$

I find that b is increasing in $\bar{E}$, decreasing in $\underline{E}$ and increasing in the degree of naivety, γ. Similar to my analysis in Section 1.4.2, I could also examine how a change in inequality will affect the equilibrium cutoff $\hat{y}^*$ and the sorting fee b (which would in this case correspond to house prices). I defer this analysis to later research. In the following sections, I will examine how the two types of misperceptions (under- and overestimating inequality) differ in terms of their implications for welfare and monopolist revenue (if the sorting technology is offered by a profit-maximizing monopolist).

1.7.11 Welfare comparison: Underestimating inequality vs. overestimating inequality

In this section, I will compare total welfare in equilibrium with different types of misperceptions. For reasons of simplicity let me denote by "Case 1" the situation where misperceptions are such that both groups underestimate inequality, and by "Case 2" the opposite situation, where both groups overestimate inequality.
If society is segregated with cutoff $\hat{y}$, total welfare can be calculated as[29]

$$W_S = \int_0^{\hat{y}} y\underline{E} f(y)dy + \int_{\hat{y}}^{y_{\max}} y\bar{E} f(y)dy - (1 - F(\hat{y}))b. \tag{1.36}$$

If people are unbiased, the sorting fee b must satisfy

$$b = \hat{y}(\bar{E} - \underline{E}).$$

If people are biased according to Case 1, where both groups underestimate the benefits of sorting, the sorting fee at the equilibrium cutoff is

$$b = \hat{y}^*(\bar{E} - \underline{E}_r) = \hat{y}^*(\bar{E}_p - \underline{E}).$$

As

$$\hat{y}^*(\bar{E} - \underline{E}_r) = \hat{y}^*(\bar{E} - \underline{E}) - \hat{y}^*\gamma F(\hat{y}^* - \underline{E}) < \hat{y}^*(\bar{E} - \underline{E}),$$

b in Case 1 is smaller than the sorting fee in the unbiased case for the same cutoff $\hat{y}^*$. Hence, welfare under sorting with misperceptions according to Case 1 delivers a higher total welfare than unbiased sorting at the same cutoff.
If people are biased according to Case 2, where both groups overestimate the benefits of sorting, the sorting fee at the equilibrium cutoff is again

$$b = \hat{y}^*(\bar{E} - \underline{E}_r) = \hat{y}^*(\bar{E}_p - \underline{E})$$

[29] As in Levy and Razin (2015), total welfare from a particular partition takes into consideration the sorting fee paid (as deadweight loss to society, or benefitting only a negligible proportion of society). If the sorting fee would not be considered, perfect sorting would always be efficient, because the utility from a match is supermodular (see Becker (1974)).

However, in Case 2 we get

$$\hat{y}^*(\bar{E} - \underline{E}_r) = \hat{y}^*(\bar{E} - \underline{E}) + \hat{y}^*\gamma F\underline{E} > \hat{y}^*(\bar{E} - \underline{E})$$

and hence the sorting fee is higher than in the unbiased case for the same cutoff $\hat{y}^*$.

Proposition 1.11 *If people are biased according to Case 1, where both groups underestimate the benefits of sorting, total welfare of sorting at the equilibrium cutoff $\hat{y}^*$ is higher than unbiased sorting at the same cutoff.*

Proposition 1.12 *If people are biased according to Case 2, where both groups overestimate the benefits of sorting, total welfare of sorting at the equilibrium cutoff $\hat{y}^*$ is lower than unbiased sorting at the same cutoff.*

Welfare and increasing inequality

Suppose that the sorting technology is offered by a benevolent social planner who wants to maximize welfare. When deciding whether or not to offer the sorting technology, she will evaluate total welfare under no sorting and compare it to total welfare with two groups for the equilibrium cutoff $\hat{y}^*$.

If there is no segregation in society, no sorting fees are paid and everybody interacts with everybody else. Hence a person with income y_i gets utility $y_i E$ and total welfare in society is

$$W_{NS} = \int_0^{y_{\max}} yE(y)f(y)dy = E^2.$$

Welfare of sorting at some cutoff $\hat{y}$ is given by (1.36). If people are unbiased, the difference between welfare of sorting at some $\hat{y}$ and welfare of no sorting can be written as

$$W_S - W_{NS} = F\underline{E}^2 + (1-F)\bar{E}^2 - \hat{y}^*(1-F)(\bar{E} - \underline{E}) - E^2. \tag{1.37}$$

Levy and Razin (2015) show that expression (1.37) can be written as

$$(1-F)(\bar{E} - \underline{E})(\bar{E} - \hat{y}^* - E)$$

which will be positive for all $\hat{y}^*$ iff

$$\bar{E} - E > \hat{y}^* \quad \forall \hat{y}^* \tag{1.38}$$

A distribution function $F(y)$ that satisfies (1.38) is called *new worse than under expectations (NWUE).* If F is NWUE, welfare of sorting at any cutoff $\hat{y}$ is higher than welfare of no sorting.

If people are biased according to Case 1, welfare of sorting at $\hat{y}^*$ can be rewritten as

$$W_S^1 = F\underline{E}^2 + (1-F)\bar{E}^2 - \hat{y}^*(1-F)(\bar{E} - \underline{E}) + \hat{y}^*(1-F)\gamma F(\hat{y}^* - E).$$

Hence, in this case the welfare difference between a situation with sorting (at equilibrium cutoff $\hat{y}^*$) and a situation with no sorting is

$$W_S^1 - W_{NS} = F\underline{E}^2 + (1-F)\bar{E}^2 - \hat{y}^*(1-F)(\bar{E} - \underline{E}) + \hat{y}^*(1-F)\gamma F(\hat{y}^* - \underline{E}) - E^2. \tag{1.39}$$

Compared to the unbiased case, the welfare difference now contains the extra term $\hat{y}^*(1 - F)\gamma F(\hat{y}^* - \underline{E})$, which is positive. Hence, $F(.)$ being NWUE is a sufficient condition for welfare being higher under sorting than under no sorting (for any cutoff) if people are biased according to Case 1.

Corollary 1.3 *If people are biased according to Case 1 and F is NWUE, a benevolent (utilitarian) social planner prefers sorting (at any cutoff) to no sorting.*

For the particular case where the poor and the rich are equally naive, expression (1.39) can be further simplified using the fact that $\hat{y}^* = E$ if $a = 1$. The welfare difference between sorting (at $\hat{y}^* = E$) and no sorting is then

$$W_S^1 - W_{NS} = F\underline{E}^2 + (1-F)\bar{E}^2 - E(1-F)(\bar{E} - \underline{E}) + E(1-F)\gamma F(E - \underline{E}) - E^2$$

$$= (1-F)(\bar{E} - \underline{E})(\bar{E} - 2E) + E(1-F)\gamma F(E - \underline{E})$$

$$= (E - \underline{E})\left(\bar{E} - 2E + \gamma E(1-F)F\right).$$

Hence

$$W_S^1 - W_{NS} > 0 \iff (E - \underline{E})\left(\bar{E} - 2E + \gamma E(1-F)F\right) > 0$$

$$\iff \bar{E} > E(2 - \gamma(1-F)F). \tag{1.40}$$

Now suppose that (1.40) is not satisfied at first, but then inequality increases in the sense of a monotone mean-preserving spread that keeps $F(E)$ constant. It is straightforward to see that this increases the RHS of (1.40), while leaving the LHS constant. Thus, offering segregation can become efficient if inequality increases.

Proposition 1.13 *If people are biased according to Case 1, an increase in inequality in the sense of a monotone mean-preserving spread that keeps $F(E)$ constant increases the welfare difference between a situation with segregation and a situation without segregation. Hence, such an increase in inequality can make it desirable for a benevolent planner to switch from a society without segregation to a society with segregation.*

Comparing the situation where inequality increases in Case 1 to the situation of increasing inequality when people are unbiased at the same cutoff (E, which is the equilibrium cutoff in Case 1 if $a = 1$), I find that sorting at E will be efficient in the unbiased case iff the income distribution is such that

$$\bar{E} > 2E.$$

In Case 1, sorting is efficient already at a lower degree of inequality (measured as $\bar{E} - \underline{E}$), namely if the income distribution is such that

$$\bar{E} > E(2 - \gamma(1-F)F).$$

The reason is that for the same cutoff welfare is always higher in Case 1 than if people are unbiased, because the sorting fee is lower, hence there will be degrees of inequality where sorting is efficient in Case 1 but not efficient if people are unbiased.
If people are biased according to Case 2, welfare can be written as

$$W_S^2 = F\underline{E}^2 + (1-F)\bar{E}^2 - (1-F)\hat{y}^*(\bar{E} - \underline{E}) - (1-F)\hat{y}^*\gamma F\underline{E}.$$

and it is immediate to see that for any cutoff, welfare in Case 2 is lower than in Case 1 and in the unbiased case. F being NWUE is a necessary and sufficient condition for sorting to be efficient (at any cutoff) in the unbiased case (and a sufficient condition in Case 1), but in Case

2 NWUE is not enough to guarantee that sorting at any cutoff yields higher welfare than no sorting, because the sorting fee is higher than in the unbiased case.
The difference between welfare of sorting at $\hat{y}^*$ in Case 2 and welfare of no sorting can be written as

$$W_S^2 - W_{NS} = (1-F)(\bar{E} - \underline{E})(\bar{E} - \hat{y}^* - E) - (1-F)\hat{y}^*\gamma F\underline{E}.$$

We can now again look what happens to this difference after a mean-preserving spread and compare Case 2 where the equilibrium cutoff is at E to Case 1 and unbiased sorting at E : If the income distribution is such that E is the equilibrium cutoff in Case 2, the welfare benefit from sorting compared to no sorting is

$$(1-F)(\bar{E} - \underline{E})(\bar{E} - 2E) - (1-F)E\gamma F\underline{E}.$$

After a monotone mean-preserving spread that leaves $F(E)$ constant, the first summand will increase, while the term that is subtracted will decrease, and therefore the welfare benefit from sorting will increase, and can go from positive to negative. However, compared to the unbiased case and Case 1, this will happen only for larger degrees of inequality (as measured by $\bar{E} - \underline{E}$), because the sorting fee is higher.

Proposition 1.14 *If $a = 1$ and the income distribution is such that the equilibrium cutoff is $\hat{y}^* = E$ in Case 2, an increase in inequality in the form of a monotone mean-preserving spread that leaves $F(E)$ constant makes sorting at E efficient (compared to no sorting) in Case 1 already for lower levels of inequality (as measured by $\bar{E} - \underline{E}$) than in the unbiased case and it makes sorting at E efficient in the unbiased case already for lower levels than in Case 2. Hence, there exist levels of inequality such that sorting at E is efficient in Case 1 but not in the other cases. There exist levels of inequality such that sorting at E is efficient in Case 1 and in the unbiased case, but not in Case 2.*

1.7.12 Monopolist profit comparison

In the following section, I will compare the profits of a monopolist who offers the sorting technology for different types of misperceptions. Again, let "Case 1" be the situation where misperceptions are such that both groups underestimate inequality, and "Case 2" the opposite situation, where both groups overestimate inequality.
Suppose a profit-maximizing monopolist who has a fixed cost $c > 0$ of offering the sorting technology can decide whether or not to become active. If people are biased according to Case 1, the monopolist's profit from offering sorting is

$$\hat{y}^*(\bar{E}(\hat{y}^*) - \underline{E}_r(\hat{y}^*))(1 - F(\hat{y}^*)) - c.$$

Given that the equilibrium cutoff is at E and substituting for $\underline{E}_r$, this can be rewritten as

$$E(E - \underline{E}(E))[1 - \gamma F(E)(1 - F(E))] - c. \tag{1.41}$$

Suppose that initially the income distribution is such that

$$E(E - \underline{E}(E))[1 - \gamma F(E)(1 - F(E))] - c < 0$$

and hence the monopolist would prefer to stay out of the market. If inequality increases (again in the sense of a monotone mean-preserving spread of the income distribution that leaves $F(E)$ constant), $E - \underline{E}$ increases. This means that if the increase in inequality is sufficiently large, the profits from offering the sorting technology will become positive and the society will become segregated. Thus, a large enough increase in inequality will lead to economic segregation.

If people are biased according to Case 2, the monopolist's profit from offering the sorting technology is again

$$\hat{y}^*(\bar{E} - \underline{E}_r)(1 - F(\hat{y}^*)) - c$$

but now people overestimate the benefits of sorting and hence this expression can be rewritten as

$$\hat{y}^*(\bar{E}(\hat{y}^*) - (1 - \gamma F(\hat{y}^*))\underline{E}(\hat{y}^*))(1 - F(\hat{y}^*)) - c$$

It depends on the shape of the income distribution whether the monopolist's profit increases or decreases due to a mean-preserving spread. Remember that the equilibrium cutoff in Case 2 if $a = 1$ is given by

$$1 - F(\hat{y}^*) = \frac{E}{y_{\max}}.$$

Hence, the equilibrium cutoff need not be at average income E in this case, the exact location of $\hat{y}^*$ depends on the income distribution. This means that in general, a mean-preserving spread will not change only $\bar{E}$ and $\underline{E}$ but also the equilibrium cutoff. Therefore, the overall effect of a monotone mean-preserving spread on the monopolist's profits is not clear: $\bar{E} - \underline{E}_r$ increases, but the equilibrium cutoff may go up or down and what happens to the overall sorting fee b and to the monopolist's profits depends on the shape of the income distribution. Thus, we cannot in general compare whether the monopolist will be quicker to enter than in Case 1 if inequality increases.

However, if the income distribution is such that the equilibrium cutoff is also at E in Case 2, then a monotone mean-preserving spread (that leaves $F(E)$ constant) will not affect the location of the equilibrium cutoff and the monopolist's profits will increase due to a monotone mean-preserving spread. Moreover, the monopolist will offer the sorting technology for lower degrees of inequality (as measured by $\bar{E} - \underline{E}$) than the monopolist in Case 1 (with the same fixed costs c), because her revenue $(1 - F(\hat{y}))b$ is higher for any cutoff (and therefore also for $\hat{y} = E$) than in Case 1, because the sorting fee b is higher.

If people are unbiased, the sorting fee for a given cutoff $\hat{y}$ amounts to

$$b = \hat{y}(\bar{E}(\hat{y}) - \underline{E}(\hat{y}))$$

and hence the monopolist's profits from offering the sorting technology at cutoff $\hat{y}$ are

$$\Pi(\hat{y}) = (1 - F(\hat{y}))\hat{y}(\bar{E}(\hat{y}) - \underline{E}(\hat{y})) - c.$$

As the sorting fee lies in between the one for Case 2 and the one for Case 1, the monopolists profits will be lower than in Case 2 and higher than in Case 1 for any cutoff $\hat{y}$.

In order to compare the effects of an increase in inequality in this case to the effect in Case 1 and Case 2, I will again assume that the cutoff is at average income. (Note that if people are unbiased, the monopolist can set the cutoff anywhere in Y and will therefore set it such that her profits are maximized. This means that at the optimal cutoff we need[30]

$$\Pi'(\hat{y}) = 0$$

[30] It is straightforward to see that this maximization problem has an interior solution, because $\Pi(0) = \Pi(y_{\max}) = -c$ whereas any interior $\hat{y}$ yields $\Pi(\hat{y}) > -c$.

which can be rewritten as

$$\bar{E}(\hat{y}) - E = \hat{y}\frac{f(\hat{y})}{1 - F(\hat{y})}(\hat{y} - \underline{E}(\hat{y})). \tag{1.42}$$

Plugging $\hat{y} = E$ into 1.42 and rearranging, I find that for average income to be the optimal cutoff, the income distribution must be such that $\frac{F(E)}{f(E)} = E$.) Suppose that income is initially distributed relatively equally, such that the difference between average income of the rich and average income of the poor is small, i.e. $\bar{E} - \underline{E}$ is low, and the monopolist's profits are negative. If income inequality increases in the form of a monotone mean-preserving spread that leaves $F(E)$ and $f(E)$ constant then the equilibrium cutoff will not change (because it is at average income), but $\bar{E} - \underline{E}$ will increase, and hence also the monopolist's profits. If the mean-preserving spread is large enough, the monopolist will find it profitable to offer the sorting technology. Offering sorting will become profitable for smaller degrees of inequality than in Case 1 and for larger degrees of inequality than in Case 2.
This analysis yields the following Propositions:

Proposition 1.15 *Let people be biased according to Case 1 and suppose that the income distribution is initially such that a profit maximizing monopolist with fixed costs $c > 0$ does not find it profitable to offer the sorting technology. Then for any $c < \infty$ there exists a mean-preserving spread of the income distribution such that the monopolist's profits become positive.*

Proposition 1.16 *Let people be biased according to Case 2 and let the income distribution be such that the equilibrium cutoff is at average income. Suppose that the income distribution is initially such that a profit maximizing monopolist with fixed costs $c > 0$ does not find it profitable to offer the sorting technology. Then for any $c < \infty$ there exists a mean-preserving spread of the income distribution such that the monopolist's profits become positive.*

Proposition 1.17 *Let people be unbiased and let the income distribution be such that the monopolist's optimal profit is at average income. Suppose that the income distribution is initially such that a profit maximizing monopolist with fixed costs $c > 0$ does not find it profitable to offer the sorting technology. Then for any $c < \infty$ there exists a mean-preserving spread of the income distribution such that the monopolist's profits become positive.*

Proposition 1.18 *The monopolist's profits will be higher in Case 2 than in Case 1 for any cutoff $\hat{y}^*$ and for any degree of inequality of the income distribution. Therefore, as income inequality increases (in the sense of a monotone mean-preserving spread), the monopolist's profits in Case 2 will become positive already for smaller degrees of inequality (in terms of $\bar{E} - \underline{E}$) than necessary for her profits in Case 1 to be positive.*

Proposition 1.19 *For any income distribution, the profits for the unbiased case lie in between the profits for Case 2 and Case 1 (for the same cutoff). Therefore, as income inequality increases (in the sense of a monotone mean-preserving spread), the monopolist's profits in the unbiased case will become positive already for smaller degrees of inequality (in terms of $\bar{E} - \underline{E}$) than necessary for her profits in Case 1 to be positive. However, higher degrees of inequality are needed for her profits to be positive than in Case 2.*

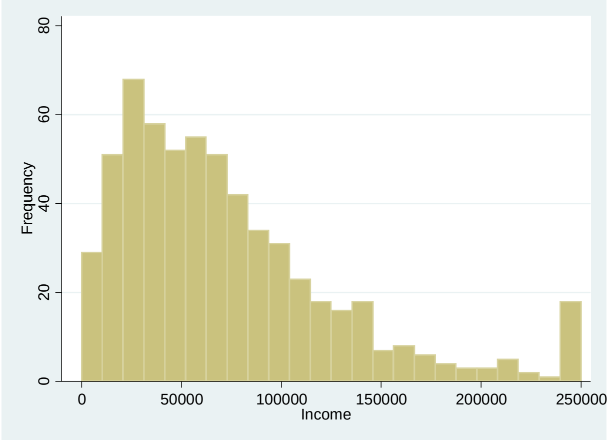

Figure 1.5: Sample household income distribution

1.8 Appendix B: Empirical Appendix

1.8.1 Working with Amazon Mechanical Turk

For tax reasons, it is not possible for researchers living outside the United States to use Amazon Mechanical Turk directly. Therefore, I used the Amazon requester MTurkData to publish my survey via Amazon Mechanical Turk. They check the survey for compliance with Amazon's Terms and Conditions, publish it on MTurk and deal with the payment of the workers afterwards.
The advantages and disadvantages of working with Amazon Mechanical Turk have been discussed by Kuziemko et al (2015) in their online appendix. I agree with them that a major advantage of using MTurk is the speed of gathering responses: In my case, it took less than two hours to get 600 responses. There might in general be doubts about the quality of the responses, but it is possible to screen the MTurk workers based on their ratings for previous tasks. Using MTurk is also relatively cheap, as researchers design the survey themselves, instead of having it designed by a professional survey company. (Note also that I did not keep costs low at the expense of the respondents: they were all paid an hourly wage of 9 dollars.) One disadvantage of using MTurk is definitely that the obtained sample is usually not as representative as other, more expensive, online panel surveys (see below for a description of my own sample). However, as long as one keeps this in mind when interpreting the results, I think this is tolerable, especially when working with respondents from the United States, where MTurk is relatively well known and the pool of workers is therefore fairly representative.

1.8.2 Sample characteristics

The sample is 83% White, 8.3% Black, 5.3% Asian and 1.5% Native American (the rest is "of other ethnicity"). Average age is 36.78, 44% of respondents are married. 68% are full- or part-time employees, 17% are self-employed and 13% are unemployed or not in the labour force. The respondents are very well educated, 63% have completed some kind of college degree. Hence, compared to other (more representative and commonly used) online panel surveys cited in the online appendix of Kuziemko et al. my sample is younger, more educated and has fewer minorities. The household income distribution of the sample is roughly similar to the actual US household distribution (see Figures 1.5 and 1.6).

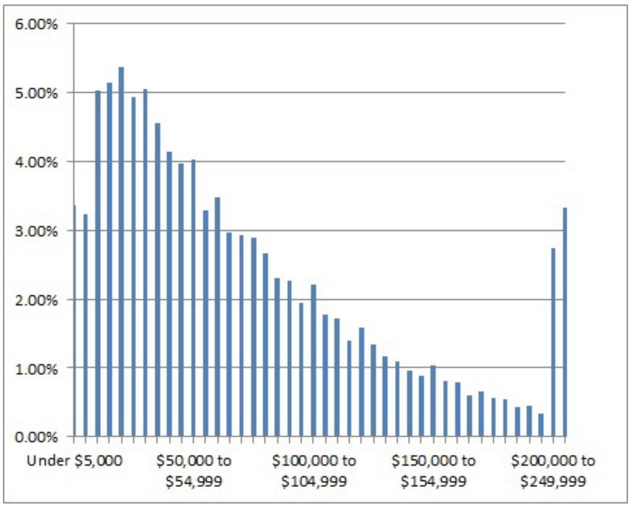

Figure 1.6: US household income distribution 2015 (Source: US Census Current Population Survey)

Social Segregation: Description of Factor Analysis

In the survey, I ask several questions about people's colleagues at work, friends and family (spouse and siblings, if applicable). This is an attempt at identifying how diverse a person's social circle is. The underlying hypothesis is that an individual is more "socially segregated" the more homogenous and similar to herself her social circle is. However, it turned out that some of the questions were practically useless for my analysis in this relatively small sample: As less than half of the respondents are married, it turned out that using spouse characteristics to categorize social segregation would exclude a big part of the sample, and a similar reason can be applied to sibling characteristics. I therefore decided to exclude those variables from my factor analysis. Furthermore, I excluded variables indicating whether friends or colleagues have the same mother tongue, because I figured out that these variables predominantly serve to identify Hispanics in the sample and do not provide much variation. Hence, the factor analysis utilizes four categorical variables classifying the similarity of friends' and colleagues' education and income level. The variables take on the value 0 if the respondent has answered that all of their friends/ colleagues are different to them in the respective area (e.g. the variable friends_educ is 0 if the respondent states that all of her friends have a different education level than herself) and is then increasing in the degree of similarity (i.e. 1 if most friends have different education levels,... up to 4 if all friends have the same education level as the respondent). Hence, the higher the value of each categorical variable, the higher the respondent's degree of social segregation.
The results of the factor analysis are presented in the main text.

Chapter 2

Sorting in the Presence of Misperceptions

2.1 Introduction

Who we choose to socialize with is often determined by our beliefs about others, about their qualities and their characteristics, and by our surmise about what effects their good and bad traits will have on ourselves. It is natural that we try to interact frequently with people who we think we can benefit from, be it in a material sense or simply because we enjoy their company.

On the other hand, our beliefs about society are likely to be influenced by our social interactions, by what and who we observe on a day-to-day basis. Depending on how diverse our social circles are, we might end up knowing a lot or very little about certain groups in society. Specifically, if we do not interact with some social groups, we are prone to develop distorted beliefs about what people are like in that group, about their characteristics and their traits. This in turn might influence who we choose to interact with in the first place and hence solidify and reinforce our attitudes and beliefs.

Take for example the question of how contact with ethnic minorities affects people's attitudes towards minorities. Dustmann and Preston (2001) show that looking at the effect of living in an area with high ethnic diversity on attitudes towards minorities can give a misleading answer. The reason for this is that we can at least to some degree decide where we want to live, and therefore people might live in ethnically diverse areas because they have a favourable attitude towards minorities in the first place - there is a two-way interaction between location choice and people's beliefs that needs to be taken into account.

Another example is parents' school choice for their kids. There is considerable evidence (see e.g. a 2007 Center on Education Policy report using National Educational Longitudinal Study (NELS) data from 1988-2000) that private schools and state schools yield relatively similar learning outcomes if we control for pupils' family background.[1] Nevertheless, parents are willing to pay a lot to live in areas with supposedly "good" schools (especially in the US) or to send their kids to private schools. However, these seem to be mainly parents who were privately educated themselves. In fact, Evans and Tilley (2011) show that in the UK parents who went to private schools are five times more likely to send their kids to private schools than state-educated parents (controlling for income). Levy and Razin (2016) describe a model in which beliefs about the benefits of private education are passed on from parents to

[1] On a related note, Abdulkadiroglu et al. (2014) show that high achieving peers and racial composition of schools have no effect on learning achievement of individual pupils.

their children and influence school choice (and subsequent success in the labour market) from generation to generation.

If we want to examine how changes in the economy - like an increase in income inequality, a reform to the education system or a surge in immigration - affect social groups and the belief system in society, we have to bear in mind that it doesn't suffice to look at the direct effect that these changes have on segregation and beliefs. Where society will end up in the long run depends also on the mutual reinforcement and interdependence of social segregation and beliefs.

The interaction of social segregation and beliefs about society is what I examine in this paper. I take the canonical model of sorting according to income as a starting point. In this model, all feasible partitions of the income distribution are monotone, i.e. social groups will be single intervals of the income distribution. It is important to note, though, that without further assumptions on the sorting fees there is nothing that pins down the exact way in which society will segregate in this model, i.e. how the social groups will look like. By varying the menu of sorting fees accordingly, any monotone partition of the income distribution is feasible. Furthermore, while this model takes into account that our beliefs about society affect our social interactions, it doesn't allow for the effect to go in the other direction: people's beliefs about the whole of society remain unchanged (and unbiased), even if people interact mainly with their own social circle.

In the present paper, I eliminate these shortcomings by adding misperceptions to the model. I demonstrate that this addresses both issues of the canonical model at the same time. First, it accounts for the fact that our social interactions shape our beliefs about society, and second it limits the amount of partitions that are feasible in equilibrium and therefore reduces multiplicity.

In my model of sorting with misperceptions I assume that, once society is segregated and people interact mainly with their own social circle, they become biased about the overall income distribution, and specifically about average income in the other groups. I define as "biased sorting equilibria" those partitions of the income distribution that are stable given people's misperceptions, i.e. partitions in which people want to stay in their chosen group, despite their acquired distorted beliefs.

I show that adding misperceptions to the model initially leads to more complications: While in the canonical model all equilibrium partitions are monotone (i.e. single, connected intervals), biased sorting equilibria can also be non-monotone and hence people in one and the same group can have very different incomes, which complicates the analysis and is at odds with empirical evidence of assortative matching. Furthermore, the issue of multiplicity of equilibrium partitions persists, and even for a given equilibrium partition the sorting fees might not be uniquely determined. Finally, people's beliefs about other groups can be inconsistent with what they see: they can be surprised by seeing people with certain incomes choosing to be in certain groups, because given their beliefs about incomes in the other groups they do not think these choices are optimal.

In order to address these problems, I introduce a refinement criterion that I call the "consistency requirement". A partition satisfies consistency if the misperceptions are such that people are not surprised by the choices of people in other groups. I show that all biased sorting equilibria with consistency are monotone, and that the menu of sorting fees is uniquely pinned down for a given equilibrium partition. Furthermore, I demonstrate that if there are two groups in society and the misperceptions satisfy a form of monotonicity, then the consistency requirement selects a unique biased sorting equilibrium out of all possible stable partitions. In that case, social groups and sorting fees are uniquely determined. In the last section I examine under which conditions on the belief function biased sorting equilibria with consistency and

more than two groups will exist.
The rest of this paper is organized as follows. Section 2 discusses related literature, Section 3 presents the canonical model of sorting with respect to income and Section 4 introduces misperceptions into that model and explains the concept of biased sorting equilibrium and the consistency requirement in its local and global form. Section 5 finds conditions on the functional form of the misperceptions that lead to existence and uniqueness of binary biased sorting equilibria with consistency. Section 6 examines under which conditions biased sorting equilibria with consistency and more than two groups can exist. Section 7 concludes.

2.2 Relation to existing literature

The canonical model of sorting and assortative matching was most famously employed by Becker (1974) to model the marriage market. Pesendorfer (1995) uses it to explain fashion cycles. Rayo (2013) examines optimal sorting from a profit-maximizing monopolist's point of view, Damiano and Li (2007) analyze the case of two or more competing firms and optimality with respect to welfare is explored in Levy and Razin (2015). In Windsteiger (2017c), I compare the optimal partitions for the monopolist and the social planner and find that the optimal type of sorting depends on the shape of the income distribution and varies depending on which entity (profit-maximizing monopolist or benevolent social planner) is assumed to offer the sorting technology.
What all the above papers have in common is that beliefs about society determine who people interact with, but social interactions do not influence beliefs. In fact, people retain perfect knowledge about society despite interacting only with a (potentially small) group of society.
Recently, the fact that segregation can affect beliefs has gained attention in the literature: On the theoretical side, Golub and Jackson (2012) present a model in which homophily (and resulting segregation) slows down convergence to a consensus in society. Concerning empirical evidence, Algan et al. (2015) show that political views converge among peers at university, and Boisjoly et al. (2006) and Burns et al. (2013) find that having roommates of a different ethnicity to one's own lowers students' prejudices.
That (potentially biased) beliefs can, in turn, have an effect on segregation, is pointed out by Dustmann and Preston (2001). They argue that estimating the effect of living in ethnically diverse neighbourhoods on attitudes towards minorities can lead to biased results, if we do not take into account how those attitudes affect neighbourhood choices in the first place. Levy and Razin (2016) present a model in which beliefs about school quality and parent's school choice for their children interact to create essentially two groups of society: a group of privately educated parents who believe in the benefits of the private school system and send their children to private school as well, and a group of state educated parents, who think private schools are not worth paying for and send their kids to state schools.
The main contribution of my paper is to present a general model in which beliefs about society and segregation choice interact to create an endogenous system of beliefs and societal groups. This general model can be used to analyze sorting according to many variables that are distributed continuously in society.[2] While in my version of the model, I assume that people sort according to income, the continuous variable could also be "ability" or "intelligence" and the model could be about sorting into different types of schools.

[2]Due to strict increasingness of U those should, however, be variables where the whole of society agrees that "more is better", such as intelligence, ability, income or wealth.

2.3 A theoretical model of economic segregation

Let income y in an economy be distributed according to some income distribution $F(y)$ on the interval $Y = [0, y_{\max}]$ where $y_{\max} < \infty$. Assume furthermore that $F(y)$ is continuous and strictly monotonic.

Suppose that an agent's utility is increasing not only in her own income but also in the average income of the people that she interacts with, which I will henceforth call her "reference group". Specifically, I will assume that a person with income y_j gets utility $U_j = U(y_j, E(y|y \in S_i))$, where S_i is individual j's reference group. If there is no economic segregation, let everybody's reference group be a representative sample of the whole population, such that $U_j = U(y_j, E(y))$. However, suppose that a person with income y_j can pay b_i ($i \in \{1, .., n\}$) to join club S_{b_i} ($i \in \{1, .., n\}$) and get utility

$$U(y_j, E[y|y \in S_{b_i}]) - b_i$$

or refrain from paying any b_i and get

$$U(y_j, E[y|y \in S_{b_0}]),$$

where S_{b_i} is the set of incomes y of people who have paid b_i and $b_0 = 0$. Let $U(.,.)$ be continuous, strictly increasing in both arguments and strictly supermodular, such that[3]

$$\forall x' > x : U(y, x') - U(y, x) \text{ is strictly increasing in } y.$$

Then I can define the following:

Definition 2.1 *A sorting equilibrium is a partition* $[S_{b_0}, S_{b_1}, .., S_{b_n}]$ *of* Y *and a menu of sorting fees* $[b_0, b_1, .., b_n]$ *(with* $b_i < b_{i+1}$ $\forall i$ *and* $b_0 = 0$*) such that* $\forall i \in \{0, .., n\}$

$$U(y, E[y|y \in S_{b_i}]) - b_i \geq U(y, E[y|y \in S_{b_k}]) - b_k \quad \forall y \in S_{b_i}, \forall i, \forall k \neq i$$

$$\Leftrightarrow U(y, E[y|y \in S_{b_i}]) - U(y, E[y|y \in S_{b_k}]) \geq b_i - b_k \quad \forall y \in S_{b_i}, \forall i, \forall k \neq i \qquad (2.1)$$

In a sorting equilibrium as defined above people stay in the group that gives them the highest utility.

Corollary 2.1 *In any sorting equilibrium, groups with a higher average income correspond to higher sorting fees.*

Proof. This immediately follows from Definition 2.1, from the assumption that all b_i are different and from the fact that U is strictly increasing in both arguments. ■

I can show that all sorting equilibria will be of a certain form:

Proposition 2.1 *All sorting equilibria will be monotone.*[4] [5]

[3] Note that this paper offers only a very reduced form model of economic segregation. That people's utility is increasing in the average income of the other people they mix with is perhaps a simplified way of saying that living in an affluent neighborhood offers many benefits, such as good schools (because people are willing to spend more on the education of their kids, and because the presence of children of rich people might increase other pupil's chances in life through various peer effects) and pleasant surroundings such as parks or leisure centres (perhaps with increased security or surveillance). Instead of modelling all this on a micro level, I subsum all these effects into a utility function which is increasing in the average income of one's peers.

[4] By *monotone* I mean that the groups S_{b_i} are single intervals of Y. (By Corollary 2.1, this implies that those groups sitting higher up on the Y scale correspond to higher sorting fees.)

[5] If some or all b_i are equal, then there exist trivial non-monotone sorting equilibria where the average income in all those groups with the same b_i is the same, so that people are indifferent about which of these groups to join. I exclude those cases from my analysis by assuming that $b_i < b_{i+1}$.

Proof. Suppose w.l.o.g. that a sorting equilibrium exists where $y_2 \in S_{b_i}$ and $y_1 \in S_{b_j}$, with $b_i < b_j$ but $y_2 > y_1$. Then I must have

$$U(y_2, E[y|y \in S_{b_j}]) - U(y_2, E[y|y \in S_{b_i}]) \leq b_j - b_i$$

and

$$U(y_1, E[y|y \in S_{b_j}]) - U(y_1, E[y|y \in S_{b_i}]) \geq b_j - b_i$$

and hence

$$U(y_1, E[y|y \in S_{b_j}]) - U(y_1, E[y|y \in S_{b_i}]) \geq U(y_2, E[y|y \in S_{b_j}]) - U(y_2, E[y|y \in S_{b_i}]).$$

But due to $y_2 > y_1$ this is a contradiction to U being strictly supermodular. ■
Proposition 2.1 allows me to rewrite the definition of a sorting equilibrium in terms of intervals of Y.

Corollary 2.2 *A sorting equilibrium is characterized by a partition* $[0, \hat{y}_1, ..., \hat{y}_n, y_{\max}]$ *of* Y *and a menu of sorting fees* $[b_0, b_1, .., b_n]$ *(with* $b_i < b_{i+1}$ $\forall i$ *and* $b_0 = 0$*) such that* $\forall i \in \{1, .., n\}$

$$U(\hat{y}_i, E[y|y \in S_{b_i}]) - b_i = U(\hat{y}_i, E[y|y \in S_{b_{i-1}}]) - b_{i-1} \quad \forall i$$

$$\Leftrightarrow U(\hat{y}_i, E[y|y \in S_{b_i}]) - U(\hat{y}_i, E[y|y \in S_{b_{i-1}}]) = b_i - b_{i-1} \quad \forall i \tag{2.2}$$

Proof. Given the fact that $S_{b_i} = [\hat{y}_i, \hat{y}_{i+1}]$ $\forall i$ and equilibrium condition (2.1), it follows that both

$$U(y, E[y|y \in S_{b_i}]) - U(y, E[y|y \in S_{b_{i-1}}]) \geq b_i - b_{i-1} \quad \forall y \in [\hat{y}_i, \hat{y}_{i+1}], \forall i$$

and

$$U(y, E[y|y \in S_{b_i}]) - U(y, E[y|y \in S_{b_{i-1}}]) \leq b_i - b_{i-1} \quad \forall y \in [\hat{y}_{i-1}, \hat{y}_i], \forall i$$

need to hold in any sorting equilibrium. This implies that a person with income $\hat{y}_i$ just at the border of two groups $S_{b_{i-1}}$ and S_{b_i} has to be exactly indifferent between joining either of the two groups in equilibrium. Hence, we get

$$U(\hat{y}_i, E[y|y \in S_{b_i}]) - U(\hat{y}_i, E[y|y \in S_{b_{i-1}}]) = b_i - b_{i-1} \quad \forall i$$

■

Corollary 2.3 *For a given equilibrium partition* $[0, \hat{y}_1, ..., \hat{y}_n, y_{\max}]$*, the menu of sorting fees* $[b_0, b_1, .., b_n]$ *(with* $b_i < b_{i+1}$ $\forall i$ *and* $b_0 = 0$*) is always unique.*

Proof. This follows immediately from equilibrium condition (2.2). ■
The above presented canonical model of sorting according to income has many positive features: The equilibrium partitions will always be monotone and therefore individual income within the resulting equilibrium groups will be similar (or at least within a simple interval), which simplifies the analysis of the model and is also compatible with empirical evidence of segregation according to income and assortative matching. Furthermore, inherent in the above definition is a notion of *consistency*: if a person sees the income of another person and knows which group this person joined, she always thinks that this person is correct in doing so, because both people evaluate the benefit of being in a certain group (given a certain income) equally. This means that no extra condition is needed to guarantee consistency, it "comes for free" in the equilibrium condition (2.2) - if a person in a certain group thinks that being in that group is best for her, then also all other people - no matter which group they belong to - will think that this is optimal for her.

However, from equilibrium condition (2.2) it is immediate to see that the model delivers no prediction about the type of segregation that will happen in a society, i.e. how the social groups will look like: For any continuous distribution function F there exists an infinite number of equilibrium partitions. More specifically, for any partition $P = [0, \hat{y}_1, ..., \hat{y}_n, y_{\max}]$ there exists a menu of sorting fees $\mathbf{b} =[0, b_1, ...b_n]$ such that $(P, \mathbf{b})$ is a sorting equilibrium. Moreover, while in this model people's beliefs about society determine their social interactions, the reverse effect is not taken into account: segregation has no effect on people's beliefs about the economy and people retain perfect knowledge about the income distribution in the whole of society, even though they interact mainly with a select group of people who are similar to them in terms of income.

One way to try and resolve the issue of multiplicity is to look at the supply side of the sorting technology: we can analyze the optimal partition that a profit-maximizing monopolist, a number of competing firms or a benevolent social planner would want to offer and thereby select "plausible" equilibria out of the infinite number of possible equilibria. I explore this path in another paper of mine and find that the form of resulting optimal partitions depends on the underlying distribution function and on which entity is assumed to provide the sorting technology (see Windsteiger (2017c)).

In this paper, I pursue a different path: I add misperceptions to the model. Specifically, I will assume that people, once they are sorted into their respective groups, become biased about average income in the other groups, and I will define partitions as biased sorting equilibria if they are such that people want to stay in their group given their misperceptions about the rest of society. I will show in the next section that this addresses both of the above mentioned limitations of the canonical model: First, it lifts the assumption that people retain perfect knowledge about society once they are sorted and therefore allows for the interaction between segregation and beliefs to go both ways. Second, restricting attention to biased sorting equilibria with the additional requirement of consistency (which I will explain below) greatly reduces the number of possible equilibrium partitions and can lead, if the misperceptions are of a certain form, even to uniqueness.

2.4 Sorting with misperceptions

Suppose that people, once they are sorted into their group S_{b_i}, become biased about average income in the other groups and hence about the overall income distribution. I will model a group's belief about the other groups as resulting from a group belief "technology". Specifically, I will assume that for any partition of Y with $n+1$ groups, people's biased perception of the average income of the other groups can be characterized by the belief function

$$B : \mathbf{P} \to Y^{(n+1)^2}$$

(where $\mathbf{P}$ is the space of all partitions $P = [S_0, S_{b_1}, ..S_{b_n}]$ of Y) that maps every partition of Y into an $(n+1)^2$-dimensional vector of beliefs

$$(E_0^0(P), E_0^1(P), ...E_0^n(P), E_1^0(P), E_1^1(P), ..., E_n^{n-1}(P), E_n^n(P)), \tag{2.3}$$

where $E_i^j(P)$ denotes group i's belief about average income in group j and hence the first $n+1$ entries of (2.3) denote group S_{b_0}'s belief about average income in S_{b_0} and all the other groups, entry $n+2$ to $2(n+1)$ denote group S_{b_1}'s belief about average income in all groups, etc.... I will assume that the belief function is such that people are correct about average income in their own group, i.e.

$$E_i^i(P) = E^i(P) \ \ \forall i,$$

where $E^i(P)$ is the true average income in S_{b_i}, i.e. $E^i(P) = E[y|y \in S_{b_i}]$.[6] Furthermore, the above definition of the belief function implies that misperceptions are constant within groups, i.e. people who are in the same group have the *same* (mis-)perception about the other groups' average (and thus misperceptions do not depend on one's own income directly, but on group membership).[7]

I also restrict the beliefs of one group about average income in another group to actually lie in that group's income range:

Assumption 2.1 *The belief function $B(P)$ is such that*

$$\inf S_{b_j} \leq E_i^j(P) \leq \sup S_{b_j} \forall i \forall j.$$

Given this belief function, I can define the following:

Definition 2.2 *A* ***biased sorting equilibrium*** *is a partition $P = [S_{b_0}, S_{b_1}, .., S_{b_n}]$ of Y and a menu of sorting fees $[b_0, b_1, .., b_n]$ (with $b_i < b_{i+1}$ $\forall i$ and $b_0 = 0$) such that $\forall i \in \{0, .., n\}$*

$$U(y, E^i(P)) - b_i \geq U(y, E_i^k(P)) - b_k \quad \forall y \in S_{b_i}, \forall i, \forall k \neq i. \tag{IC}$$

A biased sorting equilibrium is therefore a partition of Y that is "stable" given people's misperceptions about the other group. When people compare the utility they obtain in their own group and compare it to the utility they *think* they could obtain in any other group - given their misperceptions about average income in the other groups - they come to the conclusion that they reach the highest possible level of utility in their own group and therefore they do not want to move to another group.

Corollary 2.4 *In any biased sorting equilibrium, groups whose members perceive that the average income in their group is high compared to other groups will have high sorting fees compared to other groups.*

Proof. This follows immediately from Definition 2.2 and the fact that U is strictly increasing. ■

With biased perceptions, non-monotone sorting equilibria can exist, as the following example demonstrates.

Example 1 *Suppose $Y = [0, 1]$ and income y is uniformly distributed, and suppose that $U(y, x) = yx$. Suppose the two groups $S_{b_0} = \left[0, \frac{1}{4}\right) \cup \left(\frac{1}{2}, \frac{3}{4}\right]$ and $S_{b_1} = \left[\frac{1}{4}, \frac{1}{2}\right] \cup \left(\frac{3}{4}, 1\right]$ would constitute an equilibrium partition if people are unbiased. The correct average income of S_{b_1} is $\frac{5}{8}$ and the average income of S_{b_0} is $\frac{3}{8}$. In an unbiased sorting equilibrium the sorting fee b_1 must be strictly positive (normalizing $b_0 = 0$), because the average in group S_{b_1} is higher than in group S_{b_0}. Now take the person with income $\frac{5}{8}$, who is in S_{b_0}. She derives utility $\frac{5}{8} \times \frac{3}{8}$ from being in S_{b_0} and utility $\frac{5}{8} \times \frac{5}{8} - b_1$ from being in S_{b_1}. Because this person is in S_{b_0} we must have*

$$\frac{5}{8} \times \frac{5}{8} - b_1 \leq \frac{5}{8} \times \frac{3}{8}$$

[6]For reasons of simplicity, I will restrict attention to partitions P such that all groups S_{b_i} have strictly positive measure.

[7]I restrict my attention to misperceptions that are constant within group because I specifically want to focus on differences in perceptions *between* groups rather than *within* groups. This restriction helps to simplify the analysis, but the main results of this paper would not change fundamentally if biases were to vary also within groups. The restriction can be deduced "naturally" from the assumption that people interact and communicate freely within their own group and hence will, within their group, reach a common belief about the other groups.

or equivalently

$$\frac{5}{8}\left(\frac{5}{8}-\frac{3}{8}\right) \leq b_1. \tag{2.4}$$

At the same time, the person who is at $\frac{3}{8}$ is in S_{b_1}, hence her utility from being in S_{b_1} must be higher than the utility from being in S_{b_0}. This yields the condition

$$\frac{3}{8}\left(\frac{5}{8}-\frac{3}{8}\right) \geq b_1 \tag{2.5}$$

It is immediate to see that (2.4) and (2.5) contradict each other, so this partition cannot be an unbiased sorting equilibrium. However, depending on the belief function, it is possible that this partition is a biased sorting equilibrium: Suppose that the belief function is such that people in S_{b_1} perceive the average in their group and in the other group correctly, but the people in S_{b_0} all think that the average in group S_{b_1} is equal to the average in their own group. Then they would not be willing to pay any $b_1 > 0$ to join group S_{b_1}, whereas everybody in group S_{b_1} is willing to pay some positive b_1, e.g. $b_1 = \frac{1}{16}$ (which makes the poorest person in S_{b_1} exactly indifferent between the two groups, while everybody else in S_{b_1} strictly prefers being there). In fact any $b_1 \in (0, \frac{1}{16}]$ would make the above partition a biased sorting equilibrium.

2.4.1 Global and local consistency

At first, adding misperceptions to the model does not simplify the analysis, but rather adds some additional problems: As the above example shows, with sufficient freedom on how to specify the groups' misperceptions, non-monotone biased sorting equilibria are possible. I consider this to be an undesirable feature because it complicates the analysis and is at odds with empirical evidence of how groups in society are formed. Additionally, the menu of sorting fees might not be uniquely determined for a given equilibrium partition (see Example 1 above). Moreover, we do not necessarily have the notion of consistency (which is inherent in the unbiased model, as explained above) in a biased sorting equilibrium. Go back to Example 1: People's beliefs about the other group are inconsistent with what they see: Everybody in group S_{b_0} wonders why anybody would want to pay b_1 to join S_{b_1}, while at the same time the people in S_{b_1} cannot understand why the people with income between $\frac{1}{2}$ and $\frac{3}{4}$ do not want to join their group.

However, the inconsistency, the non-monotonicity and the non-uniqueness of the menu of sorting fees for a given equilibrium partition vanish if I introduce what I call the *consistency requirement*. This requirement can hold either *locally* or (in its stronger version) *globally*. Let me first introduce the notion of *global consistency:* This requires that people who are in S_{b_i} think that people with different incomes, who are not in S_{b_i} but in some other group S_{b_j}, are correct in doing so. Formally, this requirement translates to

Definition 2.3 *A partition $P = [S_{b_0}, S_{b_1}, .., S_{b_n}]$ of Y and a menu of sorting fees $[b_0, b_1, .., b_n]$ satisfy **global consistency** iff $\forall i \in \{0, .., n\}$*

$$U(y, E_i^j(P)) - b_j \geq U(y, E_i^k(P)) - b_k \quad \forall y \in S_{b_j}, \forall j, \forall k. \tag{GC}$$

In words, (GC) says that a person in group S_{b_i} who looks at a person with income y in any other group S_{b_j}, thinks that this person cannot achieve higher utility by switching to a different group (and note that the person from group S_{b_i} evaluates person y's utility given her own biased perception of average group incomes E_i, the one that she has acquired in her group S_{b_i}).

As I have pointed out, in the "unbiased" sorting equilibrium that I have defined in the previous section, global consistency is implicit. Because everybody has the same (correct) understanding of average incomes in all the groups, people cannot be "puzzled" by other people's choices - everybody evaluates everybody else's utility in the same way. It is only when people have incorrect perceptions of the other groups that consistency becomes a separate issue and is not implicit in the equilibrium definition. People can be happy with their own choices (which means we are in a sorting equilibrium), while at the same time not understanding other people's choices (which means that consistency is violated). Hence, it makes sense - as a refinement to biased sorting equilibria - to define biased sorting equilibria which additionally satisfy global consistency:

Definition 2.4 *A* ***biased sorting equilibrium with global consistency*** *is a partition* $P = [S_{b_0}, S_{b_1}, .., S_{b_n}]$ *of* Y *and a menu of sorting fees* $[b_0, b_1, ...b_n]$ *(with* $b_i < b_{i+1}$ $\forall i$ *and* $b_0 = 0$*) such that* $\forall i \in \{0, .., n\}$

$$U(y, E^i(P)) - b_i \geq U(y, E_i^k(P)) - b_k \quad \forall y \in S_{b_i}, \forall k \tag{IC}$$

$$U(y, E_i^j(P)) - b_j \geq U(y, E_i^k(P)) - b_k \quad \forall y \in S_{b_j}, \forall j, \forall k \tag{GC}$$

A less restrictive requirement than global consistency is the notion of *local consistency.* To explain this concept, I first need to define what neighbouring groups are in the present context:

Definition 2.5 *If* $[S_{b_0}, S_{b_1}, ...S_{b_n}]$ *is a partition of* Y *with corresponding sorting fees* $[b_0, b_1, ...b_n]$ *(with* $b_i < b_{i+1}$ $\forall i$*) then for all* i *the neighbouring groups of* S_{b_i} *are* $S_{b_{i-1}}$ *and* $S_{b_{i+1}}$.

I define neighbouring groups in terms of sorting fees. However, as groups with higher sorting fees correspond to higher average income, this is equivalent to defining them in terms of average income: neighbouring groups are groups which are "next to each other" if ranked according to mean income. Therefore, local consistency can be interpreted in the following way: People understand the decisions of people who have income that is relatively similar to their own income, i.e. people who are in a group that is a bit richer or poorer than their own group. However, they don't think about people who are in much richer or poorer groups (and therefore do not need to think that their decisions are optimal). The reason for why people care only about neighbouring groups could be that these groups are more salient to them, because in effect those are the groups that matter also for their own individual optimal decision-making about which group to join.[8]

Local consistency only requires consistency between neighbouring groups and doesn't put any restrictions on what people believe about the optimality of other people's decisions who are not in their neighbouring group. Formally, this is equivalent to

Definition 2.6 *A partition* $[S_{b_0}, S_{b_1}, .., S_{b_n}]$ *of* Y *and a menu of sorting fees* $[b_0, b_1, .., b_n]$ *satisfy* ***local consistency*** *iff* $\forall i \in \{0, .., n\}$

$$U(y, E_i^j(P)) - b_j \geq U(y, E_i^k(P)) - b_k \quad \forall y \in S_{b_j}, \forall k, \forall j \in \{i-1, i+1\} \cap \mathbb{N} \tag{LC}$$

It is straightforward to see that global consistency is a stricter requirement than local consistency.

Corollary 2.5 *Global consistency implies local consistency.*

[8] Another way to think about is that if people only consider members of neighbouring groups, this could be because societal interactions are such that people are - outside of their own group - most likely to interact with members of neighbouring groups, perhaps due to intersecting meeting points of neighbouring groups, such as supermarkets or schools etc.

A biased sorting equilibrium with local consistency is defined as follows:

Definition 2.7 *A* ***biased sorting equilibrium with local consistency*** *is a partition* $P = [S_{b_0}, S_{b_1}, .., S_{b_n}]$ *of* Y *and a menu sorting fees* $[b_0, b_1, ...b_n]$ *(with* $b_i < b_{i+1}$ $\forall i$ *and* $b_0 = 0$*) such that* $\forall i \in \{0, .., n\}$

$$U(y, E^i(P)) - b_i \geq U(y, E_i^k(P)) - b_k \geq b_i - b_k \quad \forall y \in S_{b_i}, \forall k \tag{IC}$$

$$U(y, E_i^j(P)) - b_j \geq U(y, E_i^k(P)) - b_k \quad \forall y \in S_{b_j}, \forall k, \forall j \in \{i-1, i+1\} \cap \mathbb{N} \tag{LC}$$

Requiring (local or global) consistency eliminates non-monotone biased sorting equilibria, as the following example demonstrates:

Example 2 *(Example 1 continued) Take again the example from before. I have showed above that the partition* $S_{b_0} = [0, \frac{1}{4}) \cup (\frac{1}{2}, \frac{3}{4}]$ *and* $S_{b_1} = [\frac{1}{4}, \frac{1}{2}] \cup \left(\frac{3}{4}, 1\right]$ *together with* $b_1 = \frac{1}{16}$*,* $b_0 = 0$*,* $E_1^1(P) = E^1(P)$*,* $E_1^0(P) = E^0(P)$*,* $E_0^1(P) = E_0^0(P) = E^0(P)$ *is a biased sorting equilibrium. However, local consistency*[9] *does not hold here: All people in* S_{b_0} *think that the people in* S_{b_1} *are wrong to pay* b_1 *and join* S_{b_1}*, and equally all people in* S_{b_1} *do not understand why the people with* $y \in (\frac{1}{2}, \frac{3}{4})$ *do not want to join* S_{b_1}*.*

The finding that non-monotone biased sorting equilibria do not satisfy local consistency can be generalized.

Proposition 2.2 *All biased sorting equilibria with local consistency satisfy monotonicity.*

Proof. Suppose a non-monotone equilibrium exists. Then there must be two neighbouring groups S_{b_i} and $S_{b_{i-1}}$ with $y_2 \in S_{b_{i-1}}$ and $y_1 \in S_{b_i}$ but $y_2 > y_1$. Then (IC) requires

$$b_i - b_{i-1} \geq U(y_2, E_{i-1}^i(P)) - U(y_2, E_{i-1}^{i-1}(P)) \tag{2.6}$$

and

$$U(y_1, E_i^i(P)) - U(y_1, E_i^{i-1}(P)) \geq b_i - b_{i-1}. \tag{2.7}$$

These two conditions combined give

$$U(y_1, E_i^i(P)) - U(y_1, E_i^{i-1}(P)) \geq U(y_2, E_{i-1}^i(P)) - U(y_2, E_{i-1}^{i-1}(P))$$

and as E_{i-1} and E_i can be chosen freely, this inequality can hold for some E_i and E_{i-1}, even though $y_1 < y_2$. However, the consistency requirement (LC) yields the additional conditions

$$U(y_1, E_{i-1}^i(P)) - U(y_1, E_{i-1}^{i-1}(P)) \geq b_i - b_{i-1} \tag{2.8}$$

and

$$b_i - b_{i-1} \geq U(y_2, E_i^i(P)) - U(y_2, E_i^{i-1}(P)). \tag{2.9}$$

But (2.7) and (2.9) together imply

$$U(y_1, E_i^i(P)) - U(y_1, E_i^{i-1}(P)) \geq U(y_2, E_i^i(P)) - U(y_2, E_i^{i-1}(P))$$

and (2.6) and (2.8) together imply

$$U(y_1, E_{i-1}^i(P)) - U(y_1, E_{i-1}^{i-1}(P)) \geq U(y_2, E_{i-1}^i(P)) - U(y_2, E_{i-1}^{i-1}(P)).$$

[9] Both global and local consistency do not hold in this case, because in the case of two groups these two notions are identical.

Neither of these inequalities can hold if $y_2 > y_1$ and U is strictly supermodular. ■

Proposition 2.2 shows that we do not need to assume global consistency to get monotonicity, local consistency is enough to ensure that even if people are biased there cannot be any other equilibria apart from monotone ones. Hence, global and local consistency are equivalent In terms of guaranteeing monotone equilibria. However, the difference between local and global consistency will play an important role in Section 2.6, when I analyze existence of biased sorting equilibria with more than two groups.

Proposition 2.2 implies that all biased sorting equilibria with consistency (both local and/or global) can be characterized in terms of cutoffs $\hat{y}_i$ and biased perceptions E_i.

Corollary 2.6 *A* ***biased sorting equilibrium with global consistency*** *is characterized by a partition* $[0, \hat{y}_1, ..., \hat{y}_n, y_{\max}]$ *of* Y *and a menu of sorting fees* $[b_0, b_1, .., b_n]$ *(with* $b_i < b_{i+1}$ $\forall i$ *and* $b_0 = 0$*) such that* $\forall i \in \{1, .., n\}$

$$U(\hat{y}_i, E_0^i(P)) - U(\hat{y}_i, E_0^{i-1}(P)) \tag{2.10}$$

$$= U(\hat{y}_i, E_1^i(P)) - U(\hat{y}_i, E_1^{i-1}(P))$$

$$= ... = U(\hat{y}_i, E_n^i(P)) - U(\hat{y}_i, E_n^{i-1}(P))$$

$$= b_i - b_{i-1} \quad \forall i \in \{1, .., n\}$$

Corollary 2.7 *A* ***biased sorting equilibrium with local consistency*** *is characterized by a partition* $[0, \hat{y}_1, ..., \hat{y}_n, y_{\max}]$ *of* Y *and a menu of sorting fees* $[b_0, b_1, .., b_n]$ *(with* $b_i < b_{i+1}$ $\forall i$ *and* $b_0 = 0$*) such that* $\forall i \in \{1, .., n\}$

$$U(\hat{y}_i, E^i(P)) - U(\hat{y}_i, E_i^{i-1}(P)) \tag{2.11}$$

$$= U(\hat{y}_i, E_{i-1}^i(P)) - U(\hat{y}_i, E^{i-1}(P))$$

$$= b_i - b_{i-1}$$

and[10]

$$U(y, E^i(P)) - b_i \geq U(y, E_i^k(P)) - b_k \quad \forall y \in S_{b_i}, \forall k \neq i \tag{2.12}$$

Proof. (IC) and (LC) imply that

$$U(y, E^i(P)) - b_i \geq U(y, E_i^{i-1}(P)) - b_{i-1} \quad \forall y \in S_{b_i}, \forall i$$

$$\iff U(y, E^i(P)]) - U(y, E_i^{i-1}(P)) \geq b_i - b_{i-1} \quad \forall y \in S_{b_i}, \forall i \tag{2.13}$$

and

$$U(y, E^i(P)) - b_i \leq U(y, E_i^{i-1}(P)) - b_{i-1} \quad \forall y \in S_{b_{i-1}}, \forall i$$

$$\iff U(y, E^i(P)) - U(y, E_i^{i-1}(P)) \leq b_i - b_{i-1} \quad \forall y \in S_{b_{i-1}}, \forall i \tag{2.14}$$

(2.13) and (2.14) together with the fact that $S_{b_i} = [\hat{y}_i, \hat{y}_{i+1}]$ and $S_{b_{i-1}} = [\hat{y}_{i-1}, \hat{y}_i]$ imply that

$$U(\hat{y}_i, E^i(P)) - U(\hat{y}_i, E_i^{i-1}(P)) = b_i - b_{i-1} \quad \forall i \tag{2.15}$$

Furthermore, (IC) and (LC) also imply that

$$U(y, E_{i-1}^i(P)) - b_i \geq U(y, E^{i-1}(P)) - b_{i-1} \quad \forall y \in S_{b_i}, \forall i$$

[10] Condition (2.11) ensures that condition (2.12) holds for $k \in \{i-1, i+1\}$ but we need to ensure that people prefer their group S_{b_i} to *any* other group S_{b_k}. This condition is not needed for the definition of biased sorting equilibrium with global consistency, because it is guaranteed by condition (2.10).

$$\iff U(y, E_{i-1}^{i}(P)) - U(y, E^{i-1}(P)) \geq b_i - b_{i-1} \quad \forall y \in S_{b_i}, \forall i \tag{2.16}$$

and

$$U(y, E_{i-1}^{i}(P)) - b_i \leq U(y, E^{i-1}(P)) - b_{i-1} \quad \forall y \in S_{b_{i-1}}, \forall i$$

$$\iff U(y, E_{i-1}^{i}(P)) - U(y, E^{i-1}(P)) \leq b_i - b_{i-1} \quad \forall y \in S_{b_{i-1}}, \forall i \tag{2.17}$$

(2.16) and (2.17) together with $S_{b_i} = [\hat{y}_i, \hat{y}_{i+1}]$ and $S_{b_{i-1}} = [\hat{y}_{i-1}, \hat{y}_i]$ imply that

$$U(\hat{y}_i, E_{i-1}^{i}(P)) - U(\hat{y}_i, E^{i-1}(P)) = b_i - b_{i-1} \quad \forall i \tag{2.18}$$

Combined, (2.15) and (2.18) give the equilibrium condition (2.11). ■

Hence, a biased sorting equilibrium with local consistency is a partition where the perceived benefit of being in group S_{b_i} rather than $S_{b_{i-1}}$ of the person just to the right of every cutoff $\hat{y}_i$ (which is $U(\hat{y}_i, E^i(P)) - U(\hat{y}_i, E_i^{i-1}(P))$) is the same as the perceived benefit of being in group S_{b_i} rather than $S_{b_{i-1}}$ of the person just to the left of every cutoff $\hat{y}_i$ ($U(\hat{y}_i, E_{i-1}^{i}(P)) - U(\hat{y}_i, E^{i-1}(P))$). The equilibrium conditions (2.11) and (2.12) restrict the sets of belief functions which imply equilibrium existence. In the next sections, I put more structure on the functional form of the belief functions and find conditions on these functions such that equilibrium exists and is unique. If the misperceptions are such that the biased sorting equilibrium partition with consistency is unique, it follows from (2.11) that the corresponding schedule of sorting fees will also be unique (as long as b_0 is normalized to 0).

2.5 Existence and uniqueness of binary biased sorting equilibria with consistency

What kind of biased sorting equilibria can exist for different types of misperceptions? In the following section, I will focus on (monotone) partitions with two groups $P = [S_0, S_b]$ of Y that can be uniquely characterized by a cutoff $\hat{y} \in Y$ (with the convention that $S_0 = [0, \hat{y})$ and $S_b = [\hat{y}, y_{\max}]$). I will henceforth call the people in S_0 "the poor" and the people in S_b "the rich".

The belief function B in the two-group case is a continuous function that maps all monotone partitions of Y (and note that any monotone partition can be uniquely characterized by the cutoff $\hat{y}$) into a four-dimensional vector of beliefs

$$B(\hat{y}) = (\underline{E}(\hat{y}), \bar{E}_p(\hat{y}), \underline{E}_r(\hat{y}), \bar{E}(\hat{y}))$$

where the first two entries denote the poor group's belief about average income in the poor and the rich group respectively and the last two entries denote the rich group's belief about average income in the poor and the rich group. $\underline{E}(\hat{y})$ is the true average income in the poor group, i.e. $\underline{E}(\hat{y}) = E[y|y < \hat{y}]$ and $\bar{E}(\hat{y})$ is the correct average income in the rich group, $\bar{E}(\hat{y}) = E[y|y \geq \hat{y}]$. The poor's belief about average income in the rich group is $\bar{E}_p(\hat{y})$ and the rich's belief about average income in the poor group is $\underline{E}_r(\hat{y})$.

In the following analysis I will restrict my attention to misperceptions where the direction of the bias does not vary with the cutoff, i.e. for any configuration of groups one group either always overestimates or underestimates average income in the other group and groups do not switch between over- and underestimating each other depending on group size or shape. Formally, this means I look at belief functions B that satisfy

$$\bar{E}_p(\hat{y}) < (>) \bar{E}(\hat{y}) \quad \forall \hat{y} \in [0, y_{\max})$$

and

$$\underline{E}_r(\hat{y}) > (<) \underline{E}(\hat{y}) \quad \forall \hat{y} \in (0, y_{\max}].$$

Note, though, that by Assumption 2.1 a group is correct about average income in the other group if they see the whole income distribution, i.e. the poor are correct at $y_{\max}$,

$$\bar{E}_p(\hat{y}) = \bar{E}(\hat{y}) \quad \text{if} \;\; \hat{y} = y_{\max},$$

and the rich are correct at 0,

$$\underline{E}_r(\hat{y}) = \underline{E}(\hat{y}) \quad \text{if} \;\; \hat{y} = 0.$$

Then there are four possible combinations of biases:

Case 1: $\underline{E}_r(\hat{y}) > \underline{E}(\hat{y})$ and $\bar{E}_p(\hat{y}) < \bar{E}(\hat{y}) \;\; \forall \hat{y} \in (0, y_{\max})$: The rich overestimate average income in the poor group and the poor underestimate average income in the rich group.

Case 2: $\underline{E}_r(\hat{y}) < \underline{E}(\hat{y})$ and $\bar{E}_p(\hat{y}) > \bar{E}(\hat{y}) \;\; \forall \hat{y} \in (0, y_{\max})$: The rich underestimate average income in the poor group and the poor overestimate average income in the rich group.

Case 3: $\underline{E}_r(\hat{y}) < \underline{E}(\hat{y})$ and $\bar{E}_p(\hat{y}) < \bar{E}(\hat{y}) \;\; \forall \hat{y} \in (0, y_{\max})$: Both groups underestimate each other's average income

Case 4: $\underline{E}_r(\hat{y}) > \underline{E}(\hat{y})$ and $\bar{E}_p(\hat{y}) > \bar{E}(\hat{y}) \;\; \forall \hat{y} \in (0, y_{\max})$: Both groups overestimate each other's income.

In the first case, both groups underestimate the difference between groups, while in the second case they both overestimate it. In the latter two cases the misperceptions work in opposite directions for the two groups: one group overestimates the difference, the other group underestimates it. In Appendix 2.8.1 I analyze these four combinations to see whether biased sorting equilibria (with and without consistency) can exist and I provide sufficient conditions for existence and uniqueness. (Note that global and local consistency are the same concept in the case of two groups, which is why I talk only about "consistency" in this section.) It turns out that interior biased sorting equilibria with consistency can only exist in two of the four possible combinations: either both groups think the other group is more similar to themselves or both groups think the other group is more different to themselves. The reason for this is that the equilibrium condition (2.11) requires both groups to have the same perception of the benefits of sorting at the equilibrium cutoff, and hence it cannot be the case that one group underestimates the difference between the groups for any cutoff, while the other group overestimates it.

Proposition 2.3 *No cutoff $\hat{y} \in Y$ constitutes a biased sorting equilibrium in Case 4.*

Proposition 2.4 *Any cutoff $\hat{y} \in Y$ constitutes a biased sorting equilibrium in Case 3, but no cutoff $\hat{y} \in Y$ constitutes a biased sorting equilibrium with consistency.*

Proof. See Appendix 2.8.1. ∎

While biased sorting equilibria with consistency cannot exist in Case 3 and 4, it turns out that they will always exist in Case 1 and 2.

Proposition 2.5 *There always exists a biased sorting equilibrium with consistency in Case 1 and 2.*

Proof. See Appendix 2.8.1. ∎

The intuition for why Case 1- and Case 2-type misperceptions guarantee the existence of equilibria is that the form of the misperceptions ensures that the groups' perceived benefits of

sorting cross at least once: At $\hat{y} = 0$ the poor's perceived benefits of sorting are lower (higher) than the rich's, and the reverse is true at $\hat{y} = y_{\max}$. As all functions are continuous, there must be an interior cutoff $\hat{y} \in (0, y_{\max})$ such that they are equal.
If B and $U(.,.)$ are such that the perceived benefits of sorting of the two groups intersect only once, the interior equilibrium cutoff is unique. Sufficient conditions for this are stated in Appendix IC2. If U is linear in both arguments[11], the sufficient conditions for uniqueness simplify to the following:

Proposition 2.6 *(Case 1 Uniqueness) If $U(.,.)$ is linear and people are biased according to Case 1, the biased sorting equilibrium with consistency is unique if the misperceptions monotonically converge to the truth, i.e.*

$$\frac{d(\bar{E}(\hat{y}) - \bar{E}_p(\hat{y}))}{d\hat{y}} < 0 \quad \textit{and} \quad \frac{d(\underline{E}_r(\hat{y}) - \underline{E}(\hat{y}))}{d\hat{y}} > 0 \quad \forall \hat{y} \in (0, y_{\max})$$

Proof. See Appendix IC2. ■

Proposition 2.7 *(Case 2 Uniqueness) If $U(.,.)$ is linear and people are biased according to Case 2, the interior equilibrium is unique if the misperceptions monotonically converge to the truth, i.e.*

$$\frac{d(\bar{E}_p(\hat{y}) - \bar{E}(\hat{y}))}{d\hat{y}} < 0 \quad \textit{and} \quad \frac{d(\underline{E}(\hat{y}) - \underline{E}_r(\hat{y}))}{d\hat{y}} > 0 \quad \forall \hat{y} \in (0, y_{\max})$$

Proof. See Appendix IC2. ■
The condition that the misperceptions converge to the truth monotonically as the cutoff goes to 0 resp. $\hat{y}$, i.e.

$$\frac{d\left|\bar{E}(\hat{y}) - \bar{E}_p(\hat{y})\right|}{d\hat{y}} < 0 \textit{ and } \frac{d\left|\underline{E}_r(\hat{y}) - \underline{E}(\hat{y})\right|}{d\hat{y}} > 0$$

can be interpreted as people being less biased, the more they see of the income distribution: as $\hat{y}$ increases, the poor see a bigger part of the income distribution and their belief about average income in the other group becomes more accurate. The opposite happens for the rich: as the cutoff increases, they see a smaller part of society and therefore become more biased. In the limit this yields what I have already assumed in Assumption 2.1: the poor are correct at $y_{\max}$ and the rich at 0.

2.6 Existence of biased sorting equilibria with consistency and more than two groups

In the following section, I will assume that the utility function $U(.,.)$ is linear in both arguments, because it greatly simplifies my analysis. However, the results would not change qualitatively for a general strictly increasing and strictly supermodular utility function.
Whether biased sorting equilibria with consistency can exist and will be unique in the case of more than two groups depends, as in the two-group case, on the belief function. In addition, existence and uniqueness depend in general on the underlying income distribution. Finally, for more than two groups it also matters whether we want to require local or global consistency (while in the two-group case those two concepts are the same). For example, if the belief function is such that the perceived difference between group average incomes decreases in the distance (in terms of average income) of one's own group to the observed groups

[11] If U is linear then it is not strictly increasing whenever one of the arguments is 0. See Appendix IC2 for a specific analysis of linear utility functions at 0.

(perhaps because all groups that are far away from one's own group appear roughly similar), no equilibrium with more than two groups can exist if global consistency is required. The reason is that under global consistency, perceived differences between groups have to be the same across groups. For example, in the case of three groups the equilibrium conditions for a sorting equilibrium with global consistency would be[12]

$$E_0^1 - E_0^0 = E_1^1 - E_1^0 = E_2^1 - E_2^0$$

and

$$E_0^2 - E_0^1 = E_1^2 - E_1^1 = E_2^2 - E_2^1.$$

If the perceived difference in group averages is decreasing with distance, these equalities obviously cannot hold (for example, we would have $E_0^2 - E_0^1 < E_1^2 - E_1^1$ etc.). The same problem would occur for four or more groups. Hence, biased sorting equilibria with global consistency and more than two groups are not possible if the belief function specifies decreasing perceived group average differences. Of course, the same result holds for belief functions with increasing perceived group average differences.
On the other hand, biased sorting equilibria with local consistency are still possible with the above beliefs, because non-neighbouring groups don't matter in this case. However, it is important that neighbouring groups assess the differences between each other in the same way, so beliefs cannot be such that one group overestimates the difference to a neighbouring group, while this neighbouring group underestimates it.
In the following analysis I want to focus on a particular form of beliefs: belief functions which are such that misperceptions about a group's average income do not depend on whether this group is a neighbouring group or not, and thus the severity of the bias depends only on group size, irrespective of which group people look at (and so e.g. people in the poorest group are "equally" biased about the average income in all other groups, no matter whether it is their neighbouring group or a very rich group). Specifically, I will for the rest of this section focus on beliefs that are of a dertain functional form, which I will call *Proportional Biased Beliefs*. *Proportional Biased Beliefs of Type 1* are such that all groups underestimate the differences between their own group and other groups, whereas *Proportional Biased Beliefs of Type 2* are such that all groups overestimate the differences between their own group and other groups.

Definition 2.8 *A belief function $B(P)$ generates Proportional Biased Beliefs of Type 1 if group i's belief about average income in group $j < i$ are*[13]

$$E_i^j = \beta(1 - F_{i+1} + F_i)\hat{y}_{j+1} + (1 - \beta(1 - F_{i+1} + F_i))E^j \ \forall i \forall j$$

and group i's beliefs about average income in group $k > i$ are

$$E_i^k = \beta(1 - F_{i+1} + F_i)\hat{y}_k + (1 - \beta(1 - F_{i+1} + F_i))E^k \ \forall i \forall k.$$

Definition 2.9 *A belief function $B(P)$ generates Proportional Biased Beliefs of Type 2 if group i's belief about average income in group $j < i$ are*

$$E_i^j = \beta(1 - F_{i+1} + F_i)\hat{y}_j + (1 - \beta(1 - F_{i+1} + F_i))E^j \ \forall i \forall j$$

[12] To improve readability I will from now on write E_i^j instead of $E_i^j(P)$, but the group beliefs continue to be functions of the partition.
[13] To improve readability I will from now on write F_i instead of $F(\hat{y}_i)$.

and group i's beliefs about average income in group $k > i$ are

$$E_i^k = \beta(1 - F_{i+1} + F_i)\hat{y}_{k+1} + (1 - \beta(1 - F_{i+1} + F_i))E^k \ \forall i \forall k.$$

Remark 2.1 *Note that by "group i" I mean people in group S_{b_i}, i.e. people with income between cutoffs $\hat{y}_i$ and $\hat{y}_{i+1}$. Average income between y_i and y_{i+1} is denoted as E^i.*

With Proportional Biased Beliefs of Type 1 (Type 2), people tend to underestimate (overestimate) differences between their own group and other groups (in terms of average group income), and their misperceptions are more severe, the smaller their group (i.e. the less they see of the whole income distribution).
These beliefs satisfy the necessary and sufficient conditions for existence and uniqueness of binary biased sorting equilibria as stated in Section 2.5: They are either Case 1- or Case 2- type misperceptions and satisfy Assumption 2.1. In Appendix 2.8.2 I show that with Proportional Biased Beliefs, biased sorting equilibria with global consistency do not exist for more than three groups. The reason for this is of a technical nature: With global consistency and this particular specification of the bias, everybody must have the same perceived benefits of being in one group versus being in another group. However, as everybody knows the average income in their own group, but misperceives average incomes in the other groups, this creates a difference between perceived benefits of being in one's own group versus being in any other group, and perceived benefits of being in one group versus another group while not being a member of either group. As these different perceived benefits have to be the same across groups, this yields a contradiction. For three groups, existence with global consistency and the above defined bias depends on the underlying distribution. For instance, if the income distribution is uniform, an equilibrium partition with three groups cannot exist, as shown below.
In contrast, it turns out that with the same belief function, there *can* be more than three groups if only local consistency is required. For example, if the income distribution is uniform, any partition with equidistant (finitely many) cutoffs is a biased sorting equilibrium with local consistency. I prove this in Appendix 2.8.2.
Let me illustrate these findings with a simple example: Suppose y is uniformly distributed on $[0, a]$. A biased sorting equilibrium with local consistency and three groups needs to satisfy the following equilibrium conditions

$$E_0^1 - E_0^0 = E_1^1 - E_1^0 \tag{2.19}$$
$$E_1^2 - E_1^1 = E_2^2 - E_2^1 \tag{2.20}$$

and

$$\hat{y}_1(E_0^2 - E_0^0) \leq b_2 \tag{2.21}$$

$$\hat{y}_2(E^2 - E_2^0) \geq b_2 \tag{2.22}$$

with

$$b_2 = \hat{y}_2(E^2 - E_2^1) + b_1 = \hat{y}_2(E_1^2 - E^1) + b_1$$

and

$$b_1 = \hat{y}_1(E^1 - E_1^0) = \hat{y}_1(E_0^1 - E^0).$$

Suppose that people have proportional biased beliefs of type 1, i.e.

$$\begin{aligned}
E_0^0 &= E^0 \\
E_0^1 &= \beta(1-F(\hat{y}_1))\hat{y}_1 + (1-\beta(1-F(\hat{y}_1)))\,E^1 \\
E_0^2 &= \beta(1-F(\hat{y}_1))\hat{y}_2 + (1-\beta(1-F(\hat{y}_1)))\,E^2 \\
E_1^0 &= \beta(1-F(\hat{y}_2)+F(\hat{y}_1))\hat{y}_1 + (1-\beta(1-F(\hat{y}_2)+F(\hat{y}_1)))\,E^0 \\
E_1^1 &= E^1 \\
E_1^2 &= \beta(1-F(\hat{y}_2)+F(\hat{y}_1))\hat{y}_2 + (1-\beta(1-F(\hat{y}_2)+F(\hat{y}_1)))\,E^2 \\
E_2^0 &= \beta F(\hat{y}_2)\hat{y}_1 + (1-\beta F(\hat{y}_2))E^0 \\
E_2^1 &= \beta F(\hat{y}_2)\hat{y}_2 + (1-\beta F(\hat{y}_2))E^1 \\
E_2^2 &= E^2
\end{aligned}$$

Given the functional form of the misperceptions, the equilibrium conditions (2.19) and (2.20) yield

$$\hat{y}_1 = \frac{(1-F_2+F_1)E^0 + (1-F_1)E^1}{2-F_2}$$

and

$$\hat{y}_2 = \frac{F_2E^1 + (1-F_2+F_1)E^2}{1+F_1}$$

Using the fact that F is uniformly distributed on $[0,a]$ yields the unique solution

$$\begin{aligned}
\hat{y}_1 &= \frac{a}{3} \\
\hat{y}_2 &= \frac{2a}{3}.
\end{aligned}$$

Note that this implies that in equilibrium the three groups will all be of equal size, i.e.

$$F_1 = F_2 - F_1 = 1 - F_2 = \frac{1}{3}.$$

Also conditions (2.21) and (2.22) are satisfied in this case: with equidistant cutoffs and a uniform distribution we get

$$b_1 = \left(\frac{a}{3}\right)^2\left[1-\beta\frac{1}{3}\right]$$

and

$$b_2 = 3\left(\frac{a}{3}\right)^2\left[1-\beta\frac{1}{3}\right] = 3b_1$$

and hence condition (2.21) becomes

$$\frac{a}{3}\left[\frac{5a}{6} - \beta\frac{2}{3}\frac{a}{6} - \frac{a}{6}\right] \leq 3\left(1-\beta\frac{1}{3}\right)\left(\frac{a}{3}\right)^2$$

which reduces to

$$\beta\frac{2}{3} \leq 1$$

and is therefore always satisfied. Condition (2.22) becomes

$$\frac{2a}{3}\left[\frac{5a}{6} - \frac{a}{6} - \beta\frac{2}{3}\frac{a}{6}\right] \geq 3\left(1-\beta\frac{1}{3}\right)\left(\frac{a}{3}\right)^2$$

which reduces to

$$\beta\frac{1}{3} \geq -1$$

which always holds. Hence, for the uniform distribution there exists a unique biased sorting equilibrium with local consistency and three groups if the misperceptions are as defined above. However, with those same misperceptions, there doesn't exist a sorting equilibrium with three groups if global consistency is required. Global consistency yields the equilibrium conditions

$$E_0^1 - E_0^0 = E_1^1 - E_1^0 = E_2^1 - E_2^0 \tag{2.23}$$

and

$$E_0^2 - E_0^1 = E_1^2 - E_1^1 = E_2^2 - E_2^1. \tag{2.24}$$

If F is uniform, this translates to the following four conditions:

$$E^1 + \frac{(2F_2 - F_1 - 1)}{F_2}(\hat{y}_1 - E^0) = \hat{y}_2 \tag{2.25}$$

$$\frac{2F_1 - F_2}{1 - F_1}\left(E^2 - \hat{y}_2\right) + E^1 = \hat{y}_1 \tag{2.26}$$

$$\hat{y}_1 = \frac{(1 - F_2 + F_1)E^0 + (1 - F_1)E^1}{2 - F_2} \tag{2.27}$$

$$\hat{y}_2 = \frac{F_2 E^1 + (1 - F_2 + F_1)E^2}{1 + F_1} \tag{2.28}$$

Conditions (2.27) and (2.28) are the same as for the equilibrium with local consistency, and we know they yield the unique solution

$$\begin{aligned} \hat{y}_1 &= \frac{a}{3} \\ \hat{y}_2 &= \frac{2a}{3} \end{aligned}$$

with groups of equal size. However, from (2.25) it follows that

$$2F_2 - F_1 - 1 > 0$$

in equilibrium, because y_2 must be greater than E^1. This can be rewritten as

$$F_2 - F_1 > 1 - F_2,$$

which is a contradiction to groups being of equal size. Therefore, no biased sorting equilibrium with global consistency and three groups can exist for the uniform distribution if misperceptions are defined as above.

2.7 Conclusion

In this paper I have introduced a new framework for modelling situations in which beliefs and group choice interact endogenously: a model of sorting in the presence of misperceptions that takes into account the two-way interaction between beliefs about society and social segregation. Furthermore, I have defined a new equilibrium concept, the biased sorting equilibrium, which can be interpreted as characterizing partitions which are stable given the misperceptions that people acquire once they are segregated and interact only with members of their own social group. I have also introduced a refinement concept, the consistency requirement, which can be stated in a local and a global version, and I have showed that it guarantees that the biased sorting equilibrium partitions will be monotone. It also adds structure to the model:

not every monotone partition of the income space can constitute a biased sorting equilibrium with consistency - existence and uniqueness of equilibrium will depend on the functional form that is used to model people's beliefs about the other groups. For the two-group case I have proved that the consistency requirement guarantees uniqueness of the equilibrium partition if people's misperceptions are such that they converge to the truth monotonically as the size of their group increases. In the case of multiple groups, the main contribution of this paper is to demonstrate that existence of equilibria depends both on the functional form of the belief function and on what type of consistency (local or global) is required. Focusing on a specific functional form of the misperceptions, which I call Proportional Biased Beliefs, I have demonstrated that global consistency precludes existence of equilibria with more than two groups and I have examined conditions for equilibria with local consistency. My findings in this section show that the types of stable partitions that are possible if we allow for more than two groups in equilibrium depend on the situation we want to model, and how we assume the sorting process to look like.
To conclude, let me note that the framework that I have developed in this paper is very general and lends itself to many further applications. In the present paper, the variable according to which people sort is income, but any other variable where everybody in society agrees that "more is better" would work as well. For instance, if we want to examine school choice, the sorting variable could be "ability" or "intelligence". With a different utility function (e.g. one that accounts for homophily), the same framework could be applied to model ethnic, religious or cultural segregation.

2.8 Appendix

2.8.1 Necessary and sufficient conditions for existence and uniqueness of a binary biased sorting equilibrium with consistency

In this section I will derive necessary and sufficient conditions for existence and uniqueness of a binary biased sorting equilibrium with consistency in the four cases mentioned in Section 2.5.
The conditions for a binary biased sorting equilibrium can be written as

$$U(y,\bar{E}_p(\hat{y}^*)) - U(y,\underline{E}(\hat{y}^*)) \quad < \quad b \;\; \forall y < \hat{y}^* \tag{IC1}$$
$$U(y,\bar{E}(\hat{y}^*)) - U(y,\underline{E}_r(\hat{y}^*)) \quad \geq \quad b \;\; \forall y \geq \hat{y}^* \tag{IC2}$$

for some $b > 0$. Due to supermodularity of U, these two conditions can be simplified to

$$\begin{aligned} U(\hat{y}^*,\bar{E}_p(\hat{y}^*)) - U(\hat{y}^*,\underline{E}(\hat{y}^*)) \quad &\leq \quad b \\ U(\hat{y}^*,\bar{E}(\hat{y}^*)) - U(\hat{y}^*,\underline{E}_r(\hat{y}^*)) \quad &\geq \quad b \end{aligned}$$

which implies that at the equilibrium cutoff we must have

$$U(\hat{y}^*,\bar{E}(\hat{y}^*)) - U(\hat{y}^*,\underline{E}_r(\hat{y}^*)) \geq U(\hat{y}^*,\bar{E}_p(\hat{y}^*)) - U(\hat{y}^*,\underline{E}(\hat{y}^*)) \tag{2.29}$$

(and note that as $U(\hat{y}^*,\bar{E}(\hat{y}^*)) - U(\hat{y}^*,\underline{E}_r(\hat{y}^*)) \geq b$ the RHS of this inequality must be strictly positive in any biased sorting equilibrium).
Inequality (2.29) says that at any equilibrium cutoff it must be the case that the rich group's perceived benefit of sorting (LHS) is greater than the poor group's perceived benefit of sorting (RHS). It follows immediately that we cannot find a $b > 0$ such that a biased sorting equi-

librium exists at any cutoff in Case 4, in which the rich group underestimates the difference between groups for any cutoff and the poor group overestimates the difference between groups for any cutoff. The reason is that the fact that U is strictly increasing in both arguments implies that the rich's perceived benefit of sorting lies below the poor's benefit of sorting for every cutoff $\hat{y}$ in this case, and therefore no positive sorting fee can be found such that the rich would be willing to pay the fee and be in the rich group and the poor would not be willing to join.

Proposition 2.8 *No cutoff $\hat{y} \in Y$ constitutes a biased sorting equilibrium in Case 4.*

Proof. The fact that in Case 4 we have $\bar{E}(\hat{y}) < \bar{E}_p(\hat{y})$ and $\underline{E}_r(\hat{y}) > \underline{E}(\hat{y})$ $\forall \hat{y} \in (0, y_{\max})$ combined with strict increasingness of U implies that inequality (2.29) cannot be satisfied for any $\hat{y} \in Y$. ■

The opposite to Case 4 happens in Case 3: if the rich group overestimates the group difference for every cutoff, while the poor group underestimates it, any cutoff $\hat{y} \in Y$ is a biased sorting equilibrium.

Proposition 2.9 *Any cutoff $\hat{y} \in Y$ constitutes a biased sorting equilibrium in Case 3.*[14]

Proof. Case 3 implies that $\bar{E}(\hat{y}) > \bar{E}_p(\hat{y})$ and $\underline{E}_r(\hat{y}) < \underline{E}(\hat{y})$ $\forall \hat{y} \in (0, y_{\max})$. Together with the fact that U is strictly increasing in both arguments, this implies that inequality (2.29) holds for all $\hat{y} \in Y$. ■

If we require the equilibrium partition to satisfy consistency, the equilibrium cutoff must satisfy

$$U(\hat{y}^*, \bar{E}(\hat{y}^*)) - U(\hat{y}^*, \underline{E}_r(\hat{y}^*)) = U(\hat{y}^*, \bar{E}_p(\hat{y}^*)) - U(\hat{y}^*, \underline{E}(\hat{y}^*)) = b \tag{2.30}$$

(and therefore both differences must be strictly positive because $b > 0$).

Proposition 2.10 *No cutoff $\hat{y} \in Y$ constitutes a biased sorting equilibrium with consistency in Case 3.*

Proof. Case 3 implies that $\bar{E}(\hat{y}) < \bar{E}_p(\hat{y})$ and $\underline{E}_r(\hat{y}) > \underline{E}(\hat{y})$ $\forall \hat{y} \in (0, y_{\max})$. If U is strictly increasing in both arguments, condition (2.30) cannot be satisfied for any $\hat{y} \in Y$. ■

Only Case 1 and Case 2 allow for the existence of an interior biased sorting equilibrium with consistency. In fact, I find that in those two cases an interior equilibrium always exists:

Proposition 2.11 *A biased sorting equilibrium with consistency always exists in Case 1 and Case 2.*

Proof. Remember that the rich are correct at 0 and the poor at $y_{\max}$. Therefore, it holds that

$$U(0, \bar{E}(0)) - U(0, \underline{E}_r(0)) > (<) U(0, \bar{E}_p(0)) - U(0, \underline{E}(0))$$

and

$$U(y_{\max}, \bar{E}(y_{\max})) - U(y_{\max}, \underline{E}_r(y_{\max})) < (>) U(y_{\max}, \bar{E}_p(y_{\max})) - U(y_{\max}, \underline{E}(y_{\max})).$$

As both U and the belief function are continuous, there must be an interior $\hat{y}^* \in (0, y_{\max})$ such that

$$U(\hat{y}^*, \bar{E}(\hat{y}^*)) - U(\hat{y}^*, \underline{E}_r(\hat{y}^*)) = U(\hat{y}^*, \bar{E}_p(\hat{y}^*)) - U(\hat{y}^*, \underline{E}(\hat{y}^*)).$$

[14] If I would be meticulously diligent, I would have to exclude $\hat{y} = 0$ as a potential equilibrium cutoff in this and all following Propositions. The reason is that strictly speaking $\hat{y} = 0$ cannot be an equilibrium cutoff due to my definition of the partition as $\{[0, \hat{y}), [\hat{y}, y_{\max}]\}$. $\hat{y} = 0$ would imply that S_0 would be empty, which is not possible because the empty set cannot be an element of a partition. Therefore, $\hat{y} = 0$ is technically not even included in my sorting equilibrium definition. If I had defined the partition the other way round, i.e. $\{[0, \hat{y}], (\hat{y}, y_{\max}]\}$, then the same would hold for $y_{\max}$.

■

Linear utility function

If U is linear in both arguments, it is strictly speaking not in my set of analyzed utility functions, because it is not strictly increasing (and also not strictly supermodular) whenever one of the arguments is 0. If $U(x,y) = xy$ the equilibrium condition translates to

$$\hat{y}^*[\bar{E} - \underline{E}_r] = \hat{y}^*[\bar{E}_p - \underline{E}].$$

It is immediate to see that this condition will always be satisfied at $\hat{y} = 0$ (also for Case 3 and Case 4). However, $\hat{y} = 0$ cannot be a biased sorting equilibrium cutoff according to my definition, because I require the sorting fee b to be strictly positive, and hence $U(\hat{y}^*, \bar{E}(\hat{y}^*)) - U(\hat{y}^*, \underline{E}_r(\hat{y}^*))$ must be strictly positive in any equilibrium (this follows from condition (IC2)) Therefore, $\hat{y} = 0$ can never constitute a biased sorting equilibrium cutoff if U is linear.
For reasons of completeness, I therefore restate Propositions 2.8 - 2.10 for a linear utility function (Proposition 2.11 doesn't change):

Proposition 2.12 *No cutoff $\hat{y} \in Y$ constitutes a biased sorting equilibrium in Case 4.*

Proof. The fact that in Case 4 we have $\bar{E}(\hat{y}^*) < \bar{E}_p(\hat{y}^*)$ and $\underline{E}_r(\hat{y}^*) > \underline{E}(\hat{y}^*)$ $\forall \hat{y}^* \in (0, y_{\max})$ combined with strict increasingness of U implies that inequality (2.29) cannot be satisfied for any $\hat{y} \in Y$. If U is linear (and hence not strictly increasing at $\hat{y} = 0$) then inequality (2.29) is trivially satisfied for $\hat{y} = 0$, but $\hat{y} = 0$ cannot be a biased sorting equilibrium cutoff according to my definition, because I require the sorting fee b to be strictly positive, and hence $U(\hat{y}^*, \bar{E}(\hat{y}^*)) - U(\hat{y}^*, \underline{E}_r(\hat{y}^*))$ must be strictly positive in any equilibrium. ■

Proposition 2.13 *Any cutoff $\hat{y} \in (0, y_{\max}]$ constitutes a biased sorting equilibrium in Case 3.*

Proof. Case 3 implies that $\bar{E}(\hat{y}^*) > \bar{E}_p(\hat{y}^*)$ and $\underline{E}_r(\hat{y}^*) < \underline{E}(\hat{y}^*)$ $\forall \hat{y}^* \in (0, y_{\max})$. If U is strictly increasing in both arguments, this implies that inequality (2.29) holds for all $\hat{y} \in Y$. If U is linear (and hence not strictly increasing at $\hat{y} = 0$) then inequality (2.29) is trivially satisfied for $\hat{y} = 0$, but $\hat{y} = 0$ cannot be a biased sorting equilibrium cutoff. ■

Proposition 2.14 *No cutoff $\hat{y} \in (0, y_{\max})$ constitutes a biased sorting equilibrium with consistency in Case 3.*

Proof. Case 3 implies that $\bar{E}(\hat{y}^*) < \bar{E}_p(\hat{y}^*)$ and $\underline{E}_r(\hat{y}^*) > \underline{E}(\hat{y}^*)$ $\forall \hat{y}^* \in (0, y_{\max})$. If U is strictly increasing everywhere, condition (2.30) cannot be satisfied for any $\hat{y} \in Y$. If U is linear then the first equality of condition (2.30) is trivially satisfied for $\hat{y} = 0$, but the second equality can never be satisfied, because $b > 0$.. ■

Uniqueness: Linear utility function

For Case 1 and Case 2, the following sufficient conditions for uniqueness of an interior biased sorting equilibrium with consistency can be stated if U is linear:

Proposition 2.15 *If people are biased according to Case 1 or Case 2, U is linear in both arguments and people's misperceptions converge to the truth monotonically, i.e.*

$$\frac{d\left|\bar{E}(\hat{y}) - \bar{E}_p(\hat{y})\right|}{d\hat{y}} < 0 \quad \text{and} \quad \frac{d\left|\underline{E}_r(\hat{y}) - \underline{E}(\hat{y})\right|}{d\hat{y}} > 0 \;\; \forall \hat{y} \in (0, y_{\max}) \tag{2.31}$$

there always exists a unique biased sorting equilibrium with consistency.

Proof. The equilibrium cutoff must satisfy

$$\bar{E}(\hat{y}^*) - \underline{E}_r(\hat{y}^*) = \bar{E}_p(\hat{y}^*) - \underline{E}(\hat{y}^*).$$

Suppose people are biased according to Case 1 (the argument can be made analogously for Case 2). Then the conditions in (2.31) become

$$\frac{d(\bar{E}(\hat{y}) - \bar{E}_p(\hat{y}))}{d\hat{y}} < 0 \quad \text{and} \quad \frac{d(\underline{E}_r(\hat{y}) - \underline{E}(\hat{y}))}{d\hat{y}} > 0 \;\; \forall \hat{y} \in (0, y_{\max})$$

This implies that the distance between the correct difference in average group incomes, $\underline{E}(\hat{y}) - \underline{E}(\hat{y})$, and the poor's perceived group difference, $\bar{E}_p(\hat{y}) - \underline{E}(\hat{y})$, which can be written as $\bar{E}(\hat{y}) - \bar{E}_p(\hat{y})$, is monotonically decreasing in $\hat{y}$, while the opposite holds for the distance between the correct group difference and the rich's perceived group difference (which can be written as $\underline{E}_r(\hat{y}) - \underline{E}(\hat{y})$). This means that there can be only one $\hat{y}$ for which the distance between the correct group differences and the group's perceived group differences is the same, and therefore (2.31) guarantees that the perceived benefits of sorting of the rich and of the poor only cut once in the interval $(0, y_{\max})$. ■

Uniqueness in Case 1: General utility function

For a general utility function, we also need to impose conditions on the shape of the utility function to ensure uniqueness.

Proposition 2.16 *If people are biased according to Case 1 and people's misperceptions converge to the truth monotonically, i.e.*

$$\frac{d\left|\bar{E}(\hat{y}) - \bar{E}_p(\hat{y})\right|}{d\hat{y}} < 0 \quad \textit{and} \quad \frac{d\left|\underline{E}_r(\hat{y}) - \underline{E}(\hat{y})\right|}{d\hat{y}} > 0 \quad \forall \hat{y} \in (0, y_{\max})$$

and additionally it holds that at any $\hat{y}^*$ for which

$$U(\hat{y}^*, \bar{E}(\hat{y}^*)) - U(\hat{y}^*, \underline{E}_r(\hat{y}^*)) = U(\hat{y}^*, \bar{E}_p(\hat{y}^*)) - U(\hat{y}^*, \underline{E}(\hat{y}^*))$$

holds we have that

$$U_1(\hat{y}^*, \bar{E}(\hat{y}^*)) - U_1(\hat{y}^*, \underline{E}_r(\hat{y}^*)) \leq U_1(\hat{y}^*, \bar{E}_p(\hat{y}^*)) - U_1(\hat{y}^*, \underline{E}(\hat{y}^*))$$

$$U_2(\hat{y}^*, \underline{E}_r(\hat{y}^*)) \geq U_2(\hat{y}^*, \underline{E}(\hat{y}^*))$$

and

$$U_2(\hat{y}^*, \bar{E}(\hat{y}^*)) \leq U_2(\hat{y}^*, \bar{E}_p(\hat{y}^*))$$

there always exists a unique biased sorting equilibrium with consistency.

Proof. At any equilibrium cutoff $\hat{y}^*$ such that

$$U(\hat{y}^*, \bar{E}(\hat{y}^*)) - U(\hat{y}^*, \underline{E}_r(\hat{y}^*)) = U(\hat{y}^*, \bar{E}_p(\hat{y}^*)) - U(\hat{y}^*, \underline{E}(\hat{y}^*))$$

these conditions ensure that the slope of the function on the LHS is smaller than the slope of the function on the RHS, which implies that the two functions can only intersect once. ■

Remark 2.2 *The conditions in Proposition 2.16 are sufficient conditions for uniqueness, because they ensure that the slope of the left hand side of the equilibrium condition is strictly smaller than the slope of the right hand side at any intersection.*

Remark 2.3 *Similar sufficient conditions can be found for Case 2 type misperceptions, where both groups overestimate group differences.*

2.8.2 Biased sorting equilibria with more than two groups: General proofs

Proof that with the uniform distribution, arbitrarily many groups are possible with Proportional Biased Beliefs

Suppose people have proportional biased beliefs of type 1. If a partition $[0, \hat{y}_1, .., \hat{y}_{n-1}, y_{\max}]$ is a biased sorting equilibrium with local consistency, it needs to satisfy

$$E^i - E_i^{i-1} = E_{i-1}^i - E^{i-1} \quad \forall i \tag{2.32}$$

and

$$yE^i(P) - yE_i^k(P) \geq b_i - b_k \quad \forall y \in S_{b_i}, \forall i, \forall k \neq i. \tag{2.33}$$

Given the functional form of the misperceptions, (2.32) can be written as

$$(\hat{y}^i - E^{i-1})(1 - F_{i+1} + F_i) = (1 - F_i + F_{i-1})(E^i - \hat{y}_i).$$

If the income distribution is uniform on $[0, a]$, this expression becomes

$$\frac{(a - \hat{y}_{i+1} + \hat{y}_i)}{a}\left(\frac{\hat{y}_i - \hat{y}_{i-1}}{2}\right) = \frac{(a - \hat{y}_i + \hat{y}_{i-1})}{a}\left(\frac{\hat{y}_{i+1} - \hat{y}_i}{2}\right) \quad \forall i$$

which, after simplifying, yields

$$\hat{y}_i = \frac{\hat{y}_{i+1} + \hat{y}_{i-1}}{2} \quad \forall i$$

and therefore that all cutoffs need to be equidistant in equilibrium. If all cutoffs are equidistant, the cutoffs will all be of the form

$$y_i = \frac{ia}{n}$$

(if there's n groups in total). Also, because all groups are of equal size (in terms of F), their misperceptions about the other groups are equally severe. Hence, all groups below a given group i have equal (wrong) beliefs about average income in the groups above group i and equally, all groups above group i have equal (wrong) beliefs about the groups below group i (because the severity of the misperceptions depends on group size and all groups are of equal size (in terms of F) in this case). Therefore, in order to check whether condition (2.33) holds, it suffices to look at one group (e.g. group 0) and check whether everybody in this group prefers staying where they are to switching to any other group with higher average income, and vice versa whether everybody in those groups prefers staying to switching to group 0. If group 0 wants to stay where they are and no other groups want to switch to group 0, then this is also satisfied for any other group i, because what matters for this decision is not the group's location on the income line itself, but rather the distance (in terms of F) to the other groups. As cutoffs are equidistant, all groups consider the same range of distances to each other when they decide whether they want to switch or not, so it suffices to consider one "model" group's decisions.

With the uniform distribution and equidistant cutoffs, it turns out that all b_is are of the form

$$b_i = \left(\frac{a}{n}\right)^2 \left(1 - \beta\frac{n-1}{2n}\right) \frac{i(i+1)}{2}$$

Furthermore, group i's belief about average income in another group j is

$$E_i^j = E^j + \beta\frac{n-1}{n}\frac{a}{2n}$$

if $i > j$ and

$$E_i^j = E^j - \beta\frac{n-1}{n}\frac{a}{2n}$$

if $i < j$.

If group 0 doesn't want to switch to group 2, it has to hold that

$$\hat{y}_1(E_0^2 - E^0) \le b_2$$

which translates to

$$\left(\frac{a}{n}\right)^2\left(2-\beta\frac{n-1}{2n}\right) \le 3\left(1-\beta\frac{n-1}{2n}\right)\left(\frac{a}{n}\right)^2 \tag{2.34}$$

and therefore

$$\beta\frac{n-1}{n} \le 1$$

which is always satisfied because $\beta \in (0,1)$. As only group distances matter in this case, condition (2.34) ensures that any group doesn't want to switch to a group that is two cutoffs above their own (i.e. if condition (2.34) is satisfied, then group 1 doesn't want to switch to group 3, group 2 doesn't want to switch to group 4 and so on).

If group 2 doesn't want to switch to group 0 it has to hold that

$$\hat{y}_2(E^2 - E_2^0) \ge b_2$$

which translates to

$$2\left(\frac{a}{n}\right)^2\left(2-\beta\frac{n-1}{2n}\right) \ge 3\left(1-\beta\frac{n-1}{2n}\right)\left(\frac{a}{n}\right)^2$$

and therefore

$$\beta\frac{n-1}{2n} \ge -1$$

which always holds. Again, this condition also ensures that group 3 doesn't want to switch to group 1, group 4 to group 2 etc. In a similar way we can check conditions for larger distances (group 0 to group 3, etc.). It turns out that these conditions are even slacker than the ones above, and therefore that all conditions of the form (2.33) are satisfied if F is uniform and the cutoffs are equidistant.

It is immediate to see that the proof works in the analogous way if people have proportional biased beliefs of type 2.

Proposition 2.17 *If people have proportional biased beliefs, and the income distribution is uniform, any equidistant partition is a biased sorting equilibrium with local consistency.*

General income distribution, local consistency

Concerning general income distributions, I will first analyze the case of three groups. I will assume that people have proportional biased beliefs of type 1 (Everything can be done analogously for proportional biased beliefs of type 2):

$$\begin{aligned}
E_0^0 &= E^0 \\
E_0^1 &= \beta(1-F(\hat{y}_1))\hat{y}_1 + (1-\beta(1-F(\hat{y}_1)))\,E^1 \\
E_0^2 &= \beta(1-F(\hat{y}_1))\hat{y}_2 + (1-\beta(1-F(\hat{y}_1)))\,E^2 \\
E_1^0 &= \beta(1-F(\hat{y}_2)+F(\hat{y}_1))\hat{y}_1 + (1-\beta(1-F(\hat{y}_2)+F(\hat{y}_1)))\,E^0 \\
E_1^1 &= E^1 \\
E_1^2 &= \beta(1-F(\hat{y}_2)+F(\hat{y}_1))\hat{y}_2 + (1-\beta(1-F(\hat{y}_2)+F(\hat{y}_1)))\,E^2 \\
E_2^0 &= \beta F(\hat{y}_2)\hat{y}_1 + (1-\beta F(\hat{y}_2))E^0 \\
E_2^1 &= \beta F(\hat{y}_2)\hat{y}_2 + (1-\beta F(\hat{y}_2))E^1 \\
E_2^2 &= E^2
\end{aligned}$$

A biased sorting equilibrium with local consistency and three groups needs to satisfy the following equilibrium conditions

$$E_0^1 - E_0^0 = E_1^1 - E_1^0 \tag{2.35}$$

$$E_1^2 - E_1^1 = E_2^2 - E_2^1 \tag{2.36}$$

and

$$\hat{y}_1(E_0^2 - E^0) \leq b_2 \tag{2.37}$$

$$\hat{y}_2(E^2 - E_2^0) \geq b_2 \tag{2.38}$$

Given the functional form of the misperceptions defined above, conditions (2.35) and (2.36) can be rewritten as

$$G(\hat{y}_1, \hat{y}_2) = \frac{(1-F_2+F_1)E^0 + (1-F_1)E^1}{2-F_2} - \hat{y}_1 = 0 \tag{2.39}$$

$$H(\hat{y}_1, \hat{y}_2) = \frac{F_2E^1 + (1-F_2+F_1)E^2}{1+F_1} - \hat{y}_2 = 0 \tag{2.40}$$

An equilibrium will exist if there exists a pair $(\hat{y}_1^*, \hat{y}_2^*)$ that satisfy both (2.39) and (2.40). My proof for existence proceeds in the following way: first I show that $(0,0)$ and $(E, y_{\max})$ satisfy (2.39), then I show that $(y_{\max}, y_{\max})$ and $(0, E)$ satisfy (2.40). This implies that, if G is such that $G(\hat{y}_1, \hat{y}_2) = 0$ implicitly describes a continuous function $\hat{y}_2(\hat{y}_1)$ and H is such that $H(\hat{y}_1, \hat{y}_2) = 0$ implicitly describes a continuous function $\hat{y}_1(\hat{y}_2)$, then the two must cross at some point in the $(\hat{y}_1, \hat{y}_2)$ space (and in fact this crossing must be where $\hat{y}_1 \leq E$ and $\hat{y}_2 \geq E$). I then invoke the implicit function theorem to show that these two continuous functions exist.
Step 1: $(0,0)$ and $(E, y_{\max})$ satisfy (2.39):
If I set $\hat{y}_1$ to 0 it follows that $F_1 = 0$ and $E^0 = 0$ and therefore

$$G(0, \hat{y}_2) = \frac{E^1}{1-F_2} - 0$$

and hence (2.39) is satisfied if $\hat{y}_2 = 0$ (which implies that $E^1 = 0$ and therefore $G(0,0) = 0$).
If I set $\hat{y}_2$ to $y_{\max}$ it follows that $F_2 = 1$ and therefore

$$G(\hat{y}_1, y_{\max}) = F_1E^0 + (1-F_1)E^1 - \hat{y}_1$$

and hence (2.39) is satisfied if $\hat{y}_1 = E$.
Step 2: $(y_{\max}, y_{\max})$ and $(0, E)$ satisfy (2.40):

If I set $\hat{y}_2$ to $y_{\max}$ then $F_2 = 1$ and $E^2 = y_{\max}$ and therefore

$$H(\hat{y}_1, y_{\max}) = \frac{E^1 + F_1 y_{\max}}{1 + F_1} - y_{\max}$$

and hence (2.40) is satisfied if $\hat{y}_1 = y_{\max}$ (because then $E^1 = y_{\max}$ and $F_1 = 1$ and hence $H(y_{\max}, y_{\max}) = \frac{2y_{\max}}{2} - y_{\max} = 0$).
If I set $\hat{y}_1$ to 0 it follows that $F_1 = 0$ and therefore

$$H(0, \hat{y}_2) = F_2 E^1 + (1 - F_2)E^2 - \hat{y}_2$$

and hence (2.40) is satisfied if $\hat{y}_2 = E$.
Step 3: $G(\hat{y}_1, \hat{y}_2) = 0$ describes a continuous function $\hat{y}_2(\hat{y}_1)$ for all $(\hat{y}_1, \hat{y}_2) \neq (0, 0)$ satisfying $G(\hat{y}_1, \hat{y}_2) = 0:$
Note that[15]

$$\begin{aligned}\frac{\partial G}{\partial \hat{y}_2} &= \frac{\left[-f_2 E^0 + (1 - F_1)\frac{\partial E^1}{\partial \hat{y}_2}\right](2 - F_2) + \left[(1 - F_2 + F_1)E^0 + (1 - F_1)E^1\right] f_2}{(2 - F_2)^2} \\ &= \frac{f_2}{(2 - F_2)^2}\left[(1 - F_1)(E^1 - E^0) + \frac{(1 - F_1)(2 - F_2)(\hat{y}_2 - E^1)}{F_2 - F_1}\right]\end{aligned}$$

By the implicit function theorem, $G(\hat{y}_1, \hat{y}_2) = 0$ describes a continuous function $\hat{y}_2(\hat{y}_1)$ for all $(\hat{y}_1, \hat{y}_2)$ satisfying $G(\hat{y}_1, \hat{y}_2) = 0$ such that

$$\frac{\partial G(\hat{y}_1, \hat{y}_2)}{\partial \hat{y}_2} \neq 0.$$

If $f_2 \neq 0$ (I assume that F is such that this holds whenever $\hat{y}_2 \neq 0$) then

$$\frac{\partial G}{\partial \hat{y}_2} \neq 0 \iff (1 - F_1)(E^1 - E^0) + \frac{(1 - F_1)(2 - F_2)(\hat{y}_2 - E^1)}{F_2 - F_1} \neq 0$$

In fact, it is easy to see that the latter inequality would not hold if

$$1 - F_1 = 0$$

but this cannot happen, because it would imply that $\hat{y}_1 = y_{\max}$ and $G(y_{\max}, \hat{y}_2) = 0$ is not satisfied for any $\hat{y}_2$. Furthermore, this expression is zero if

$$E^1 = E^0 \text{ and } \hat{y}_2 = E^1.$$

The only point at which this would be satisfied is where $(\hat{y}_1, \hat{y}_2) = (0, 0)$. On all points $(\hat{y}_1, \hat{y}_2) \neq (0, 0)$ for which $G(\hat{y}_1, \hat{y}_2) = 0$ it is actually the case that

$$\frac{\partial G(\hat{y}_1, \hat{y}_2)}{\partial \hat{y}_2} > 0$$

because $E^2 - E^1 > 0$ and $\hat{y}_2 - E^2 > 0$.
Step 4: $H(\hat{y}_1, \hat{y}_2) = 0$ describes a continuous function $\hat{y}_1(\hat{y}_2)$ for all points $(\hat{y}_1, \hat{y}_2) \neq (y_{\max}, y_{\max})$ and $(\hat{y}_1, \hat{y}_2) \neq (0, E)$ for which $H(\hat{y}_1, \hat{y}_2) = 0:$

$$\frac{\partial H}{\partial \hat{y}_1} = \frac{\left[f_1 E^2 + F_2 \frac{\partial E^1}{\partial \hat{y}_1}\right](1 + F_1) - F_2 E^1 + (1 - F_2 + F_1)E^2]f_1}{(1 + F_1)^2}$$

[15] Notation: $f_i = f(\hat{y}_i)$

$$= \frac{f_1}{(1+F_1)^2}\left[F_2(E^2 - E^1) + F_2\frac{(1+F_1)(E^1 - \hat{y}_1)}{F_2 - F_1}\right]$$

By the implicit function theorem, $H(\hat{y}_1, \hat{y}_2) = 0$ describes a continuous function $\hat{y}_1(\hat{y}_2)$ for all $(\hat{y}_1, \hat{y}_2)$ satisfying $H(\hat{y}_1, \hat{y}_2) = 0$ such that

$$\frac{\partial H(\hat{y}_1, \hat{y}_2)}{\partial \hat{y}_1} \neq 0$$

If $f_1 \neq 0$ (I assume that F is such that this holds whenever $\hat{y}_1 \neq 0$) then

$$\frac{\partial H}{\partial \hat{y}_1} \neq 0 \iff F_2(E^2 - E^1) + F_2\frac{(1+F_1)(E^1 - \hat{y}_1)}{F_2 - F_1} \neq 0$$

In fact, it is easy to see that the latter inequality would not hold if

$$F_2 = 0$$

which implies that in order for $\frac{\partial H}{\partial \hat{y}_1} \neq 0$ we need $\hat{y}_2 \neq 0$. However, note that $\hat{y}_2$ will never be 0 along $H(\hat{y}_1, \hat{y}_2) = 0$ because $H(\hat{y}_1, 0) = 0$ is never satisfied. Furthermore, the expression is zero if

$$E^2 = E^1 \text{ and } E^1 = \hat{y}_1.$$

The only point at which this would be satisfied is at $(y_{\max}, y_{\max})$. On all points $(\hat{y}_1, \hat{y}_2) \neq (y_{\max}, y_{\max})$ for which $H(\hat{y}_1, \hat{y}_2) = 0$ it is actually the case that

$$\frac{\partial H}{\partial \hat{y}_1} > 0$$

because $E^2 - E^1 > 0$ and $E^1 - \hat{y}_1 > 0$.

The above steps establish that there exists a pair $(\hat{y}_1, \hat{y}_2)$ that satisfy (2.39) and (2.40). Whether they also satisfy (2.37) and (2.38) and therefore constitute a biased sorting equilibrium with local consistency depends on β: We know that

$$b_2 = \hat{y}_2(E^2 - E^1_2) + b_1 = \hat{y}_2(E^2_1 - E^1) + b_1 \tag{2.41}$$

and

$$b_1 = \hat{y}_1(E^1 - E^0_1) = \hat{y}_1(E^1_0 - E^0)$$

Note that (2.37) and (2.41) can be combined to give

$$\hat{y}_1(E^2_0 - E^1_0) \leq \hat{y}_2(E^2_1 - E^1) \tag{2.42}$$

(rewrite $\hat{y}_1(E^2_0 - E^0) = \hat{y}_1(E^2_0 - E^1_0 + E^1_0 - E^0)$ and note that $\hat{y}_1(E^1_0 - E^0) = b_1$ and (2.36) can be written as $\hat{y}_2(E^2_1 - E^1_1) = \hat{y}_2(E^2_2 - E^1_2) = b_2 - b_1$). Denoting by $1 - x$ and $1 - w$ the size of group 0 and group 1 respectively, (2.42) can be rearranged to give

$$\hat{y}_1(E^2 - E^1) + \beta(x(E^1 - \hat{y}_1 - E^2 + \hat{y}_2)\hat{y}_1 + w(E^2 - \hat{y}_2)\hat{y}_2) \leq \hat{y}_2(E^2 - E^1).$$

As

$$\hat{y}_1(E^2 - E^1) < \hat{y}_2(E^2 - E^1),$$

condition (2.42) always holds if β is small enough. We can deal with (2.38) in the analogous way: Rewriting it, we get

$$\hat{y}_2(E^2 - E^1_2) + \hat{y}_2(E^1_2 - E^0_2) \geq b_2$$

and because (2.41) yields $\hat{y}_2(E^2 - E_2^1) = b_2 - b_1$ we get $\hat{y}_2(E_2^1 - E_2^0) \geq b_1$ which is equivalent to

$$\hat{y}_2(E_2^1 - E_2^0) \geq \hat{y}_1(E_0^1 - E^0). \tag{2.43}$$

Denoting by $1 - x$ and $1 - z$ the size of group 0 and 2 respectively, this condition can be rearranged to give

$$\hat{y}_2(E^1 - E^0) + \beta(z(\hat{y}_2 - E^1 - \hat{y}_1 + E^0)\hat{y}_2 + x(E^1 - \hat{y}_1)\hat{y}_1) \geq \hat{y}_1(E^1 - E^0).$$

Again, due to

$$\hat{y}_2(E^1 - E^0) > \hat{y}_1(E^1 - E^0)$$

condition (2.43) always holds for small enough β. Hence, conditions (2.37) and (2.37) will always be satisfied for small β. It follows that for small enough β there will always exist a pair $(\hat{y}_1, \hat{y}_2)$ that satisfy all four conditions (2.39) - (2.38).

Proposition 2.18 *With Proportional Biased Beliefs, a biased sorting equilibrium with local consistency and three groups will always exist for small enough* β*.*

In fact, it can be seen that the same pattern emerges for four groups:
With four groups, the conditions for equilibrium and local consistency can be reduced to

$$E_0^1 - E^0 = E^1 - E_1^0 \tag{2.44}$$
$$E_1^2 - E^1 = E^2 - E_2^1 \tag{2.45}$$
$$E_2^3 - E^2 = E^3 - E_3^2 \tag{2.46}$$

and

$$\begin{aligned}
\hat{y}_1(E_0^2 - E^0) &\leq b_2 \\
\hat{y}_2(E^2 - E_2^0) &\geq b_2 \\
\hat{y}_1(E_0^3 - E^0) &\leq b_3 \\
\hat{y}_3(E^3 - E_3^0) &\geq b_3 \\
\hat{y}_2(E_1^3 - E^1) &\leq b_3 - b_1 \\
\hat{y}_3(E^3 - E_3^1) &\geq b_3 - b_1
\end{aligned}$$

The first three conditions yield

$$\hat{y}_1 = \frac{(1 - F_2 + F_1)E^0 + (1 - F_1)E^1}{2 - F_2}$$

$$\hat{y}_2 = \frac{(1 - F_3 + F_2)E^1 + (1 - F_2 + F_1)E^2}{2 + F_1 - F_3}$$

$$\hat{y}_3 = \frac{(1 - F_3 + F_2)E^3 + (F_3)E^2}{1 + F_2}$$

Again, the equilibrium cutoffs are weighted averages of the average incomes of the two neighbouring groups and such cutoffs can always be found. The additional six conditions (which ensure that people don't want to switch to another non-neighbouring group) are again satisfied if β is small enough (the proof proceeds in the same way as for the three-group case above). This pattern of equilibrium cutoffs that are weighted averages of neighbouring groups' average incomes emerges for every finite number of groups. Thus, if only local consistency is required, existence of multi-group biased sorting equilibria is guaranteed for small β. As I demonstrate above, if the income distribution is uniform, any equidistant partition is a sorting equilibrium, irrespective of the size of β.

General income distribution, global consistency

In Section 2.6, I have pointed out that for the existence of biased sorting equilibria with global consistency and more than two groups, it is crucial how exactly the misperceptions are modelled and I have demonstrated that equilibria with more than two groups do not exist under the assumption that perceived differences between groups decline with group distance. In the following, I will show that if people have proportional biased beliefs there cannot exist more than three groups in equilibrium.

Suppose there are four groups: A poor group with income from 0 to $\hat{y}_1$, a lower middle-class group with incomes from $\hat{y}_1$ to $\hat{y}_2$, an upper middle-class group with incomes from $\hat{y}_2$ to $\hat{y}_3$ and a rich group with incomes from $\hat{y}_3$ to $y_{\max}$. As always, I assume that people perceive the average income in their own group correctly, but they are biased about average income in the other groups. Suppose that people have proportional biased beliefs of type 1:[16]

$$
\begin{aligned}
E_0^0 &= E^0\\
E_0^1 &= \beta(1-F(\hat{y}_1))\hat{y}_1 + (1-\beta(1-F(\hat{y}_1)))\,E^1\\
E_0^2 &= \beta(1-F(\hat{y}_1))\hat{y}_2 + (1-\beta(1-F(\hat{y}_1)))\,E^2\\
E_0^3 &= \beta(1-F(\hat{y}_1))\hat{y}_3 + (1-\beta(1-F(\hat{y}_1)))\,E^3\\
E_1^0 &= \beta(1-F(\hat{y}_2)+F(\hat{y}_1))\hat{y}_1 + (1-\beta(1-F(\hat{y}_2)+F(\hat{y}_1)))\,E^0\\
E_1^1 &= E^1\\
E_1^2 &= \beta(1-F(\hat{y}_2)+F(\hat{y}_1))\hat{y}_2 + (1-\beta(1-F(\hat{y}_2)+F(\hat{y}_1)))\,E^2\\
E_1^3 &= \beta(1-F(\hat{y}_2)+F(\hat{y}_1))\hat{y}_3 + (1-\beta(1-F(\hat{y}_2)+F(\hat{y}_1)))\,E^3\\
E_2^0 &= \beta(1-F(\hat{y}_3)+F(\hat{y}_2))\hat{y}_1 + (1-\beta(1-F(\hat{y}_3)+F(\hat{y}_2)))E^0\\
E_2^1 &= \beta(1-F(\hat{y}_3)+F(\hat{y}_2))\hat{y}_2 + (1-\beta(1-F(\hat{y}_3)+F(\hat{y}_2)))E^1\\
E_2^2 &= E^2\\
E_2^3 &= \beta(1-F(\hat{y}_3)+F(\hat{y}_2))\hat{y}_3 + (1-\beta(1-F(\hat{y}_3)+F(\hat{y}_2)))E^3\\
E_3^0 &= \beta F(\hat{y}_3)\hat{y}_1 + (1-\beta F(\hat{y}_3))E^0\\
E_3^1 &= \beta F(\hat{y}_3)\hat{y}_2 + (1-\beta F(\hat{y}_3))E^1\\
E_3^2 &= \beta F(\hat{y}_3)\hat{y}_3 + (1-\beta F(\hat{y}_3))E^2\\
E_3^3 &= E^3
\end{aligned}
$$

Then in order for this partition to constitute a biased sorting equilibrium with global consistency, the following conditions have to hold:

For a biased sorting equilibrium, I need:

[16] The analysis for proportional biased beliefs of type 2 can be done in the analogous way.

$$
\begin{aligned}
y(E_0^1 - E_0^0) &\leq b_1 \quad \forall y \leq \hat{y}_1 \\
y(E_0^2 - E_0^0) &\leq b_2 \quad \forall y \leq \hat{y}_1 \\
y(E_0^2 - E_0^0) &\leq b_3 \quad \forall y \leq \hat{y}_1 \\
y(E_1^1 - E_1^0) &\geq b_1 \quad \forall y \in [\hat{y}_1, \hat{y}_2] \\
y(E_1^2 - E_1^1) &\leq b_2 - b_1 \quad \forall y \in [\hat{y}_1, \hat{y}_2] \\
y(E_1^3 - E_1^1) &\leq b_3 - b_1 \quad \forall y \in [\hat{y}_1, \hat{y}_2] \\
y(E_2^2 - E_2^0) &\geq b_2 \quad \forall y \in [\hat{y}_2, \hat{y}_3] \\
y(E_2^2 - E_2^1) &\geq b_2 - b_1 \quad \forall y \in [\hat{y}_2, \hat{y}_3] \\
y(E_2^3 - E_2^2) &\leq b_3 - b_2 \quad \forall y \in [\hat{y}_2, \hat{y}_3] \\
y(E_3^3 - E_3^0) &\geq b_3 \quad \forall y \geq \hat{y}_3 \\
y(E_3^3 - E_3^1) &\geq b_3 - b_1 \quad \forall y \geq \hat{y}_3 \\
y(E_3^3 - E_3^2) &\geq b_3 - b_2 \quad \forall y \geq \hat{y}_3
\end{aligned}
$$

For global consistency, the additional conditions are:

$$
\begin{aligned}
y(E_0^1 - E_0^0) &\geq b_1 \quad \forall y \in [\hat{y}_1, \hat{y}_2] \\
y(E_0^2 - E_0^1) &\leq b_2 - b_1 \quad \forall y \in [\hat{y}_1, \hat{y}_2] \\
y(E_0^3 - E_0^1) &\leq b_3 - b_1 \quad \forall y \in [\hat{y}_1, \hat{y}_2] \\
y(E_0^2 - E_0^1) &\geq b_2 - b_1 \quad \forall y \in [\hat{y}_2, \hat{y}_3] \\
y(E_0^2 - E_0^0) &\geq b_2 \quad \forall y \in [\hat{y}_2, \hat{y}_3] \\
y(E_0^3 - E_0^2) &\leq b_3 - b_2 \quad \forall y \in [\hat{y}_2, \hat{y}_3] \\
y(E_0^3 - E_0^0) &\geq b_3 \quad \forall y \geq \hat{y}_3 \\
y(E_0^3 - E_0^2) &\geq b_3 - b_2 \quad \forall y \geq \hat{y}_3 \\
y(E_0^3 - E_0^1) &\geq b_3 - b_1 \quad \forall y \geq \hat{y}_3
\end{aligned}
$$

$$
\begin{aligned}
y(E_1^1 - E_1^0) &\leq b_1 \quad \forall y \leq \hat{y}_1 \\
y(E_1^2 - E_1^0) &\leq b_2 \quad \forall y \leq \hat{y}_1 \\
y(E_1^3 - E_1^0) &\leq b_3 \quad \forall y \leq \hat{y}_1 \\
y(E_1^2 - E_1^1) &\geq b_2 - b_1 \quad \forall y \in [\hat{y}_2, \hat{y}_3] \\
y(E_1^2 - E_1^0) &\geq b_2 \quad \forall y \in [\hat{y}_2, \hat{y}_3] \\
y(E_1^3 - E_1^2) &\leq b_3 - b_2 \quad \forall y \in [\hat{y}_2, \hat{y}_3] \\
y(E_1^3 - E_1^0) &\geq b_3 \quad \forall y \geq \hat{y}_3 \\
y(E_1^3 - E_1^2) &\geq b_3 - b_2 \quad \forall y \geq \hat{y}_3 \\
y(E_1^3 - E_1^1) &\geq b_3 - b_1 \quad \forall y \geq \hat{y}_3
\end{aligned}
$$

$$
\begin{aligned}
y(E_2^1 - E_2^0) &\leq b_1 \quad \forall y \leq \hat{y}_1 \\
y(E_2^2 - E_2^0) &\leq b_2 \quad \forall y \leq \hat{y}_1 \\
y(E_2^3 - E_2^1) &\leq b_3 \quad \forall y \leq \hat{y}_1 \\
y(E_2^1 - E_2^0) &\geq b_1 \quad \forall y \in [\hat{y}_1, \hat{y}_2] \\
y(E_2^2 - E_2^1) &\leq b_2 - b_1 \quad \forall y \in [\hat{y}_1, \hat{y}_2] \\
y(E_2^3 - E_2^1) &\leq b_3 - b_1 \quad \forall y \in [\hat{y}_1, \hat{y}_2] \\
y(E_2^3 - E_2^0) &\geq b_3 \quad \forall y \geq \hat{y}_3 \\
y(E_2^3 - E_2^2) &\geq b_3 - b_2 \quad \forall y \geq \hat{y}_3 \\
y(E_2^3 - E_2^1) &\geq b_3 - b_1 \quad \forall y \geq \hat{y}_3
\end{aligned}
$$

$$
\begin{aligned}
y(E_3^1 - E_3^0) &\leq b_1 \quad \forall y \leq \hat{y}_1 \\
y(E_3^2 - E_3^0) &\leq b_2 \quad \forall y \leq \hat{y}_1 \\
y(E_3^3 - E_3^0) &\leq b_3 \quad \forall y \leq \hat{y}_1 \\
y(E_3^1 - E_3^0) &\geq b_1 \quad \forall y \in [\hat{y}_1, \hat{y}_2] \\
y(E_3^2 - E_3^0) &\leq b_2 - b_1 \quad \forall y \in [\hat{y}_1, \hat{y}_2] \\
y(E_3^3 - E_3^0) &\leq b_3 - b_1 \quad \forall y \in [\hat{y}_1, \hat{y}_2] \\
y(E_3^2 - E_3^1) &\geq b_2 - b_1 \quad \forall y \in [\hat{y}_2, \hat{y}_3] \\
y(E_3^2 - E_3^0) &\geq b_2 \quad \forall y \in [\hat{y}_2, \hat{y}_3] \\
y(E_3^3 - E_3^2) &\leq b_3 - b_2 \quad \forall y \in [\hat{y}_2, \hat{y}_3]
\end{aligned}
$$

These conditions can be combined to give

$$
E_0^1 - E_0^0 = E_1^1 - E_1^0 = E_2^1 - E_2^0 = E_3^1 - E_3^0 \tag{2.47}
$$

$$
E_0^2 - E_0^1 = E_1^2 - E_1^1 = E_2^2 - E_2^1 = E_3^2 - E_3^1 \tag{2.48}
$$

$$
E_0^3 - E_0^2 = E_1^3 - E_1^2 = E_2^3 - E_2^2 = E_3^3 - E_3^2 \tag{2.49}
$$

For simplicity of notation, let me denote

$$
x = 1 - F(\hat{y}_1), \quad m = 1 - F(\hat{y}_2) + F(\hat{y}_1), \quad z = 1 - F(\hat{y}_3) + F(\hat{y}_2), \quad w = F(\hat{y}_3)
$$

Then (2.49) implies that

$$
E_2^3 - E_2^2 = E_0^3 - E_0^2 \iff
$$

$$
\beta z\hat{y}_3 + (1 - \beta z)E^3 - E^2 = \beta x\hat{y}_3 + (1 - \beta x)E^3 - \beta x\hat{y}_2 - (1 - \beta x)E^2
$$

$$
\iff \hat{y}_2 = E^2 + \frac{(z - x)}{x}(E^3 - \hat{y}_3) \tag{2.50}
$$

(2.49) also implies that

$$
E_2^3 - E_2^2 = E_1^3 - E_1^2 \iff
$$

$$
\beta z\hat{y}_3 + (1 - \beta z)E^3 - E^2 = \beta m\hat{y}_3 + (1 - \beta m)E^3 - \beta m\hat{y}_2 - (1 - \beta m)E^2
$$

$$
\iff \hat{y}_2 = E^2 + \frac{(z - m)}{m}(E^3 - \hat{y}_3) \tag{2.51}
$$

There are 2 conclusions that follow from (2.50) and (2.51): First, it has to be the case that $z < x$ and $z < m$ (implying that group 3 needs to be larger than both group 1 and group 2). Second, combining the two equations we find that $m = x$.
(2.48) implies that

$$E_0^2 - E_0^1 = E_1^2 - E_1^1 \iff$$

$$\beta m\hat{y}_2 + (1-\beta m)E^2 - E^1 = \beta x\hat{y}_2 + (1-\beta x)E^2 - \beta x\hat{y}_1 - (1-\beta x)E^1$$

$$\iff \hat{y}_1 = \left(\frac{m-x}{x}\right)(E^2 - \hat{y}_2) + E^1$$

which implies that $m < x$ because $\hat{y}_1 < E^1$. Hence, (2.48) and (2.49) cannot be satisfied at the same time, and a biased sorting equilibrium with consistency cannot exist for four groups. In fact, this argument holds also for more than four groups, because nowhere in this proof did I make the assumption that there are only four groups. The conditions that contradict each other would be the same with $n \geq 4$ groups. The analogous analysis can be conducted for proportional biased beliefs of type 2 to yield the exact same result: no biased sorting equilibria with global consistency will exist for more than three groups.

The possibility of 3 groups

Suppose there are 3 groups: A poor group with income from 0 to $\hat{y}_1$, a middle group with incomes from $\hat{y}_1$ to $\hat{y}_2$ and a rich group with incomes from $\hat{y}_2$ to $y_{\max}$. Suppose that the belief function is of the same type as in the previous section, and hence

$$\begin{aligned}
E_0^0 &= E^0 \\
E_0^1 &= \beta(1-F(\hat{y}_1))\hat{y}_1 + (1-\beta(1-F(\hat{y}_1)))\,E^1 \\
E_0^2 &= \beta(1-F(\hat{y}_1))\hat{y}_2 + (1-\beta(1-F(\hat{y}_1)))\,E^2 \\
E_1^0 &= \beta(1-F(\hat{y}_2)+F(\hat{y}_1))\hat{y}_1 + (1-\beta(1-F(\hat{y}_2)+F(\hat{y}_1)))\,E^0 \\
E_1^1 &= E^1 \\
E_1^2 &= \beta(1-F(\hat{y}_2)+F(\hat{y}_1))\hat{y}_2 + (1-\beta(1-F(\hat{y}_2)+F(\hat{y}_1)))\,E^2 \\
E_2^0 &= \beta(1-F(\hat{y}_3)+F(\hat{y}_2))\hat{y}_1 + (1-\beta(1-F(\hat{y}_3)+F(\hat{y}_2)))E^0 \\
E_2^1 &= \beta(1-F(\hat{y}_3)+F(\hat{y}_2))\hat{y}_2 + (1-\beta(1-F(\hat{y}_3)+F(\hat{y}_2)))E^1 \\
E_2^2 &= E^2
\end{aligned}$$

Then in order for this partition to constitute a biased sorting equilibrium with global consistency, the following conditions have to hold:
For a biased sorting equilibrium, I need:

$$\begin{aligned}
y(E_0^1 - E_0^0) &\leq b_1 \quad \forall y \leq \hat{y}_1 \\
y(E_0^2 - E_0^0) &\leq b_2 \quad \forall y \leq \hat{y}_1 \\
y(E_1^1 - E_1^0) &\geq b_1 \quad \forall y \in [\hat{y}_1, \hat{y}_2] \\
y(E_1^2 - E_1^1) &\leq b_2 - b_1 \quad \forall y \in [\hat{y}_1, \hat{y}_2] \\
y(E_2^2 - E_2^0) &\geq b_2 \quad \forall y \geq \hat{y}_2 \\
y(E_2^2 - E_2^1) &\geq b_2 - b_1 \quad \forall y \geq \hat{y}_2
\end{aligned}$$

For global consistency, the additional conditions are:

$$
\begin{aligned}
y(E_0^1 - E_0^0) &\geq b_1 \quad \forall y \in [\hat{y}_1, \hat{y}_2] \\
y(E_0^2 - E_0^1) &\leq b_2 - b_1 \quad \forall y \in [\hat{y}_1, \hat{y}_2] \\
y(E_0^2 - E_0^1) &\geq b_2 - b_1 \quad \forall y \geq \hat{y}_2 \\
y(E_0^2 - E_0^0) &\geq b_2 \quad \forall y \geq \hat{y}_2 \\
jky(E_1^1 - E_1^0) &\leq b_1 \quad \forall y \leq \hat{y}_1 \\
y(E_1^2 - E_1^0) &\leq b_2 \quad \forall y \leq \hat{y}_1 \\
y(E_1^2 - E_1^1) &\geq b_2 - b_1 \quad \forall y \geq \hat{y}_2 \\
y(E_1^2 - E_1^0) &\geq b_2 \quad \forall y \geq \hat{y}_2 \\
y(E_2^1 - E_2^0) &\leq b_1 \quad \forall y \leq \hat{y}_1 \\
y(E_2^2 - E_2^0) &\leq b_2 \quad \forall y \leq \hat{y}_1 \\
y(E_2^2 - E_2^1) &\leq b_2 - b_1 \quad \forall y \in [\hat{y}_1, \hat{y}_2] \\
y(E_2^1 - E_2^0) &\geq b_1 \quad \forall y \in [\hat{y}_1, \hat{y}_2]
\end{aligned}
$$

These conditions can be combined to

$$E_0^1 - E_0^0 = E_1^1 - E_1^0 = E_2^1 - E_2^0 \tag{2.52}$$

$$E_0^2 - E_0^1 = E_1^2 - E_1^1 = E_2^2 - E_2^1 \tag{2.53}$$

Hence, $\hat{y}_1$ and $\hat{y}_2$ must be such that (2.52) and (2.53) are satisfied.
For simplicity of notation, let me denote

$$x = 1 - F(\hat{y}_1), \quad m = 1 - F(\hat{y}_2) + F(\hat{y}_1), \quad z = F(\hat{y}_2)$$

Then (2.52) implies

$$E_1^1 - E_1^0 = E_2^1 - E_2^0 \iff$$

$$E^1 - \beta m\hat{y}_1 - (1 - \beta m)E^0 = \beta z\hat{y}_2 + (1 - \beta z)E^1 - \beta z\hat{y}_1 - (1 - \beta z)E^0$$

$$\iff zE^1 + (z - m)(\hat{y}_1 - E^0) = z\hat{y}_2$$

which implies that $z > m$ because $\hat{y}_2 > E^1$.
Furthermore, from (2.53) I get

$$E_0^2 - E_0^1 = E_1^2 - E_1^1 \iff$$

$$\beta x\hat{y}_2 + (1 - \beta x)E^2 - \beta x\hat{y}_1 - (1 - \beta x)E^1 = \beta m\hat{y}_2 + (1 - \beta m)E^2 - E^1$$

$$\iff \frac{m - x}{x}\left(E^2 - \hat{y}_2\right) + E^1 = \hat{y}_1$$

and this implies $m < x$ because $\hat{y}_1 < E^1$.
Moreover,

$$E_0^1 - E_0^0 = E_1^1 - E_1^0$$

gives

$$\hat{y}_1 = \frac{xE^1 + mE^0}{x + m}.$$

Also,

$$E_1^2 - E_1^1 = E_2^2 - E_2^1$$

gives

$$\hat{y}_2 = \frac{zE^1 + mE^2}{z+m}.$$

All together, these are 4 equations in 2 unknowns.

$$zE^1 + (z-m)(\hat{y}_1 - E^0) = z\hat{y}_2 \tag{2.54}$$

$$\frac{m-x}{x}\left(E^2 - \hat{y}_2\right) + E^1 = \hat{y}_1 \tag{2.55}$$

$$\hat{y}_1 = \frac{xE^1 + mE^0}{x+m} \tag{2.56}$$

$$\hat{y}_2 = \frac{zE^1 + mE^2}{z+m} \tag{2.57}$$

Will such an equilibrium exist and be unique? Note that (2.56) and (2.57) are the same conditions as the equilibrium conditions (2.39) and (2.40) for the equilibrium with local consistency above. As I have argued, a pair $(\hat{y}_1^*, \hat{y}_2^*)$ satisfying these two conditions can exist (e.g. if the income distribution is uniform). It depends on $F(.)$ whether a pair $(\hat{y}_1^*, \hat{y}_2^*)$ satisfying all four of these equations exists. For instance, I have demonstrated above that such a pair $(\hat{y}_1^*, \hat{y}_2^*)$ cannot be found for the uniform distribution.

To conclude, the analysis in this section shows that biased sorting equilibria with local consistency can, depending on the shape of the income distribution, exist for every (finite) number of groups. Whether biased sorting equilibria with global consistency with more than two groups can exist depends crucially on how misperceptions about non-neighbouring groups are specified. If they take the form of proportional biased beliefs, then equilibria with more than three groups can be ruled out. If the perceived difference in average incomes between groups is decreasing in group distance, equilibria with more than two groups can be ruled out.

Chapter 3

Monopolistic Supply of Sorting, Inequality and Welfare

3.1 Introduction

In recent years, we have observed a rise in social segregation in many industrialized countries. People tend to interact increasingly with others who are not too different from themselves in terms of income, education and political beliefs.[1] Moreover, evidence suggests that segregation and income inequality tend to move jointly. Several studies for the US show that both income inequality and segregation have increased in most metropolitan areas over the past 40 years.[2] The reasons for this co-movement haven't been explored widely so far. While the presence of assortative matching and (positive) sorting has been extensively discussed in the economics and sociology literature, little research has been done so far on the supply side of segregation and the relationship between inequality and the supply of segregation.

Given the trend of mounting social segregation, an important question is also the social desirability of sorting. If people benefit from interacting with wealthy and influential people, poor people who are deprived of these contacts due to social seclusion will suffer. But sorting might not be universally beneficial for the rich either: Especially if inequality is high, it might be the case that they have to pay huge sums to separate themselves off from the rest of society (e.g. via gated communities or private schools). While Becker (1974) shows that assortative matching always maximizes total surplus in society, Levy and Razin (2015) and Hoppe et al. (2009) demonstrate that segregation is not necessarily beneficial for welfare if we count these "sorting fees" as deadweight loss and subtract them from the surplus.

Finally, it is important to note that the interests of a supplier of the sorting technology might be different from society's interests, and that the way sorting is implemented need not be optimal for society. In addition, an increase in inequality is likely to have different effects on the supplier of the sorting technology and on welfare.

In the present paper, I make a first attempt to analyze the relationship between income inequality and the supply of sorting and to examine how well the interests of the supplier of the sorting technology and of society as a whole are aligned, especially in the face of rising inequality.

In my analysis, I deploy a simple model in which income is distributed unequally in society and people can pay a "fee" to join a group and interact only with members of that group henceforth. I examine how this fee will be set if a profit-maximizing monopolist offers this

[1] See e.g. Forman and Koch (2013) and Bishop (2008) for evidence on the US.

[2] See e.g. Reardon and Bischoff (2011) and Watson (2009).

sorting technology, and I analyze the monopolist's profits and society's total welfare resulting from this split into groups. I show that an increase in inequality increases monopolist profits from offering people the possibility to segregate, and potentially also welfare from segregation. However, I demonstrate that there is often a conflict between welfare and monopolist profits, in the sense that different partitions of society would be optimal for profits and welfare - the way in which the monopolist splits up society is in general not efficient (i.e. welfare maximizing). This conflict tends to intensify as inequality increases: monopolist profits increase, while welfare from sorting decreases as income inequality climbs high. At the end of the paper I argue that there is a sense in which this finding holds also if we allow the monopolist to offer more than just one group.

The rest of the paper is organized as follows: Section 2 presents related literature, Section 3 introduces the model of sorting according to income and examines how changes in inequality affect monopolist profits and welfare. Section 4 uses a stylized income distribution (the symmetric atom distribution) to demonstrate that there can be a conflict between monopolist profits and welfare as inequality increases, and generalizes this result to other types of income distributions. Section 5 examines the effect of increasing inequality on monopolist profits and welfare if the monopolist can offer as many cutoffs as she wants and Section 6 concludes.

3.2 Related Literature

The standard model of sorting and assortative matching is outlined and analyzed in Becker (1974). Levy and Razin (2015) examine total welfare and preferences for redistribution in the presence of costly income sorting without explicitly modelling the supply side of the sorting technology. Rayo (2013) characterizes optimal sorting if a profit-maximizing monopolist without costs chooses the sorting schedule, while Damiano and Li (2007) analyze the case of two or more competing firms. My paper carries elements of both Levy and Razin (2015) (in the sense that I analyze the normative aspects of segregation, in particular its effects on welfare) and of Rayo (2013) (because I assume that the sorting technology is offered by a profit-maximizing monopolist). The main contribution of my paper is that I examine how optimal sorting varies with inequality and how this affects the (potential) conflict between welfare and monopolist profit.

My paper is also related to the literature of costly signalling (see e.g. Hoppe, Moldovanu and Sela (2009)) and conspicuous consumption (see e.g. Pesendorfer (1995), Bagwell and Bernheim (1996) and Veblen (1899)) and to the literature of educational segregation via private schools (see e.g. Fernandez and Rogerson (2003), Epple and Romano (1998) and Levy and Razin (2016)).

3.3 Inequality, monopolist profit and welfare

Let income y in an economy be distributed according to an income distribution $F(y)$, on the interval $Y = [0, y_{\max}]$ (where $y_{\max} < \infty$ unless explicitly mentioned otherwise). Assume furthermore that $F(y)$ is continuous and strictly monotonic. Suppose that an agent's utility is increasing not only in her own income but also in the average income of the people that she interacts with, which I will henceforth call her "reference group". Specifically, a person with income y_j gets utility $U_j = y_j E(y|y \in S_i)$, where S_i is individual j's reference group. If there is no economic segregation, everybody's reference group is a representative sample of the whole population, such that $U_j = y_j E(y)$. However, a person with income y_j can pay a fee $b > 0$ to join group S_b and get utility

$$y_j E[y|y \in S_b] - b$$

or refrain from paying b and get

$$y_j E[y|y \in S_0]$$

where S_b is the set of incomes y of people who have paid b and S_0 is the set of incomes y of people who haven't paid b. Then I can define the following:

Definition 3.1 *A sorting equilibrium is a partition $[S_0, S_b]$ of Y and a sorting fee $b > 0$ such that*

$$yE[y|y \in S_b] - b \leq yE[y|y \in S_0] \quad \forall y \in S_0 \tag{3.1}$$

$$yE[y|y \in S_b] - b \geq yE[y|y \in S_0] \quad \forall y \in S_b \tag{3.2}$$

In a sorting equilibrium as defined above people stay in the group that gives them the highest utility.

In Windsteiger (2017b), I discuss this model in detail and show that in any sorting equilibrium, group S_b must have a higher average income than group S_0, and that all sorting equilibria will be *monotone,* meaning that the groups S_0 and S_b are single intervals of Y (where group S_b must lie to the right of group S_0 on the Y scale).

Therefore, I will from now on call people in S_b "the rich" and people in S_0 "the poor". Furthermore, the fact that all equilibria are monotone allows me to rewrite the definition of a sorting equilibrium in terms of a cutoff $\hat{y}$, where everybody with income below the cutoff is in the poor group and everybody with income above the cutoff is in the rich group. For simplicity of notation I will denote average income in the rich group, $E[y|y \in S_b]$, by $\bar{E}(\hat{y})$ and average income in the poor group, $E[y|y \in S_0]$, by $\underline{E}(\hat{y})$. In Windsteiger (2017b) I show the following:

Corollary 3.1 *A sorting equilibrium is characterized by a cutoff $\hat{y} \in Y$ and a sorting fee b such that*

$$\hat{y}\bar{E}(\hat{y}) - \hat{y}\underline{E}(\hat{y}) = b \tag{3.3}$$

A person with income $\hat{y}$ just at the border of the two groups S_b and S_0 has to be exactly indifferent between joining either of the two groups in equilibrium. For the remainder of the paper I will choose the convention that people with income $\hat{y}$ (who are indifferent between the two groups) stay in the poor group.

It can immediately be seen from (3.3) that the sorting fee is uniquely determined by the equilibrium cutoff $\hat{y}$, i.e. for a given equilibrium partition $\{[0, \hat{y}], (\hat{y}, y_{\max}]\}$, the sorting fee b is unique. The reverse statement is not true in general: For a given b, there might be multiple cutoffs $\hat{y}$ that satisfy $\hat{y}(\bar{E}(\hat{y}) - \underline{E}(\hat{y})) = b$ (this could happen if the distribution is such that $\hat{y}(\bar{E}(\hat{y}) - \underline{E}(\hat{y}))$ is not strictly increasing or decreasing for all $\hat{y} \in Y$[3]). For a given sorting fee, there could therefore be several monotone partitions of society that would be sorting equilibria given this fee. When I model the supply side below, I thus require that whoever offers the sorting technology chooses the cutoff optimally and I implicitly assume that the supplier can then ensure that the agents coordinate on the equilibrium that yields the highest payoff for the supplier (which, in the case of a profit-maximizing firm, would always be the lowest cutoff $\hat{y}$ such that $\hat{y}(\bar{E}(\hat{y}) - \underline{E}(\hat{y})) = b$, because it yields the largest mass of customers).

[3]It can be shown that a sufficient condition for $\hat{y}(\bar{E} - \underline{E})$ to be monotone is that the income distribution is *new worse than used in expectations (NWUE).* For a definition of the NWUE property see Section 3.3.2.

3.3.1 Monopolist profit

The model outlined above shows how the sorting fee has to be set in order to generate a certain partition of society. But who determines how the groups in society look like? Who offers the sorting technology and chooses the cutoff?

For the remainder of this paper I will assume that the sorting technology is offered by a profit-maximizing monopolist and I will examine the implications of an increase in inequality for the monopolist's profits and for total welfare. In the next sections I will focus on the model of sorting with two groups as described above. The monopolist can therefore only decide between offering one cutoff or staying inactive, but she cannot offer more than one cutoff. This could be modelled explicitly by assuming that the costs of offering more than one cutoff are prohibitively high. In the last section of this paper, I will discuss what happens if the monopolist's costs are negligible and she can therefore offer as many cutoffs as she wants.

If the monopolist faces fixed costs $c > 0$ of operating, her profits from offering sorting are

$$\Pi(\hat{y}^*) = R(\hat{y}^*) - c,$$

where $R(\hat{y}^*)$ is the revenue from offering sorting at cutoff $\hat{y}^*$ and $\hat{y}^*$ is placed optimally,

$$\hat{y}^* = \arg\max_{\hat{y}} R(\hat{y}).$$

Revenue at cutoff $\hat{y}$ is given by

$$R(\hat{y}) = \hat{y}(\bar{E}(\hat{y}) - \underline{E}(\hat{y}))(1 - F(\hat{y})) = \hat{y}(E - \underline{E}(\hat{y})).$$

It is straightforward to see that the solution to the revenue maximization problem must be interior, because $R(0) = R(y_{\max}) = 0$ whereas $R(\hat{y})$ is strictly positive for any interior $\hat{y}$.

Suppose that the income distribution and the fixed costs c are such that $\Pi(\hat{y}^*) > 0$ and hence it is profitable for the monopolist to offer the sorting technology. What happens to her profits as inequality increases? In the following, I will show that the monopolist's profits always rise if inequality increases in the form of a particular type of mean-preserving spread of the income distribution. I shall say that a mean-preserving spread is *monotone* if $\bar{E}(\hat{y})$ increases and $\underline{E}(\hat{y})$ decreases for any interior cutoff $\hat{y}$ (while of course, as implied by the definition of a mean-preserving spread, average income E doesn't change.)

Proposition 3.1 *A monotone mean-preserving spread of the income distribution increases the monopolist's profits from offering sorting.*

Proof. If inequality increases in the form of a monotone mean-preserving spread of the income distribution, the difference $E - \underline{E}$ will increase. This implies a rise in $\hat{y}^*(E - \underline{E})$, keeping $\hat{y}^*$ constant at the optimal choice for the initial income distribution. It is very likely that the optimal cutoff will also change for the monopolist, but even with keeping the old cutoff, her revenues increase, and they will do even more so if the monopolist also chooses the cutoff optimally. ■

Remark 3.1 *A mean-preserving spread of the income distribution always implies an increase in the Gini-coefficient (see Dalton (1920) and Cowell (2000)).*

Remark 3.2 *In order for the monopolist's profits to increase, the mean-preserving spread does not have to be such that $\bar{E}$ increases and $\underline{E}$ decreases for any cutoff - it suffices if this holds for the initially optimal cutoff. The proposition therefore states sufficient conditions for an increase in the monopolist's profits.*

Note that the definition of a general mean-preserving spread of a distribution requires that mass from the middle of the distribution is transferred to the tails in such a way that the mean of the distribution remains constant (see Rothschild and Stiglitz (1970) or Atkinson (1970)). Formally, we say that $G(y)$ is a mean-preserving spread of $F(y)$ if (1) $\int dG(y) = \int dF(y)$ and (2) $\int_0^{\hat{y}} [F(y) - G(y)]dy \leq 0 \ \forall \hat{y} \in Y$ with strict inequality for some $\hat{y}$. It is immediate to see that this definition doesn't imply that $\bar{E}$ increases and $\underline{E}$ decreases for *all* cutoffs. For instance, suppose we take mass from the interval $[a, b]$ (where $0 < a < b < E$) and transfer it to the interval $[a', b']$ (where $a' < a$ and $b' < b$) and do a symmetric shift of mass to the upper tail from an interval above the mean such that the mean stays constant. This transformation would qualify as a mean-preserving spread, but the conditional expectations at any cutoff below a wouldn't change (or in other words, $\int_0^{\hat{y}} [F(y) - G(y)]dy = 0 \ \forall \hat{y} < a$). We can ensure that the mean-preserving spread increases $\bar{E}$ and decreases $\underline{E}$ for *any* cutoff (and is therefore what I call "monotone") if we require that for *all* values of y smaller than E, weight shifts downwards to lower values, and for all values of y larger than E, weight shifts upwards to higher values. Formally, this would mean that $F(E) = G(E)$ and that F and G intersect only once, where F cuts G from below ("single-crossing"), and instead of $\int_0^{\hat{y}} [F(y) - G(y)]dy \leq 0$ $\forall \hat{y} \in Y$ we require $\int_0^{\hat{y}} [F(y) - G(y)]dy < 0 \ \forall \hat{y} \in (0, y_{\max})$.[4]

If the income distribution and the fixed cost are initially such that $\Pi(\hat{y}^*) < 0$, an increase in inequality can have an effect on the monopolist's decision of whether or not to offer sorting at some $\hat{y}$, where she compares the profits from offering the sorting technology to 0 (the profits she would make if she stays inactive). An increase in inequality of the form described above, if it is large enough, will make the monopolist's profits positive, which in turn leads the monopolist to become active. As a result, society will become segregated due to an increase in inequality in the form of a mean-preserving spread of the income distribution.

Corollary 3.2 *If society is not segregated initially, a sufficiently high increase in inequality in the form of a monotone mean-preserving spread will make it profitable for a monopolist to offer sorting.*

A mean-preserving spread is not the only type of increase in inequality that increases the monopolist's profits from offering sorting. In fact, from examining the expression for the monopolist's profits, $\hat{y}^*(E - \underline{E}(\hat{y}^*)) - c$, it is straightforward to see that any increase in inequality that increases $E - \underline{E}(\hat{y}^*)$ for the initially optimal cutoff $\hat{y}^*$ will raise the monopolist's profits. In Appendix 3.7.1 I show that if F is lognormal, an increase in the log-variance will also increase the monopolist's profits (if σ is large enough).

Proposition 3.2 *If income is lognormally distributed and the log-variance σ is sufficiently large, an increase in σ leads to an increase in the monopolist's maximal revenue from offering sorting.*

Proof. See Appendix 3.7.1. ■

Remark 3.3 *There is a 1-to-1 relationship between σ and the Gini coefficient. An increase in σ amounts to a median-preserving spread of the income distribution.*

[4] Such a mean-preserving spread can always be constructed if the initial distribution is strictly monotonic. The easiest way is to just transfer mass from the middle of the distribution to the very endpoints of it (i.e. 0 and $y_{\max}$) in such a way that the mean doesn't change.

3.3.2 Welfare

The above section shows that an increase in inequality in the form of a mean-preserving spread increases the monopolist's profit. But what happens to welfare? Total welfare under no sorting is

$$TW^P = \int yEf(y)dy = E^2.$$

Total welfare with two groups and cutoff $\hat{y}$ is[5]

$$TW(\hat{y}) = \underline{E}(\hat{y}) \int_0^{\hat{y}} yf(y)dy + \bar{E}(\hat{y}) \int_{\hat{y}}^{y_{\max}} yf(y)dy - (1 - F(\hat{y}))\hat{y}(\bar{E}(\hat{y}) - \underline{E}(\hat{y})) \tag{3.4}$$

$$= F(\hat{y})(\underline{E}(\hat{y}))^2 + (1 - F(\hat{y}))(\bar{E}(\hat{y}))^2 - (1 - F(\hat{y}))\hat{y}(\bar{E}(\hat{y}) - \underline{E}(\hat{y}))$$

$$= F(\hat{y})(\underline{E}(\hat{y}))^2 + (1 - F(\hat{y}))(\bar{E}(\hat{y}))^2 - \hat{y}(E - \underline{E}(\hat{y}))$$

Levy and Razin (2015) characterize distributions for which sorting is always more efficient than no sorting, irrespective of the cutoff. They show the difference between welfare of sorting at cutoff $\hat{y}$ and welfare of no sorting can be written as

$$TW(\hat{y}) - TW^P = (E - \underline{E}(\hat{y}))(\bar{E}(\hat{y}) - E - \hat{y}) \tag{3.5}$$

and thus two groups yield higher welfare than one group for any $\hat{y}$ iff the income distribution is such that

$$\bar{E}(\hat{y}) - E > \hat{y} \quad \forall \hat{y}. \tag{3.6}$$

This condition is what has in reliability theory been termed the *new worse than used in expectations (NWUE)* property. A distribution F is NWUE if condition (3.6) is satisfied, and *new better than used in expectations (NBUE)* if the opposite holds, i.e.

$$\bar{E}(\hat{y}) - E < \hat{y} \quad \forall \hat{y}.$$

It is immediate to conclude the following:

1. If F is NWUE, sorting at any cutoff is more efficient than no sorting.
2. If F is NBUE, no sorting yields higher welfare than sorting at any $\hat{y}$.
3. If F is not NBUE, then there will always exist some cutoff $\hat{y}$ at which sorting yields a higher welfare than no sorting.

Unless F is NBUE, sorting at some cutoff $\hat{y}$ always yields higher welfare than no sorting. For instance, the lognormal distribution is not NBUE (for no parameter values), hence there always exists a cutoff $\hat{y}$ at which sorting is more efficient than no sorting. On the other hand, the uniform distribution is NBUE, hence no sorting yields higher welfare than sorting at any cutoff.

It is immediate to show that the same mean-preserving spread that increases the monopolist's profits also increases welfare at certain cutoffs $\hat{y}$.

Proposition 3.3 *A monotone mean-preserving spread of the income distribution increases welfare from sorting at those cutoffs where $\bar{E}(\hat{y}) - E > \hat{y}$.*

[5] As in Levy and Razin (2015), total welfare from a particular partition takes into consideration the sorting fee paid (as a deadweight loss to society, or benefitting only a negligible proportion of society). If the sorting fee would not be considered, perfect sorting would always be efficient, because the utility from a match is supermodular (see Becker (1974)).

Proof. If $\bar{E}(\hat{y}) - E > \hat{y}$ then (3.5) tells us that the difference between welfare of sorting at $\hat{y}$ and welfare of no sorting increases due to this mean-preserving spread (both $E - \underline{E}(\hat{y})$ and $\bar{E}(\hat{y}) - E - \hat{y}$ increase). As welfare of no sorting is E^2 and thus doesn't change due to a mean-preserving spread, this implies that welfare of sorting at $\hat{y}$ must increase. ■

Note that no general predictions can be made for welfare at those cutoffs where $\bar{E}(\hat{y}) - E < \hat{y}$: On the one hand, $E - \underline{E}(\hat{y})$ increases, but on the other hand $\bar{E}(\hat{y}) - E - \hat{y}$ is negative (even though the mean-preserving spread will decrease this term in absolute value). The total effect of the mean-preserving spread on (3.5) is thus ambiguous and will depend on the shape of the analyzed income distribution.

If F is NBUE and hence there *is* no cutoff such that $\bar{E}(\hat{y}) - E > \hat{y}$, a mean-preserving spread can make sorting efficient for some cutoffs.

Proposition 3.4 *If F is initially NBUE, a sufficiently large monotone mean-preserving spread of the income distribution will make sorting efficient at some cutoff $\hat{y}$.*

Proof. The mean-preserving spread will increase $\bar{E}(\hat{y}) - E$ for all $\hat{y}$, which will eventually make $\bar{E}(\hat{y}) - E - \hat{y}$ positive for some $\hat{y}$. ■

An increase in inequality will therefore increase welfare of sorting at those cutoffs for which $\bar{E}(\hat{y}) - E - \hat{y} > 0$ and can make sorting at *some* cutoff efficient if F is initially NBUE. Importantly, though, it is not necessarily the case that sorting *at the cutoff that the monopolist chooses* after the increase in inequality yields higher welfare than before As described above, a mean-preserving spread of the income distribution increases welfare of sorting at those cutoffs for which $\bar{E}(\hat{y}) - E - \hat{y} > 0$, but what happens to welfare of sorting at the other cutoffs depends on the shape of the income distribution. Furthermore, even if the monopolist's optimal cutoff is initially such that $\bar{E}(\hat{y}) - E - \hat{y} > 0$, the change in the shape of the income distribution can imply that the monopolist chooses a different cutoff after the mean-preserving spread, at which welfare is lower than before.

The relationship between the monopolist's profit and welfare at the monopolist's optimally chosen cutoff will be the focus of the next section.

3.4 Increasing inequality and the conflict between monopolist profit and welfare

The above analysis shows that an increase in inequality in the form of a monotone mean-preserving spread increases both the monopolist's profit and total welfare from sorting at some cutoffs $\hat{y}$. However, the cutoffs at which the monopolist's profit increases do not have to be the same as the ones where welfare increases. Indeed, if the monopolist chooses to offer sorting at some cutoff due to an increase in inequality, welfare from sorting at this cutoff is not necessarily higher than before - a monopolist's and a benevolent planner's interests are in general not aligned. As I will demonstrate below, total welfare of sorting at the monopolist's optimal cutoff can indeed decline with inequality. In order to show this, I will first analyze how the monopolist's optimal decision (i.e. her optimal cutoff $\hat{y}^*$) is affected by an increase in inequality, for a broad class of income distributions.

At first I will use a simple income distribution to illustrate the potential conflict between monopolist profits and welfare due to increasing inequality. I call this distribution the *symmetric atom distribution*.[6]

[6] This distribution, and also some of the distributions analyzed later in this paper don't satisfy all the conditions that I require in the initial setup of the model, i.e. F is in general not continuous and strictly monotonic. However, this is not a problem for the below calculations.

3.4.1 Symmetric atom distribution

Suppose F has two atoms at 0 and $y_{\max}$, each with mass z, and is uniformly distributed in between.[7] Then average income is $E(y) = \frac{y_{\max}}{2}$ and the conditional expectations are

$$\underline{E}(\hat{y}) = \frac{\left(\frac{1-2z}{y_{\max}}\right)\frac{\hat{y}^2}{2}}{z + \left(\frac{1-2z}{y_{\max}}\right)\hat{y}}$$

and

$$\bar{E}(\hat{y}) = \frac{zy_{\max} + \left(\frac{1-2z}{y_{\max}}\right)\left(\frac{y_{\max}^2}{2} - \frac{\hat{y}^2}{2}\right)}{z + \left(\frac{1-2z}{y_{\max}}\right)(y_{\max} - \hat{y})}.$$

Note that z must be in the interval $[0, 0.5]$ and that $z = 0$ implies that F is uniformly distributed. Furthermore, z parameterizes inequality (in the sense of the difference $\bar{E} - \underline{E}$ for any cutoff), and an increase in z is a monotone mean-preserving spread of the income distribution (and therefore implies an increase in the Gini-coefficient of the distribution).
From Proposition 3.1 we know that the monopolist's profit is increasing in z. In order to identify how the monopolist's optimal cutoff is affected by an increase in inequality, I derive the following Lemma:

Lemma 3.1 *If the income distribution is such that it can be written as $F(y, z)$, where z parameterizes inequality and an increase in z is a monotone mean-preserving spread of the income distribution, then an increase in z increases the monopolist's profit-maximizing cutoff if the income distribution is such that*

$$\frac{\partial^2 \underline{E}(\hat{y}^*, z)}{\partial \hat{y} \partial z} \leq 0 \quad \textit{and} \quad \frac{\partial^2 \underline{E}(\hat{y}^*, z)}{(\partial \hat{y})^2} \geq 0.$$

Proof. If the monopolist's maximization problem has an interior solution, the monopolist's optimal cutoff is characterized via the first order condition

$$\frac{dR(\hat{y}^*, z)}{d\hat{y}} = 0.$$

The monopolist's revenue is

$$R(\hat{y}, z) = \hat{y}(E - \underline{E}(\hat{y}, z))$$

and the optimal cutoff is thus given by

$$E - \underline{E}(\hat{y}^*, z) = \hat{y}^* \frac{\partial \underline{E}(\hat{y}^*, z)}{\partial \hat{y}}.$$

Taking the derivative with respect to z gives

$$-\frac{\partial \underline{E}(\hat{y}^*, z)}{\partial \hat{y}}\frac{d\hat{y}^*}{dz} - \frac{\partial \underline{E}(\hat{y}^*, z)}{\partial z} = \hat{y}\frac{\partial^2 \underline{E}(\hat{y}^*, z)}{(\partial \hat{y})^2}\frac{d\hat{y}^*}{dz} + \frac{\partial \underline{E}(\hat{y}^*, z)}{\partial \hat{y}}\frac{d\hat{y}^*}{dz} + \hat{y}\frac{\partial^2 \underline{E}(\hat{y}^*, z)}{\partial \hat{y} \partial z}$$

and therefore

$$\frac{-\frac{\partial \underline{E}(\hat{y}^*,z)}{\partial z} - \hat{y}^* \frac{\partial^2 \underline{E}(\hat{y}^*,z)}{\partial \hat{y} \partial z}}{\hat{y}^* \frac{\partial^2 \underline{E}(\hat{y}^*,z)}{(\partial \hat{y})^2} + 2\frac{\partial \underline{E}(\hat{y}^*,z)}{\partial \hat{y}}} = \frac{d\hat{y}^*}{dz}.$$

Because an increase in z is a monotone mean-preserving spread, we have that $\frac{\partial \underline{E}(\hat{y},z)}{\partial z} < 0$. Furthermore, an increase in the cutoff always increases average income below the cutoff,

[7] This distribution is very simple and of course not usually encountered in real-life economics. However, I use it because it is easy to handle and - despite its stylized shape - can be deployed to analyze the implications of a society that is "drifting apart", where the rich are getting richer and the poor are becoming poorer.

therefore $\frac{\partial \underline{E}(\hat{y},z)}{\partial \hat{y}} > 0$. Sufficient conditions for

$$\frac{d\hat{y}^*}{dz} > 0$$

are therefore

$$\frac{\partial^2 \underline{E}(\hat{y}^*,z)}{\partial \hat{y} \partial z} \leq 0 \quad \text{and} \quad \frac{\partial^2 \underline{E}(\hat{y}^*,z)}{(\partial \hat{y})^2} \geq 0.$$

The monopolist's profit maximization problem is guaranteed to have an interior solution if the revenue function is strictly concave in $\hat{y}$, i.e. $\frac{\partial^2 R(\hat{y},z)}{(\partial \hat{y})^2} < 0$ for all $\hat{y}$. We have that

$$\frac{\partial^2 R(\hat{y},z)}{(\partial \hat{y})^2} = -2\frac{\partial \underline{E}(\hat{y},z)}{\partial \hat{y}} - \hat{y}\ \frac{\partial^2 \underline{E}(\hat{y},z)}{(\partial \hat{y})^2}.$$

$\frac{\partial \underline{E}(\hat{y},z)}{\partial \hat{y}}$ is always positive, hence the whole expression is negative for sure if $\frac{\partial^2 \underline{E}(\hat{y},z)}{(\partial \hat{y})^2} \geq 0$, which is exactly one of the sufficient conditions above. Hence, this condition ensures both that the monopolist's optimal cutoff is interior and (together with the condition for the cross derivative) that this optimal cutoff increases with inequality. ■

It is straightforward to show that the sufficient conditions from Lemma 3.1 hold for the symmetric atom distribution, and hence the monopolist's optimal cutoff is increasing in z.

Proposition 3.5 *The monopolist's optimal cutoff is increasing in z. For $z = 0$ the optimal cutoff is at $\frac{y_{\max}}{2}$. Hence, the monopolist's optimal cutoff is located in the interval $[\frac{y_{\max}}{2}, y_{\max}]$ for all z.*

Proof. See Appendix 3.7.2. ■

Total welfare without sorting is independent of inequality, it is $E^2 = \frac{y_{\max}^2}{4}$ for all z. I find that for strictly positive z, sorting at small but positive $\hat{y}$ yields higher welfare than no sorting, but sorting at the monopolist's optimal cutoff (which, as Proposition 3.5 shows, is always greater than $\frac{y_{\max}}{2}$) is always less efficient than no sorting. Total welfare is always highest at $\hat{y} = 0$, i.e. if everybody except the mass of people with 0 income is in the rich group.

Proposition 3.6

1. *If $z = 0$ (uniform distribution), maximal total welfare is achieved with no sorting.*
2. *If $z > 0$, maximum welfare is attained at $\hat{y} = 0$ for all z, i.e. it is optimal for the rich group to consist of everybody except people with 0 income. Furthermore, welfare of sorting at $\hat{y} = 0$ is increasing in z.*
3. *If $z > 0$, there is a range of $\hat{y} \geq 0$ for which sorting at these $\hat{y}$ yields higher welfare than no sorting. This range increases with z and becomes $[0, \frac{y_{\max}}{2})$ if $z = 0.5$. No sorting is therefore always more efficient than sorting at the monopolist's optimal cutoff (which is always above $\frac{y_{\max}}{2}$).*

Proof. See Appendix 3.7.2. ■

For the symmetric atom distribution, there exists a conflict between welfare and profit maximization, in the sense that no sorting is always more efficient than sorting at the monopolist's optimal cutoff. The following Proposition shows that this conflict increases with inequality:

Proposition 3.7 *Welfare at the monopolist's optimum is decreasing in z if z is large enough.*

Proof. See Appendix 3.7.2. ■

In addition to analyzing total welfare, I will also examine how welfare of the richest varies with z. The reason why this is interesting is that it gives us an upper bound on how much anybody in society benefits from sorting at some $\hat{y}$ compared to no sorting, due to the following Proposition:

Proposition 3.8 *The utility difference between sorting at some cutoff $\hat{y}$ and no sorting is increasing in y, i.e. if a person with income y prefers no sorting to sorting at some $\hat{y}$, then also everybody with income smaller than y prefers no sorting to sorting.*

Proof. Utility from sorting at $\hat{y}$ for a person with income $y \geq \hat{y}$ is

$$y\bar{E}(\hat{y}) - \hat{y}(\bar{E}(\hat{y}) - \underline{E}(\hat{y}))$$

and utility from no sorting is

$$yE,$$

hence the utility difference amounts to

$$y\bar{E}(\hat{y}) - \hat{y}(\bar{E}(\hat{y}) - \underline{E}(\hat{y})) - yE = (\bar{E}(\hat{y}) - \underline{E}(\hat{y}))(yF(\hat{y}) - \hat{y}),$$

where a positive difference implies that sorting at $\hat{y}$ yields higher utility than no sorting. The derivative of this difference with respect to y (for given $\hat{y}$) is $F(\bar{E}(\hat{y}) - \underline{E}(\hat{y}))$ which is always positive. Hence, utility of sorting is increasing in income for members of the rich group. The people just at $\hat{y}$ - who are in the rich group - will derive utility $\hat{y}\underline{E}$ and everybody in the poor group will derive less utility and it is straightforward to see that utility in the poor group is also increasing in income. Hence, utility from sorting at cutoff $\hat{y}$ is increasing in income for everybody in the economy. ■

I find the following results for welfare of people with income $y_{\max}$ (which I denote by $W_{y_{\max}}$):

Proposition 3.9

1. *If $z = 0$, welfare of people with income $y_{\max}$ is constant and equal to $\frac{y_{\max}^2}{2}$, irrespective of whether there is sorting or not.*
2. *If $z > 0$ then welfare of people with income $y_{\max}$ is equal to $\frac{y_{\max}^2}{2}$ without sorting, but it is higher than $\frac{y_{\max}^2}{2}$ if there is sorting at any cutoff $\hat{y} \in \left[0, \frac{y_{\max}}{2}\right)$. Hence, people with income $y_{\max}$ prefer sorting at any $\hat{y} \in \left[0, \frac{y_{\max}}{2}\right)$ to no sorting. However, no sorting is always preferred to sorting at $\hat{y} > \frac{y_{\max}}{2}$.*
3. *$W_{y_{\max}}$ at those $\hat{y}$ for which sorting is better than no sorting (i.e. all $\hat{y} < \frac{y_{\max}}{2}$) increases with z and is highest if $z = 0.5$.*
4. *If $z > 0$, $W_{y_{\max}}$ is maximized at $\hat{y} = 0$, i.e. when everybody except people with zero income is in the rich group. However, no sorting is always preferred to the monopolist's optimal cutoff for $z > 0$ (because the monopolist's optimal cutoff is always larger than $\frac{y_{\max}}{2}$).*

Proof. See Appendix 3.7.2. ■

Proposition 3.10 *Welfare of the richest from sorting at the monopolist's optimum is decreasing in z.*

Proof. See Appendix 3.7.2. ■

As inequality increases, welfare of the richest in society from sorting at the monopolist's optimal cutoff goes down. An increase in inequality has two effects on the richest people in society: Their group gets richer on average (because there is more mass at the top end and because the cutoff increases) but at the same time they have to pay a higher sorting fee,

because the difference between rich and poor, which determines the sorting fee, increases. The net effect on their welfare is negative.
Finally, in addition to looking at the richest in society, I also analyze how an increase in inequality affects welfare of sorting at the monopolist's optimal cutoff for the poor group. Here, I find the following:

Proposition 3.11 *Average welfare in the poor group from sorting at the monopolist's optimal cutoff decreases due to an increase in inequality.*

Proof. See Appendix 3.7.2. ■

An increase in inequality has two effects on average welfare in the poor group: We know that the monopolist's optimal cutoff increases due to a rise in inequality, which benefits the poor group because people with higher incomes become members of their group and push average income up. However, this increase in the cutoff is not enough to counteract the negative effect of an increasing mass of poor people with zero income in their group, which pulls average income and average welfare down. The overall effect of an increase in inequality is thus negative.

3.4.2 Generalizations

We have seen for the case of the atom distribution that (unless z is very small) welfare at the monopolist's optimal cutoff is decreasing in inequality, and that both welfare of the richest in society and average welfare in the poor group decline as well. Now I want to examine which of these findings apply to a more general class of distributions. First, I will analyze five stylized types of income distributions with the same average income that differ in their implied degree of inequality (measured as $\bar{E} - \underline{E}$ for any cutoff) and analyze how these different degrees of inequality are reflected in monopolist profits and resulting net welfare. These stylized income distributions range from total equality (where everybody in society has the same income) to a distribution that I call "high inequality" (where half of the population have nothing, and half have the maximum possible income). The examined distributions are ordered according to inequality (from most equal to least equal).

- **Total equality**

  If the income distribution is one of total equality, i.e. where everybody has income $\frac{y_{\max}}{2}$, then the monopolist's profits will be 0 (because offering sorting will not be profitable with fixed costs or yield a profit of 0 without fixed costs). Total (net) welfare in this case is $\left(\frac{y_{\max}}{2}\right)^2 = \frac{(y_{\max})^2}{4}$. Note that total welfare without sorting only depends on the expected value of the income distribution. As average income is the same for all the distributions in this analysis, total welfare without sorting doesn't change, it is $\frac{(y_{\max})^2}{4}$ in all cases.

- **Triangle distribution**

  If income is distributed in a triangular (isosceles) shape on $[0, y_{\max}]$ such that the density is

$$f(y) = \frac{4}{(y_{\max})^2} y \quad \text{if } y \in \left[0, \frac{y_{\max}}{2}\right]$$
$$f(y) = \frac{4}{y_{\max}} - \frac{4}{(y_{\max})^2} y \quad \text{if } y \in \left[\frac{y_{\max}}{2}, y_{\max}\right]$$

  the profit-maximizing cutoff for the monopolist is $\hat{y}^* = \frac{3y_{\max}}{8}$ and the resulting profits are $\frac{3(y_{\max})^2}{32}$. Welfare from sorting at this cutoff amounts to $\frac{3059}{529}\frac{(y_{\max})^2}{32} < \frac{(y_{\max})^2}{4}$. Hence, welfare is maximized when there is no sorting.

- **Uniform distribution**

  If income is uniformly distributed on $[0, y_{\max}]$, the monopolist's profit maximizing cutoff is $\hat{y}^* = \frac{y_{\max}}{2}$ and the resulting profit is $\frac{(y_{\max})^2}{8}$. Welfare at this cutoff is $\frac{3(y_{\max})^2}{16} < \frac{(y_{\max})^2}{4}$. Hence, welfare is maximized with no sorting.

- **Reverse triangle distribution**

  If income is distributed in a reverse-triangular (isosceles) shape on $[0, y_{\max}]$ such that the density is

$$
\begin{aligned}
f(y) &= \frac{2}{y_{\max}} - \frac{4}{(y_{\max})^2}y \qquad \text{if } y \in \left[0, \frac{y_{\max}}{2}\right] \\
f(y) &= -\frac{2}{y_{\max}} + \frac{4}{(y_{\max})^2}y \qquad \text{if } y \in \left[\frac{y_{\max}}{2}, y_{\max}\right]
\end{aligned}
$$

  the monopolist's optimal cutoff is $\hat{y}^* = 0.64 y_{\max}$, which yields a profit of $0.1935 y_{\max}^2$. Total welfare at this cutoff is $0.163 y_{\max}^2 < \frac{(y_{\max})^2}{4}$. Again, no sorting would be best for welfare.

- **High inequality**

  If half of the population has 0 income and half of them earn $y_{\max}$, the optimal cutoff for the monopolist is $\hat{y} = y_{\max}$ with corresponding sorting fee $\frac{y_{\max}}{2}$. Note that due to the jump in F at $y_{\max}$ (F is not continuous here!) the sorting fee is not uniquely determined, any $b \in (0, \frac{y_{\max}}{2}]$ would work, and the monopolist will choose the highest in this interval to maximize her profits (and therefore the profits will be $\frac{(y_{\max})^2}{2}$). Welfare in this case would be 0. Welfare would be maximized with the same partition, i.e. a poor group with zero income and a rich group with income $y_{\max}$, but with the lowest of feasible sorting fees, i.e. b being just ϵ over 0. Resulting welfare would be $\frac{y_{\max}}{2} - \frac{\epsilon}{2}$. If the mass at both endpoints is not equal, this last result holds as well, because it is always better to separate rich and poor if the sorting fee is negligible, due to the supermodularity of utility from sorting (see Becker (1974)). The monopolist's profit in that latter case is increasing in the mass of rich people relative to poor people.

From this simple analysis I can conclude the following for these five distributions:

1. As inequality increases (in terms of discrete jumps from one distribution to another), the monopolist's profits increase.

2. As inequality increases, the monopolist's optimal cutoff increases.

3. Total welfare is independent of inequality in the absence of sorting, it depends only on average income. For all the above analyzed distributions, no sorting is more efficient than sorting at any cutoff $\hat{y}$.

4. If the monopolist chooses the cutoff, then welfare is highest in the case of total equality (because the sorting fee is 0 in that case and the situation is equal to no sorting, which is optimal for all the distributions discussed above). The next highest welfare would be achieved in the uniform case, followed by the triangular and then the reverse triangular case, and the case of total inequality would be worst for welfare (given the sorting fee that the monopolist would charge). Hence - if we exclude the case of total equality and start from a triangular distribution - welfare of sorting at the monopolist's optimal cutoff initially increases with inequality, but as inequality becomes too high the monopolist can claim a huge part of the gross benefits from sorting for herself and net welfare decreases.

For the symmetric atom distribution, I found that welfare from sorting at the monopolist's cutoff was decreasing in inequality. Here, we see that if we don't only look at mean-preserving spreads of the uniform distribution, but actually allow inequality also to be smaller than for a uniform distribution, the picture is different: Welfare increases with inequality for small rates of inequality, and decreases thereafter. In Appendix 3.7.4, I show that this is true not only for the above discrete jumps in inequality but also if we look at continuous changes in inequality for these types of distributions. In particular, I examine a distribution that is, for low levels of inequality, shaped like a house, and then as inequality increases becomes uniform and in the end looks like a reverse house (or trough). The two extreme cases are thus the triangle distribution (low inequality) and the reverse triangle distribution (high inequality) from above. I find the same results for this continuous version of the stylized distributions above: welfare of sorting at y^* increases in inequality for low rates of inequality, and decreases for high rates. In a sense, there is thus less of a conflict between profit maximization and welfare for low rates of inequality than for high rates. However, note that all these distributions, ranging from the triangle to the reverse triangle one and all degrees of inequality in between, are NBUE and hence no sorting yields higher welfare than sorting at any cutoff (see Appendix 3.7.7). For low rates of inequality, an increase in inequality increases welfare at the monopolist's optimal cutoff, but a benevolent social planner would nevertheless prefer to have no sorting at all in those cases.

In Appendix 3.7.5, I analyze the lognormal distribution and show that monopolist profit-maximization and welfare maximization are not necessarily opposed goals if inequality is low. However, also for this type of distribution the conflict between welfare and monopolist's profits increases for high rates of inequality.

3.5 Multiple groups

The previous sections examine how increasing inequality affects welfare and profits if the monopolist can choose one cutoff and thus offer segregation into two groups. I have shown that the interests of a profit-maximizing monopolist and a benevolent social planner are generally not aligned, and that the conflict between those interests increases with inequality. In the following section I compare these results to a situation where the monopolist doesn't face costs of offering segregation and can therefore offer infinitely many groups (i.e. perfect sorting) if she wants. I will demonstrate that the findings from the previous sections hold in some sense also for this more general setting: There is a way in which an increase in inequality increases the conflict between monopolist's profits and welfare (and lets the monopolist extract more surplus, if she can decide on the menu of sorting fees).

Before looking at the monopolist's optimization problem, let me first examine what is best for welfare if multiple groups are possible. Hoppe et al. (2009) show that if the income distribution is such that the coefficient of variation, which is given by

$$CV = \frac{\sqrt{Var(y)}}{E(y)},$$

is larger than 1 then perfect sorting is better than no sorting for welfare, and if $CV \leq 1$, the opposite holds:

Proposition 3.12 (Hoppe et al. (2009)) *Perfect sorting is more (less) efficient than no sorting iff* $CV \geq (\leq) 1$.

Note that the coefficient of variation is a measure of inequality - it is high if the difference between the standard deviation and the average is high, and it increases due to a mean-preserving spread of the income distribution. Hence, another way to interpret the above Proposition is in terms of inequality: For low rates of inequality, no sorting is more efficient than perfect sorting, whereas if inequality is high, perfect sorting yields higher welfare than no sorting.

It is straightforward to show that the triangle distribution, the uniform distribution and the reverse triangle distribution discussed in the previous section and the house distribution discussed in the Appendix (which encompasses all the others) are NBUE (see Appendix 3.7.7). As NBUE implies that the coefficient of variation is smaller than 1, this means that no sorting yields higher welfare than perfect sorting for these distributions.

The symmetric atom distribution is not NBUE - indeed I have shown in the previous section that for small $\hat{y}$ sorting yields higher welfare than no sorting. However, the symmetric atom distribution has $CV \leq 1$ and therefore perfect sorting always yields lower welfare than no sorting (see Appendix 3.7.6).

For the lognormal distribution, the coefficient of variation can be written as

$$CV = \sqrt{e^{\sigma^2} - 1}.$$

Hence, the coefficient of variation of a lognormal distribution depends only on σ, not on μ. This is intuitive, because the CV is an inequality measure, and inequality in the lognormal distribution depends on σ, and not on μ (there is also a 1-to-1 relationship between σ and the Gini coefficient). The coefficient of variation for the lognormal distribution is greater than 1 iff

$$\begin{aligned} CV &\geq 1 \iff \\ \sqrt{e^{\sigma^2} - 1} &\geq 1 \iff \\ \sigma &\geq \sqrt{\ln(2)} \approx 0.83 \end{aligned}$$

Hence perfect sorting yields higher welfare than no sorting iff $\sigma \geq 0.83$. (Note: If we calibrate μ and σ in the lognormal distribution to match the first and second moment of the US household distribution, we get $\mu \approx 10.85$ and $\sigma \approx 0.85$.)

After characterizing the class of distributions for which perfect sorting is more efficient than no sorting, I will now analyze the monopolist's optimization problem: What is the monopolist's optimal sorting schedule if she doesn't face any costs of offering the technology? Rayo (2013) characterizes the optimal placement of regions of pooling and perfect sorting, depending on the shape of the income distribution. In the following I want to examine the implications of changing inequality on the monopolist's optimal sorting schedule and total welfare.

Rayo shows that if (and only if) the function $h(y) = y - \frac{1-F(y)}{f(y)}$ is nondecreasing everywhere, perfect sorting is the profit-maximizing sorting schedule. If there are regions of y for which $h(y)$ is decreasing, perfect sorting is not optimal for the monopolist and she will want to introduce intervals of y for which she pools everybody into one joint group.[8]

It is immediate to see that h is always decreasing if the distribution has an increasing failure rate (IFR). Hence, if a distribution exhibits IFR, perfect sorting is optimal for the monopolist. We can therefore conclude the following:

Corollary 3.3 *If the income distribution exhibits IFR, a monopolist and a benevolent social planner have conflicting interests: No sorting is more efficient than perfect sorting or any type of finite sorting, but the monopolist wants to implement perfect sorting to maximize her profits.*

[8] Note however that there are never two pooling intervals next to each other (i.e. pooling intervals are always maximal) and that pooling is never optimal at the top end of the distribution.

Proof. Levy and Razin (2015) show that an increasing failure rate of the distribution implies that the distribution is NBUE, which in turn implies that the coefficient of variation is smaller than 1. Hence, for distributions which exhibit IFR, no sorting yields higher welfare than any finite sorting (see Levy and Razin (2015)) and perfect sorting (see Hoppe et al. (2009)). ■

What happens within the class of distributions for which perfect sorting is optimal for the monopolist (note that this class contains the family of IFR distributions, which are characterized by low inequality in terms of the coefficient of variation, because $IFR \Rightarrow CV \leq 1$) if inequality increases in the sense of a mean-preserving spread of the income distribution? We know that welfare and monopolist profit are both

$$\int \frac{y^2}{2} f(y) dy = \frac{E(y^2)}{2}$$

in this case, i.e. both the monopolist and the citizens get half of the total surplus from perfect sorting (see Rayo (2013)). Hence, whenever a change to the distribution happens such that perfect sorting is still optimal for the monopolist afterwards, welfare and profits are affected in the same way, i.e. a benevolent social planner's and a monopolist's interests are aligned. For instance, look at the effects of a mean-preserving spread: The variance increases but average income doesn't change. Because of

$$Var(y) = E(y^2) - (E(y))^2$$

this implies that $E(y^2)$ must increase due to a mean-preserving spread, which means that a mean-preserving spread increases both welfare and the monopolist's profits in this case.

Proposition 3.13 *If the income distribution is such that perfect sorting is optimal for the monopolist, welfare and monopolist profits benefit equally from an increase in inequality in the form of a mean-preserving spread.*

Proof. See above. ■

The conflict between monopolist profits and welfare is thus not further intensified as inequality increases within the class of distributions for which perfect sorting is optimal for the monopolist: A benevolent social planner would always prefer no sorting to perfect sorting, but as inequality increases, both welfare and profits increase equally.

Importantly, the above result applies to small (infinitesimal) increases in inequality, such that perfect sorting still remains optimal for the monopolist. If the shape of the distribution changes too much, perfect sorting might no longer be the optimal sorting schedule for the monopolist. For instance, it can be shown in simulations that in case of the lognormal distribution, the function $h(.)$ is everywhere increasing in y for small σ (below 1), and hence perfect sorting is optimal for the monopolist. However, as σ increases further, there is an increasing region of $\hat{y}$ for which h is decreasing, which implies that pooling some regions of Y is optimal for the monopolist.

What if the income distribution is not such that perfect sorting is optimal for the monopolist? We know that total surplus (just the sum of the utilities) is always maximized with perfect sorting, due to supermodularity of the utility function (see Becker (1974)): Pooling everybody yields a total surplus of $(E(y))^2$ while perfect sorting yields $E(y^2)$, which is always larger because $E(y^2) = Var(y) + (E(y))^2$. The same holds for pooling intervals of y. But that implies that perfect sorting is always better for welfare than any other sorting that the monopolist would design. Why? Total surplus is maximized with perfect sorting, anything else must yield either the same surplus or less. With perfect sorting, citizens and the monopolist share the surplus equally. If the monopolist decides that she would rather not do perfect sorting, it means she must expect a higher surplus with another way of sorting, which must mean that the citizens get less than half of total surplus (and that total surplus might even be lower than that of perfect sorting).

Proposition 3.14 *If the income distribution is such that the monopolist doesn't want to implement perfect sorting, a benevolent social planner would always prefer perfect sorting to the monopolist's sorting schedule. With her optimal sorting schedule, the monopolist can rake more than half of the total surplus from sorting.*

Proof. See above. ■

To conclude, the conflict between monopolist's profit and total welfare has many facets in the case of multiple groups: If the distribution exhibits IFR (which implies that $CV \leq 1$ (low inequality)), the citizens would prefer no sorting to sorting, while the monopolist wants perfect sorting, but the conflict doesn't intensify with inequality: As inequality increases (in the form of a mean-preserving spread) but we stay within the class of distributions such that the monopolist wants perfect sorting (IFR is a sufficient condition for that), welfare and monopolists profits increase equally. If we start from a situation where perfect sorting is optimal for the monopolist and inequality increases such that the monopolist implements a different sorting schedule (and pools some intervals of Y), monopolist profits will increase by more and welfare will increase by less than if sorting would still be perfect (and total surplus is either equally high or less than under perfect sorting).

What remains unexplored is what happens to profits and total welfare if we already start from a situation where perfect sorting is not optimal for the monopolist and then see an increase in inequality. I leave this question open for future research.

3.6 Conclusion

In this paper, I have discussed how changes in inequality affect socioeconomic segregation and resulting welfare in society. I have used a simple two-group model to show that a rise in inequality always increases profits of a monopolist who offers the sorting technology. Corresponding welfare in society, however, increases in line with profits only for small rates of inequality. As inequality becomes higher, a conflict between welfare and profits arises, and welfare decreases with inequality if the monopolist implements sorting to maximize her profits. At the end of the paper I discuss how these findings generalize if the monopolist is not restricted to offer only one cutoff. If the income distribution is such that perfect sorting is optimal for the monopolist initially, the prediction is clear: there is a conflict between welfare and profits, because no sorting would be welfare maximizing. The conflict doesn't intensify for small increases in inequality, such that perfect sorting remains optimal, but the monopolist is able to capture more than half of the total surplus if pooling for some intervals of Y becomes optimal. The case where perfect sorting is not optimal for the monopolist to begin with remains to be explored in future research.

In the present paper, segregation does not affect people's beliefs: they retain perfect knowledge of the whole income distribution, despite interacting only with a select (and potentially very small) group of people. In Windsteiger (2017a and 2017b) I relax this assumption and explore the impact of endogenous beliefs about society that are affected by social interactions. However, I do not model the supply side of sorting explicitly in these papers. To combine these two approaches remains a promising future avenue of research.

3.7 Appendix

3.7.1 Lognormal distribution: increase in σ

Suppose that F is lognormally distributed with parameters μ and σ. The monopolist's profits from offering sorting are

$$\Pi = R(\hat{y}^*) - c = \hat{y}^*(\bar{E} - \underline{E})(1 - F(\hat{y}^*)) - c,$$

where $\hat{y}^*$ maximizes profits. Note that $\hat{y}^*$ will never be 0 because then $R(\hat{y})$ would be 0, whereas for any $\hat{y} \in (0, \infty)$ $R(\hat{y}) > 0$. However, as in the lognormal case $y_{\max} = \infty$, I need an extra condition to ensure that $\hat{y}^* = \infty$ is not optimal. The easiest way to ensure this is to show that $R(\hat{y}) \to 0$ if $\hat{y} \to \infty$. I find that

$$\lim_{\hat{y}\to\infty} R(\hat{y}) = \lim_{\hat{y}\to\infty} \hat{y}(E - \underline{E}) = \lim_{\hat{y}\to\infty} \frac{E - \underline{E}}{\frac{1}{\hat{y}}} = \lim_{\hat{y}\to\infty} \frac{-\frac{f(\hat{y}-\underline{E})}{F}}{-\frac{1}{\hat{y}^2}} = \lim_{\hat{y}\to\infty} \frac{f\hat{y}^2(\hat{y} - \underline{E})}{F} = 0,$$

where the last step comes from the fact that the third moment of the lognormal distribution is finite, which implies that $f\hat{y}^3 \to 0$ (and thus I have shown that $R(\hat{y}) \to 0$ if $\hat{y} \to \infty$ is satisfied for any income distribution with finite third moment and continuous pdf).
If income is lognormally distributed, $R(\hat{y})$ can be written as

$$R(\hat{y}) = \hat{y}(\bar{E} - \underline{E})(1 - F(\hat{y})) = \hat{y}(E - \underline{E})$$
$$= \hat{y}E(y)\left(1 - \frac{\Phi(a-\sigma)}{\Phi(a)}\right)$$

where

$$a = \frac{\ln \hat{y} - \mu}{\sigma}$$

and

$$E(y) = e^{\mu + \frac{\sigma^2}{2}}$$

What happens to the monopolist's revenue if σ changes? Note that in general also the profit-maximizing cutoff $\hat{y}^*$ will change, but let me just look at the change in $\hat{y}(E - \underline{E})$ if σ changes but $\hat{y}$ stays the same. If the expression increases with σ while keeping $\hat{y}$ constant, then it increases even more with the new optimal cutoff, hence showing that $\hat{y}(E - \underline{E})$ is increasing in σ with constant cutoff is sufficient to show that an increase in σ increases the monopolist's revenue.
The derivative of $R(\hat{y})$ with respect to σ is

$$\frac{\partial R}{\partial \sigma} = \frac{\partial}{\partial \sigma}\left[\hat{y}e^{\mu+\frac{\sigma^2}{2}}\left(1 - \frac{\Phi(a-\sigma)}{\Phi(a)}\right)\right]$$

Let me denote

$$G := 1 - \frac{\Phi(a-\sigma)}{\Phi(a)}$$

then

$$\frac{\partial}{\partial \sigma}\left[\hat{y}e^{\mu+\frac{\sigma^2}{2}}G\right] = \hat{y}e^{\mu+\frac{\sigma^2}{2}}\left[\sigma G + \frac{\partial G}{\partial \sigma}\right] \tag{3.7}$$

Note that

$$\frac{\partial G}{\partial \sigma} = -\frac{\phi(a-\sigma)(-\frac{a}{\sigma} - 1)\Phi(a) + \Phi(a-\sigma)\phi(a)\frac{a}{\sigma}}{\Phi(a)^2} = \frac{\phi(a-\sigma)(1+\frac{a}{\sigma})}{\Phi(a)} - \frac{\Phi(a-\sigma)\phi(a)\frac{a}{\sigma}}{\Phi(a)^2}$$

and that (3.7) is positive iff

$$\sigma G + \frac{\partial G}{\partial \sigma} > 0$$

which can be written as

$$\iff \sigma\left(1 - \frac{\Phi(a-\sigma)}{\Phi(a)}\right) + \phi(a-\sigma)\left(1 + \frac{a}{\sigma}\right)\left(\frac{1}{\Phi(a)}\right) - \phi(a)\frac{a}{\sigma}\frac{\Phi(a-\sigma)}{\Phi(a)^2} > 0 \tag{3.8}$$

$$\iff \sigma\left(1 - \frac{\Phi(a-\sigma)}{\Phi(a)}\right) +$$

$$+\frac{1}{\Phi(a)}\left(\phi(a-\sigma)\left(1 + \frac{a}{\sigma}\right) - \phi(a)\frac{a}{\sigma}\frac{\Phi(a-\sigma)}{\Phi(a)}\right) > 0$$

In order for this to be positive, a sufficient condition is

$$\phi(a-\sigma)\left(1 + \frac{a}{\sigma}\right) - \phi(a)\frac{a}{\sigma}\frac{\Phi(a-\sigma)}{\Phi(a)} > 0$$

(because $1 > \frac{\Phi(a-\sigma)}{\Phi(a)}$). It is immediate to see that this always holds if $a < 0$ and $1 + \frac{a}{\sigma} \geq 0$ (i.e. σ large enough for a given $\hat{y}$). It also always holds if $a \geq 0$, because

$$\phi(a-\sigma)(1 + \frac{a}{\sigma}) - \phi(a)\frac{a}{\sigma}\frac{\Phi(a-\sigma)}{\Phi(a)} > 0$$

$$\iff \frac{\phi(a-\sigma)}{\Phi(a-\sigma)}(1 + \frac{a}{\sigma}) > \frac{\phi(a)}{\Phi(a)}\frac{a}{\sigma}$$

This last expression holds for all $a \geq 0$ because $\frac{\phi(x)}{\Phi(x)}$, which is density over distribution of a standard normal distribution, is decreasing in x.
Hence, the revenue from offering a cutoff at $\hat{y}$ is increasing in σ whenever $\hat{y} \geq e^{\mu}$ or $e^{\mu} > \hat{y} \geq e^{\mu-\sigma^2}$ (this is what $a < 0$ and $1 + \frac{a}{\sigma} \geq 0$ translates to).
Remember, technically I do not need to show that the revenue increases for any $\hat{y}$, just for the optimal $\hat{y}^*$. If σ is large enough then for a given $\hat{y}$, $\hat{y} \geq e^{\mu-\sigma^2}$ will always be satisfied. Therefore, the last remaining step to show that the monopolist's maximized revenue is increasing in σ for large enough σ is to show that $\hat{y}^*$ doesn't converge to 0 if σ increases: then there will always exist a $\bar{\sigma}$ such that if inequality is higher than $\bar{\sigma}$, the revenue from offering a cutoff $\hat{y}$ (optimally) increases if inequality increases. To show this last step, I look at the first derivative of the revenue function

$$R'(\hat{y}) = E(y)\left(1 - \frac{\Phi(a-\sigma)}{\Phi(a)} - \frac{1}{\sigma}\left(\frac{\phi(a-\sigma)}{\Phi(a)} - \frac{\Phi(a-\sigma)\phi(a)}{(\Phi(a))^2}\right)\right)$$

and determine the limit as $\sigma \to \infty$. Note that this implies that $a \to 0$ and hence $\Phi(a-\sigma) \to 0$, $\Phi(a) \to 0.5$, $\phi(a-\sigma) \to 0$ and $\phi(a) \to \phi(0)$. Hence, the expression in brackets goes to 1, and because $E(y) \to \infty$, I get that $R'(\hat{y}) \to \infty$. But this implies that $\hat{y}^*$ cannot go to zero if $\sigma \to \infty$, because this means that at any cutoff $\hat{y}$, the gain from increasing it a little bit becomes infinitely large.
Hence, if σ is large enough the profit from offering segregation is sure to increase with σ, i.e. (3.7) is > 0 for sure. This means that if inequality is already high and increases further, this will increase the monopolist's revenue.

3.7.2 Proofs for the symmetric atom distribution

Proposition 3.15 *The monopolist's optimal cutoff is increasing in z. For $z = 0$ the optimal cutoff is at $\frac{y_{\max}}{2}$. Hence, the monopolist's optimal cutoff is located in the interval $[\frac{y_{\max}}{2}, y_{\max}]$*

for all z.

Proof. As Lemma 3.1 establishes, sufficient conditions for

$$\frac{d\hat{y}^*}{dz} > 0$$

are

$$\frac{\partial^2 \underline{E}(\hat{y}^*, z)}{\partial\hat{y}\partial z} < 0 \quad \text{and} \quad \frac{\partial^2 \underline{E}(\hat{y}^*, z)}{(\partial\hat{y})^2} > 0.$$

Show that $\frac{\partial^2 \underline{E}(\hat{y},z)}{\partial\hat{y}\partial z} < 0$: (Note: To simplify the notation I set $y_{\max} = 1$ in the following calculations, but everything works analogously if the distribution is scaled up to a general $y_{\max} > 0$).

$$\underline{E}(\hat{y}, z) = \frac{(1-2z)\frac{\hat{y}^2}{2}}{z + (1-2z)\hat{y}}$$

$$\begin{aligned}
\frac{\partial \underline{E}(\hat{y}, z)}{\partial \hat{y}} &= \frac{(1-2z)\hat{y}(z+(1-2z)\hat{y}) - (1-2z)^2\frac{\hat{y}^2}{2}}{(z+(1-2z)\hat{y})^2} \\
&= \frac{(1-2z)}{z+(1-2z)\hat{y}}\left(\hat{y} - \frac{(1-2z)\frac{\hat{y}^2}{2}}{z+(1-2z)\hat{y}}\right)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 \underline{E}(\hat{y}, z)}{\partial\hat{y}\partial z} &= \frac{-2(z+(1-2z)\hat{y}) - (1-2z)(1-2\hat{y})}{(z+(1-2z)\hat{y})^2}\left(\hat{y} - \frac{(1-2z)\frac{\hat{y}^2}{2}}{z+(1-2z)\hat{y}}\right) \\
&\quad - \frac{(1-2z)}{z+(1-2z)\hat{y}}\left(\frac{\hat{y}^2}{2}\left(\frac{-2(z+(1-2z)\hat{y}) - (1-2z)(1-2\hat{y})}{(z+(1-2z)\hat{y})^2}\right)\right) \\
&= \frac{-1}{(z+(1-2z)\hat{y})^2}\left(\hat{y} - \frac{(1-2z)\hat{y}^2}{z+(1-2z)\hat{y}}\right) \\
&= \frac{-1}{(z+(1-2z)\hat{y})^2}\left(\frac{z+(1-2z)\hat{y} - (1-2z)\hat{y}}{z+(1-2z)\hat{y}}\right)\hat{y} \\
&= \frac{-1}{(z+(1-2z)\hat{y})^2}\left(\frac{z}{z+(1-2z)\hat{y}}\right)\hat{y} < 0
\end{aligned}$$

Show that $\frac{\partial^2 \underline{E}(\hat{y},z)}{(\partial\hat{y})^2} > 0$:

$$\begin{aligned}
\frac{\partial^2 \underline{E}(\hat{y}, z)}{(\partial\hat{y})^2} &= -\frac{(1-2z)^2}{(z+(1-2z)\hat{y})^2}\left(\hat{y} - \frac{(1-2z)\frac{\hat{y}^2}{2}}{z+(1-2z)\hat{y}}\right) \\
&\quad + \frac{(1-2z)}{z+(1-2z)\hat{y}}\left(1 + \frac{(1-2z)^2}{(z+(1-2z)\hat{y})^2}\frac{\hat{y}^2}{2} - \frac{(1-2z)\hat{y}}{z+(1-2z)\hat{y}}\right)
\end{aligned}$$

$$= \frac{(1-2z)}{z+(1-2z)\hat{y}}\left(1 - \frac{(1-2z)\hat{y}}{z+(1-2z)\hat{y}}\right)^2 = \frac{(1-2z)}{z+(1-2z)\hat{y}}\left(\frac{z}{z+(1-2z)\hat{y}}\right)^2 > 0$$

If $z = 0$, then

$$\underline{E}(\hat{y}) = \frac{\hat{y}}{2}$$

and therefore

$$R(\hat{y}) = \hat{y}\left(\frac{y_{\max}}{2} - \frac{\hat{y}}{2}\right) = \frac{\hat{y}y_{\max}}{2} - \frac{\hat{y}^2}{2}$$

This implies that profit is maximized at $\frac{y_{\max}}{2}$ for $z = 0$. ■

Proposition 3.16 *1. If $z = 0$ (uniform distribution), maximal total welfare is achieved* *with no sorting.*

2. *If $z > 0$, maximum welfare is attained at $\hat{y} = 0$ for all z, i.e. it is optimal for the rich group to consist of everybody except people with 0 income. Furthermore, welfare of sorting at $\hat{y} = 0$ is increasing in z.*

3. *If $z > 0$, there is a range of $\hat{y} \geq 0$ for which sorting at these $\hat{y}$ yields higher welfare than no sorting. This range increases with z and becomes $[0, \frac{y_{\max}}{2})$ if $z = 0.5$. No sorting is therefore always more efficient than sorting at the monopolist's optimal cutoff (which is always above $\frac{y_{\max}}{2}$).*

Proof. First of all, note that for strictly positive z welfare jumps at 0: If people with 0 income are included in the group (so there is only one group), total welfare is $\frac{y_{\max}^2}{4}$, if 0 is excluded, then welfare is $\frac{y_{\max}^2}{4-4z} > \frac{y_{\max}^2}{4}$.
For the remainder of this proof I will again set $y_{\max} = 1$ for simplicity of notation. Total welfare at cutoff $\hat{y}$ is then given by

$$\begin{aligned} TW(\hat{y}) &= (z + (1-2z)\hat{y}) \left(\frac{(1-2z)\frac{\hat{y}^2}{2}}{(z + (1-2z)\hat{y})} \right)^2 \\ &\quad + (z + (1-2z)(1-\hat{y})) \left(\frac{z + (1-2z)\left(\frac{1}{2} - \frac{\hat{y}^2}{2}\right)}{z + (1-2z)(1-\hat{y})} \right)^2 \\ &\quad - \hat{y} \left(\frac{1}{2} - \frac{(1-2z)\frac{\hat{y}^2}{2}}{z + (1-2z)\hat{y})} \right) \end{aligned}$$

If $z = 0$ the distribution becomes uniform and total welfare is

$$TW_{z=0}(\hat{y}) = \hat{y}\frac{\hat{y}^2}{4} + (1-\hat{y}) \left(\frac{\frac{1}{2}(1-\hat{y}^2)}{(1-\hat{y})} \right)^2 - \hat{y}\left(\frac{1}{2} - \frac{\hat{y}}{2}\right) = \frac{1}{4} - \frac{\hat{y}}{4} + \frac{\hat{y}^2}{4}$$

It is straightforward to see that this quadratic function reaches its minimum at $\hat{y} = 0.5$ and is maximized at the endpoints of the examined interval Y, i.e. $\hat{y} = 0$ and $\hat{y} = 1$, where total welfare is $\frac{1}{4}$, which is equal to the total welfare of no sorting. Hence, no sorting is (weakly) preferred to sorting at any cutoff $\hat{y} \in Y$ if $z = 0$.
For the general case, where $z \neq 0$, note first that total welfare at cutoff 0 is

$$TW(0) = \frac{\left(z + (1-2z)\frac{1}{2}\right)^2}{z + (1-2z)} = \frac{1}{4(1-z)}$$

which is increasing in z for all $z \in [0, 0.5]$. It is also straightforward to see that this expression is always larger than $\frac{1}{4}$ (welfare of no sorting) if $z > 0$.
At cutoff 1 this becomes

$$TW(1) = \frac{1-2z}{4(1-z)}$$

which is decreasing in z for all $z \in [0, 0.5]$. Note that this is always smaller than $\frac{1}{4}$ (welfare of no sorting) for all $z > 0$.
For all cutoffs in between 0 and 1, note that from the previous section we know that sorting yields higher welfare than no sorting at cutoff $\hat{y}$ iff

$$\bar{E} - E - \hat{y} > 0$$

Plugging in the expressions for $\bar{E}$ and E for the atom distribution, this condition becomes

$$\frac{(1-2z)\hat{y}^2 - \hat{y} + z}{2(z + (1-2z)(1-\hat{y}))} > 0.$$

As the numerator of this fraction is positive for all z and $\hat{y}$, the condition can be simplified to

$$(1-2z)\hat{y}^2 - \hat{y} + z > 0$$

It is immediate to see that this condition never holds if $z = 0$, holds for all $z > 0$ at cutoff 0, and holds for all $\hat{y} \leq 0.5$ if $z = 0.5$. The roots of $(1-2z)\hat{y}^2 - \hat{y} + z$ are

$$y_{1,2} = \frac{1 \pm \sqrt{1 - 4z + 8z^2}}{2 - 4z}$$

and the polynomial is positive for all $\hat{y}$ that are either smaller than the smaller of the two or larger than the larger of the two roots. As the larger root is always ≥ 1, the only relevant case for us is the range of $\hat{y}$ smaller than $y_1 = \frac{1-\sqrt{1-4z+8z^2}}{2-4z}$. The value of y_1 is 0 if $z = 0$ and is then increasing in z, until it reaches $y_1 = 0.5$ for $z = 0.5$. Hence, the range of $\hat{y}$ for which sorting is better than no sorting is $[0, y_1(z)]$ for all $z > 0$ where $y_1(z)$ is increasing in z, 0 for $z = 0$ and reaches 0.5 for $z = 0.5$. ■

Proposition 3.17 *Welfare at the monopolist's optimum is decreasing in z if z is large enough.*

Proof. The derivative of total welfare with respect to z at the monopolist's optimal cutoff $\hat{y}^*$ amounts to

$$\frac{dTW(\hat{y}^*, z)}{dz} = \left(f(\underline{E}^2 - \bar{E}^2) + F2\underline{E}\frac{\partial \underline{E}}{\partial \hat{y}} + (1-F)2\bar{E}\frac{\partial \bar{E}}{\partial \hat{y}} \right) \frac{d\hat{y}^*}{dz}$$

$$+\frac{\partial F}{\partial z}(\underline{E}^2 - \bar{E}^2) + F2\underline{E}\frac{\partial \underline{E}}{\partial z} + (1-F)2\bar{E}\frac{\partial \bar{E}}{\partial z} - \frac{d\Pi(\hat{y}^*, z)}{dz}$$

where $\Pi(\hat{y}^*, z)$ is the monopolist's maximized profit. We know that the monopolist's profit maximization always has an interior solution (see Lemma 3.1 and Proposition 3.5). Hence the optimal cutoff y^* is characterized via the first order condition

$$\frac{\partial \Pi(\hat{y}^*, z)}{\partial \hat{y}} = 0$$

This implies that

$$\frac{d\Pi(\hat{y}^*, z)}{dz} = \frac{\partial \Pi(\hat{y}^*, z)}{\partial \hat{y}} \frac{d\hat{y}^*}{dz} - \hat{y}^* \frac{\partial \underline{E}}{\partial z} = -\hat{y}^* \frac{\partial \underline{E}}{\partial z} (> 0).$$

Hence, the above expression can be simplified to

$$\frac{dTW(\hat{y}^*, z)}{dz} = f(\bar{E} - \underline{E})(\bar{E} + \underline{E} - 2\hat{y}^*)\frac{d\hat{y}^*}{dz} +$$

$$+\frac{\partial F}{\partial z}(\underline{E}^2 - \bar{E}^2) + (F2\underline{E} + \hat{y}^*)\frac{\partial \underline{E}}{\partial z} + (1-F)2\bar{E}\frac{\partial \bar{E}}{\partial z}$$

(where I also use the fact that $\frac{\partial \underline{E}(\hat{y},z)}{\partial \hat{y}} = f(\hat{y}, z)\frac{\hat{y} - \underline{E}(\hat{y},z)}{F(\hat{y},z)}$ and $\frac{\partial \bar{E}(\hat{y},z)}{\partial \hat{y}} = f(\hat{y}, z)\frac{\bar{E}(\hat{y},z) - \hat{y}}{1 - F(\hat{y},z)}$).
We have

$$\frac{\partial F}{\partial z} = 1 - 2\hat{y}$$

and

$$\frac{\partial \underline{E}}{\partial z} = \frac{-\frac{\hat{y}^2}{2}}{(z+(1-2z)\hat{y})^2}$$

and

$$\frac{\partial \bar{E}}{\partial z} = \frac{\frac{1}{2}(\hat{y}-1)^2}{(z+(1-2z)(1-\hat{y}))^2}.$$

Note that $\frac{\partial \underline{E}}{\partial z} < 0$, $\frac{\partial F}{\partial z} < 0$ (because $\hat{y}^* > \frac{y_{\max}}{2}$) and $\frac{\partial \bar{E}}{\partial z} > 0$. From Proposition 3.5 we know that $\frac{d\hat{y}^*}{dz} > 0$, hence sufficient conditions for $\frac{dTW(\hat{y}^*,z)}{dz} < 0$ are that

$$\bar{E} + \underline{E} - 2\hat{y}^* < 0$$

and

$$\frac{\partial F}{\partial z}(\underline{E}^2 - \bar{E}^2) + (F2\underline{E} + \hat{y}^*)\frac{\partial \underline{E}}{\partial z} + (1-F)2\bar{E}\frac{\partial \bar{E}}{\partial z} < 0$$

The first condition can easily be shown to always hold for $\hat{y}^* > 0.5$: Plugging in the expressions for $\bar{E}$ and $\underline{E}$ yields

$$\bar{E} + \underline{E} - 2\hat{y}^* = \left(\frac{z}{2} - z\hat{y}\right) + (2\hat{y} - 6\hat{y}^2 + 4\hat{y}^3)(\frac{1}{4} + z^2 - z)$$

We have that

$$\left(\frac{z}{2} - z\hat{y}\right) < 0 \quad \forall \hat{y} > 0.5$$

and

$$(2\hat{y} - 6\hat{y}^2 + 4\hat{y}^3) < 0 \quad \forall \hat{y} > 0.5$$

while

$$(\frac{1}{4} + z^2 - z) > 0 \quad \forall z > 0.5$$

Hence, the total expression is always negative for $\hat{y} > 0.5$. For the second condition, note that

$$\frac{\partial F}{\partial z}(\underline{E}^2 - \bar{E}^2) + (F2\underline{E} + \hat{y}^*)\frac{\partial \underline{E}}{\partial z} + (1-F)2\bar{E}\frac{\partial \bar{E}}{\partial z} =$$

$$= \quad \frac{(1-2z)^2\left(\frac{\hat{y}^4}{4} - \frac{\hat{y}^5}{2}\right) - (1-2z)\frac{\hat{y}^4}{2} - \frac{\hat{y}^3}{2}}{(z+(1-2z)\hat{y})^2} + \\ + \frac{z(1-y)^2 + (1-2z)\left(\frac{1}{2} - \frac{\hat{y}^2}{2}\right)(1-\hat{y})^2 - (1-2\hat{y})\left(z + (1-2z)\left(\frac{1}{2} - \frac{\hat{y}^2}{2}\right)\right)^2}{(z+(1-2z)(1-\hat{y}))^2}$$

The first summand of this expression is negative for all $\hat{y} > 0.5$ and all z, but the second term is always positive (note that $1 - 2\hat{y} < 0$ for all $\hat{y} > 0.5$). If $z = 0$ the sum of the two becomes $\frac{1}{4} - \frac{1}{2}\hat{y}$ which is negative for all $\hat{y} > 0.5$, however if $z > 0$ then there is a small range of $\hat{y} > 0.5$ for which the second term is higher in absolute value than the first and hence the whole expression is positive. Indeed it can be shown that the entire expression for $\frac{dTW(\hat{y}^*,z)}{dz}$ is positive for small $\hat{y}^* > 0.5$ (from numerical simulations). As $\hat{y}^*$ is close to 0.5 for small z this implies that total welfare from sorting at the monopolist's optimal cutoff increases with z for very small z. However, note that the monopolist's optimal cutoff increases with z as well, and this increase moves $\hat{y}^*$ out of the area for which total welfare increases with z quickly. It can be seen (from simulations) that for all $z > 0.05$ the small range of $\hat{y}$ for which total welfare increases with z is below $\hat{y}^*$ for all z. Hence, total welfare from sorting at the monopolist's optimal cutoff decreases with z if $z > 0.05$. ■

Proposition 3.18 *1. If $z = 0$, welfare of people with income $y_{\max}$ is constant and equal* *to $\frac{y_{\max}^2}{2}$, irrespective of whether there is sorting or not.*

2. *If $z > 0$ then welfare of people with income $y_{\max}$ is equal to $\frac{y_{\max}^2}{2}$ without sorting, but it is higher than $\frac{y_{\max}^2}{2}$ if there is sorting at any cutoff $\hat{y} \in \left[0, \frac{y_{\max}}{2}\right)$. Hence, people with income $y_{\max}$ prefer sorting at any $\hat{y} \in \left[0, \frac{y_{\max}}{2}\right)$ to no sorting. However, no sorting is always preferred to sorting at $\hat{y} > \frac{y_{\max}}{2}$.*

3. *$W_{y_{\max}}$ at those $\hat{y}$ for which sorting is better than no sorting (i.e. all $\hat{y} < \frac{y_{\max}}{2}$) increases with z and is highest if $z = 0.5$.*

4. *If $z > 0$, $W_{y_{\max}}$ is maximized at $\hat{y} = 0$, i.e. when everybody except people with zero income is in the rich group. However, no sorting is always preferred to the monopolist's optimal cutoff for $z > 0$ (because the monopolist's optimal cutoff is always larger than $\frac{y_{\max}}{2}$).*

Proof. Welfare of people with income $y_{\max}$ can be calculated as

$$W_{y_{\max}}(\hat{y}, z) = y_{\max}\bar{E} - \hat{y}(\bar{E} - \underline{E}).$$

This can be written as (again set $y_{\max} = 1$)

$$W_{y_{\max}}(\hat{y}, z) = \frac{\frac{z}{2} - \frac{3}{2}z\hat{y} + \frac{z\hat{y}^2}{2} + z\hat{y}^3 + z^2\hat{y}^2 - 2z^2\hat{y}^3 + \frac{\hat{y}}{2} - \frac{\hat{y}^2}{2}}{(z + (1-2z)\hat{y})(z + (1-2z)(1-\hat{y}))}.$$

If $z = 0$ (uniform distribution) this becomes

$$W_{y_{\max}}(\hat{y}, 0) = \frac{1}{2}$$

Note that utility of no sorting is also $\frac{1}{2}$ for people with income $y_{\max}$, they are therefore indifferent between sorting and no sorting at any cutoff if $z = 0$.
If $z > 0$: When is $W_{y_{\max}}(\hat{y}) > 0.5$ (=utility from no sorting), i.e. for what range of cutoffs is sorting preferred to no sorting for the richest people?

$$W_{y_{\max}}(\hat{y}, z) > 0.5$$

$$\iff \frac{z}{2} - \frac{3}{2}z\hat{y} + \frac{z\hat{y}^2}{2} + z\hat{y}^3 + z^2\hat{y}^2 - 2z^2\hat{y}^3 + \frac{\hat{y}}{2} - \frac{\hat{y}^2}{2} > \frac{(z + (1-2z)\hat{y})(z + (1-2z)(1-\hat{y}))}{2}$$

$$z(\hat{y} - 3\hat{y}^2 + 2\hat{y}^3) + z^2(6\hat{y}^2 - 4\hat{y}^3 - 4\hat{y} + 1) > 0$$

If $z > 0$ this becomes

$$(\hat{y} - 3\hat{y}^2 + 2\hat{y}^3) > z(-6\hat{y}^2 + 4\hat{y}^3 + 4\hat{y} - 1)$$

The RHS is positive if $\hat{y} \geq 0.5$, which is also exactly when the LHS is negative. (The polynomial on the left has roots 1 and 0.5 and is smaller than zero in between the two and larger than zero elsewhere. The polynomial on the right has only root 0.5 in the interval $[0, 1]$ and is positive above 0.5 and negative below). In other words, the inequality cannot hold for any positive z if $\hat{y} \geq 0.5$ and will always hold if $\hat{y} \leq 0.5$. This means that for any z the richest people prefer sorting to no sorting at any cutoff below 0.5. It is straightforward to see that they are always indifferent between sorting and no sorting at $\hat{y} = 0.5$. The maximum utility is reached at $\hat{y} = 0$ (meaning that the rich group consists of everybody except the poor with zero income) for any $z > 0$, which can be concluded from the fact that $\frac{dW_{y_{\max}}(\hat{y}, z)}{d\hat{y}} < 0$ for all $\hat{y} \in [0, 1]$.

Proof that $\frac{dW_{y_{\max}}(\hat{y},z)}{d\hat{y}}$ for all $\hat{y} \in [0,1]$:

$$\frac{dW_{y_{\max}}(\hat{y},z)}{d\hat{y}} = \frac{\begin{array}{c}-z^4 + (1-2z)z^3(-0.5 + 3\hat{y}^2 - 3\hat{y}) + (1-2z)^2 2z^2\hat{y}(-1+\hat{y}) \\ +(1-2z)^3 z(-\hat{y}^2 - \hat{y}^4 + 2\hat{y}^3)\end{array}}{(z + (1-2z)\hat{y})^2(z + (1-2z)(1-\hat{y}))^2}$$

The denominator is always positive, so it suffices to focus on the numerator. Analysis of the factors that multiply the potencies of $(1-2z)$ shows that they are negative for all $\hat{y} \in [0,1]$ and hence $\frac{dW_{y_{\max}}(\hat{y},z)}{d\hat{y}}$ is smaller than 0 for all $\hat{y} \in [0,1]$. The maximum welfare for the rich is therefore achieved when $\hat{y} = 0$ (i.e. the rich group consists of everybody except the poorest with income 0). This maximum welfare is increasing in z:

$$W_{y_{\max}}(0,z) = \frac{1}{2(1-z)}$$

On the other hand, welfare at $\hat{y} = y_{\max}(=1)$ is decreasing in z:

$$W_{y_{\max}}(1,z) = \frac{\frac{1}{2} - z}{(1-z)}.$$

■

Proposition 3.19 *Welfare of the richest from sorting at the monopolist's optimum is decreasing in z.*

Proof.

$$W_{y_{\max}}(\hat{y}^*, z) = (y_{\max} - \hat{y}^*)\bar{E}(\hat{y}^*) + \hat{y}^*\underline{E}(\hat{y}^*)$$

The monopolist's optimal cutoff satisfies the FOC and hence

$$E - \underline{E}(\hat{y}^*) = \hat{y}^* \frac{\partial \underline{E}(\hat{y}^*)}{\partial \hat{y}} \tag{3.9}$$

The derivative of $W_{y_{\max}}(\hat{y}^*, z)$ with respect to z is:

$$\frac{dW_{y_{\max}}(\hat{y}^*,z)}{dz} = (y_{\max} - \hat{y}^*)\left(\frac{\partial \bar{E}}{\partial z}\right) + \hat{y}^*\frac{\partial \underline{E}}{\partial z} + \left[(y_{\max} - \hat{y}^*)\frac{\partial \bar{E}}{\partial \hat{y}} - (\bar{E} - \underline{E}) + \hat{y}^*\frac{\partial \underline{E}}{\partial \hat{y}}\right]\frac{d\hat{y}^*}{dz}$$

Using (3.9) this becomes

$$\frac{dW_{y_{\max}}(\hat{y}^*,z)}{dz} = (y_{\max} - \hat{y}^*)\left(\frac{\partial \bar{E}}{\partial z}\right) + \hat{y}^*\frac{\partial \underline{E}}{\partial z} + \left[(y_{\max} - \hat{y}^*)\frac{\partial \bar{E}}{\partial \hat{y}} - (\bar{E} - E)\right]\frac{d\hat{y}^*}{dz}$$

Hence, sufficient conditions for $W_{y_{\max}}(\hat{y}^*, z)$ to be decreasing in z are that

$$(y_{\max} - \hat{y}^*)\left(\frac{\partial \bar{E}}{\partial z}\right) + \hat{y}^*\frac{\partial \underline{E}}{\partial z} < 0$$

and

$$\left[(y_{\max} - \hat{y}^*)\frac{\partial \bar{E}}{\partial \hat{y}} - (\bar{E} - E)\right] < 0$$

(because we already know that $\frac{d\hat{y}^*}{dz} > 0$). Again setting $y_{\max} = 1$ and using

$$\frac{\partial \bar{E}}{\partial z} = \frac{\frac{1}{2}(1-\hat{y})^2}{(z + (1-2z)(1-\hat{y}))^2}$$

$$\frac{\partial \underline{E}}{\partial z} = \frac{-\hat{y}^2}{2(z + (1-2z)\hat{y})^2}$$

and

$$\frac{\partial \bar{E}}{\partial \hat{y}} = (1-2z)(1-\hat{y})\frac{z+(1-2z)\left(\frac{1}{2}-\frac{\hat{y}}{2}\right)}{(z+(1-2z)(1-\hat{y}))^2},$$

it can be shown that both terms are negative for $\hat{y}^* \in [0.5, 1]$ and all $z \in [0, 0.5]$: I get that

$$(y_{\max} - \hat{y}^*)\left(\frac{\partial \bar{E}}{\partial z}\right) + \hat{y}^* \frac{\partial \underline{E}}{\partial z} = \frac{(1-\hat{y})^3}{2(z+(1-2z)(1-\hat{y}))^2} - \frac{\hat{y}^3}{2(z+(1-2z)\hat{y})^2}$$

$$= \frac{(1-\hat{y})^3(z+(1-2z)\hat{y})^2 - \hat{y}^3(z+(1-2z)(1-\hat{y}))^2}{2(z+(1-2z)(1-\hat{y}))^2(z+(1-2z)\hat{y})^2}.$$

The denominator is positive, and the numerator can be simplified to give

$$z^2(1-3\hat{y}+3\hat{y}^2-2\hat{y}^3) + (2z-4z^2)(\hat{y}-3\hat{y}^2+2\hat{y}^3) + (1-2z)^2\hat{y}^2(1-4\hat{y}+5\hat{y}^2-2\hat{y}^3).$$

It turns out that the polynomials of $\hat{y}$ in each summand are negative for all $\hat{y} \in [0.5, 1]$, hence the expression is negative for the relevant ranges of $\hat{y}$ and all z.
Furthermore,

$$\left[(y_{\max} - \hat{y}^*)\frac{\partial \bar{E}}{\partial \hat{y}} - (\bar{E} - \underline{E})\right] = (1-2z)(1-\hat{y})^2\frac{z+(1-2z)\left(\frac{1}{2}-\frac{\hat{y}}{2}\right)}{(z+(1-2z)(1-\hat{y}))^2}$$

$$-\frac{z+(1-2z)\left(\frac{1}{2}-\frac{\hat{y}^2}{2}\right)}{z+(1-2z)(1-\hat{y})} + \frac{1}{2}$$

$$= \frac{-z^2+(1-2z)z(1-4\hat{y}+3\hat{y}^2)+(1-2z)^2(1-4\hat{y}+5\hat{y}^2-2\hat{y}^3)}{2(z+(1-2z)(1-\hat{y}))^2}$$

Again the denominator is positive and all the polynomials of $\hat{y}$ in the numerator are negative $\forall \hat{y} \in [0.5, 1]$, which implies that the expression is negative for the relevant ranges of $\hat{y}$ and all z. Hence, $\frac{dW_{y_{\max}}(\hat{y}^*,z)}{dz}$ is negative for all $\hat{y}$ in $[0.5, 1]$ for all z and thus $W_{y_{\max}}(\hat{y}^*, z)$ is decreasing in z. ■

Proposition 3.20 *Average welfare in the poor group from sorting at the monopolist's optimal cutoff decreases due to an increase in inequality.*

Proof. Average welfare in the poor group amounts to $\underline{E}^2$ (note that they don't have to pay the sorting fee b). We know that

$$\frac{d\underline{E}(\hat{y}^*, z)}{dz} = \frac{\partial \underline{E}}{\partial \hat{y}}\frac{d\hat{y}^*}{dz} + \frac{\partial \underline{E}}{\partial z}$$

From above we know that

$$\frac{\partial \underline{E}}{\partial \hat{y}} = \frac{(1-2z)\left(\hat{y}z+(1-2z)\frac{\hat{y}^2}{2}\right)}{(z+(1-2z)\hat{y})^2}$$

and

$$\frac{\partial \underline{E}}{\partial z} = \frac{-\hat{y}^2}{2(z+(1-2z)\hat{y})^2}$$

and plugging in all the expressions for the derivatives in $\frac{d\hat{y}^*}{dz}$ yields

$$\frac{d\hat{y}^*}{dz} = \frac{\left(\frac{3z\hat{y}}{2}+(1-2z)\frac{\hat{y}^2}{2}\right)}{(1-2z)(3z^2+3(1-2z)\hat{y}z+(1-2z)^2\hat{y}^2)}$$

Hence, after simplifications, we get that

$$\frac{d\underline{E}(\hat{y}^*,z)}{dz} = \frac{-(1-2z)^2\frac{\hat{y}^4}{4} - (1-2z)\frac{\hat{y}^3 z}{4}}{(3z^2+3(1-2z)\hat{y}z+(1-2z)^2\hat{y}^2)\left(z+(1-2z)\hat{y}\right)^2}$$

The denominator is always positive and the numerator is always negative, hence $\frac{d\underline{E}(\hat{y}^*,z)}{dz} < 0$.
■

3.7.3 Calculations for Section 3.4.2

Triangle distribution

If the density is

$$\begin{aligned} f(y) &= \frac{4}{(y_{\max})^2}y \quad \text{if } y \in \left[0, \frac{y_{\max}}{2}\right] \\ f(y) &= \frac{4}{y_{\max}} - \frac{4}{(y_{\max})^2}y \quad \text{if } y \in \left[\frac{y_{\max}}{2}, y_{\max}\right] \end{aligned}$$

then I can calculate that if the cutoff $\hat{y}$ is in $\left[0, \frac{y_{\max}}{2}\right]$ the conditional expectations are

$$\underline{E}(\hat{y}) = \frac{2\hat{y}}{3}$$

and

$$\bar{E}(\hat{y}) = \frac{\frac{y_{\max}}{2} - \frac{4\hat{y}^3}{3y_{\max}^2}}{1 - \frac{2\hat{y}^2}{y_{\max}^2}}$$

whereas if $\hat{y} \in \left[\frac{y_{\max}}{2}, y_{\max}\right]$ the expressions become

$$\underline{E}(\hat{y}) = \frac{\frac{2\hat{y}^2}{y_{\max}} - \frac{4}{3}\frac{\hat{y}^3}{y_{\max}^2} - \frac{y_{\max}}{6}}{\frac{4\hat{y}}{y_{\max}}\left(1 - \frac{\hat{y}}{2y_{\max}}\right) - 1}$$

and

$$\bar{E}(\hat{y}) = \frac{\frac{2y_{\max}}{3} - \frac{2\hat{y}^2}{y_{\max}} + \frac{4}{3}\frac{\hat{y}^3}{y_{\max}^2}}{2 - \frac{4\hat{y}}{y_{\max}} + \frac{2\hat{y}^2}{y_{\max}^2}}$$

Monopolist profits are

$$\Pi(\hat{y}) = \hat{y}(E - \underline{E}(\hat{y})).$$

It is straightforward to show that $\Pi(.)$ reaches a local maximum at $\frac{3y_{\max}}{8}$ if $\hat{y} \le \frac{y_{\max}}{2}$ and is decreasing in $\hat{y}$ for all $\hat{y} > \frac{y_{\max}}{2}$. Hence, cutoff $\hat{y}^* = \frac{3y_{\max}}{8}$ yields the maximal profit, and $\Pi(\hat{y}^*) = \frac{3}{32}y_{\max}^2$. Welfare at this cutoff is given by

$$U^S(\hat{y}^*) = F(\hat{y}^*)\left(\underline{E}(\hat{y}^*)\right)^2 + (1-F(\hat{y}^*))\left(\bar{E}(\hat{y}^*)\right)^2 - \Pi(\hat{y}^*) = \frac{y_{\max}^2}{32}\left(\frac{3059}{529}\right) \approx 0.1807y_{\max}^2$$

Uniform distribution

We have that

$$\underline{E}(\hat{y}) = \frac{\hat{y}}{2}$$

and

$$\bar{E}(\hat{y}) = \frac{y_{\max} + \hat{y}}{2}$$

and thus

$$\Pi(\hat{y}) = \frac{y_{\max}\hat{y}}{2} - \frac{\hat{y}^2}{2}$$

which is maximized at

$$\hat{y}^* = \frac{y_{\max}}{2}.$$

Total welfare at $\hat{y}^*$ is

$$U^S(\hat{y}^*) = \frac{3y_{\max}^2}{16}.$$

Reverse triangle distribution

If the density is

$$\begin{aligned} f(y) &= \frac{2}{y_{\max}} - \frac{4}{(y_{\max})^2}y \qquad \text{if } y \in \left[0, \frac{y_{\max}}{2}\right] \\ f(y) &= -\frac{2}{y_{\max}} + \frac{4}{(y_{\max})^2}y \qquad \text{if } y \in \left[\frac{y_{\max}}{2}, y_{\max}\right] \end{aligned}$$

then I can calculate that if the cutoff $\hat{y}$ is in $\left[0, \frac{y_{\max}}{2}\right]$ the conditional expectations are

$$\underline{E}(\hat{y}) = \frac{\frac{\hat{y}}{2} - \frac{2}{3}\frac{\hat{y}^2}{y_{\max}}}{1 - \frac{\hat{y}}{y_{\max}}}$$

and

$$\bar{E}(\hat{y}) = \frac{\frac{y_{\max}}{2} - \frac{2\hat{y}}{y_{\max}}\left(\frac{\hat{y}}{2} - \frac{2}{3}\frac{\hat{y}^2}{y_{\max}}\right)}{1 - \frac{2\hat{y}}{y_{\max}}\left(1 - \frac{\hat{y}}{y_{\max}}\right)}$$

whereas if $\hat{y} \in \left[\frac{y_{\max}}{2}, y_{\max}\right]$ the expressions become

$$\underline{E}(\hat{y}) = \frac{\frac{y_{\max}}{6} - \frac{2\hat{y}^2}{y_{\max}}\left(\frac{1}{2} - \frac{2}{3}\frac{\hat{y}}{y_{\max}}\right)}{1 - \frac{2\hat{y}}{y_{\max}}\left(1 - \frac{\hat{y}}{y_{\max}}\right)}$$

and

$$\bar{E}(\hat{y}) = \frac{\frac{y_{\max}}{3} + \frac{2\hat{y}^2}{y_{\max}}\left(\frac{1}{2} - \frac{2}{3}\frac{\hat{y}}{y_{\max}}\right)}{\frac{2\hat{y}}{y_{\max}}\left(1 - \frac{\hat{y}}{y_{\max}}\right)}.$$

Monopolist profits are

$$\Pi(\hat{y}) = \hat{y}(E - \underline{E}(\hat{y})).$$

It is straightforward to show that $\Pi(.)$ reaches a local maximum at $0.64y_{\max}$ (numerically calculated) if $\hat{y} > \frac{y_{\max}}{2}$ and is decreasing in $\hat{y}$ for all $\hat{y} \leq \frac{y_{\max}}{2}$. Hence, cutoff $\hat{y}^* = 0.64y_{\max}$ yields the maximal profit, and $\Pi(\hat{y}^*) \approx 0.1935y_{\max}^2$. Welfare at this cutoff is given by

$$U^S(\hat{y}^*) = F(\hat{y}^*)\left(\underline{E}(\hat{y}^*)\right)^2 + (1 - F(\hat{y}^*))\left(\bar{E}(\hat{y}^*)\right)^2 - \Pi(\hat{y}^*) \approx 0.163y_{\max}^2$$

3.7.4 House distribution, uniform distribution, trough distribution

Suppose income is distributed according to an income distribution with pdf $f(.)$ such that

$$\begin{aligned} f(y) &= x - \frac{2z}{y_{\max}}y \qquad \text{if } y \in \left[0, \frac{y_{\max}}{2}\right] \\ f(y) &= x - 2z + \frac{2z}{y_{\max}}y \qquad \text{if } y \in \left[\frac{y_{\max}}{2}, y_{\max}\right] \end{aligned}$$

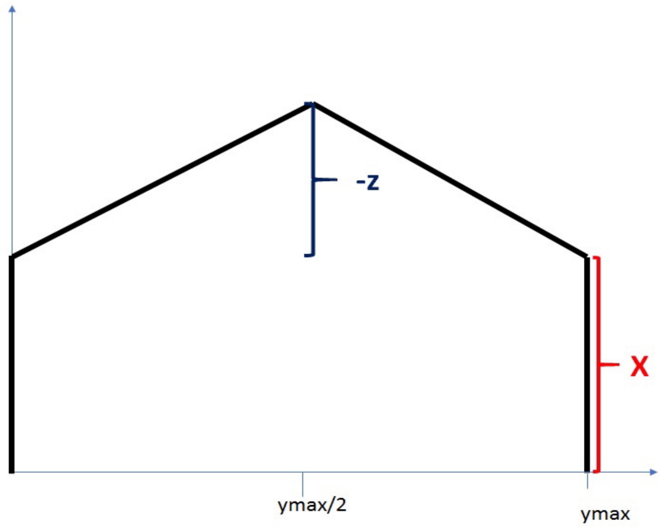

Figure 3.1: House distribution

Note that we must have

$$x = \frac{1}{y_{\max}} + \frac{z}{2} \tag{3.10}$$

in order for $F(y_{\max}) = 1$ and

$$x \in \left[0, \frac{2}{y_{\max}}\right], \quad z \in \left[-\frac{2}{y_{\max}}, \frac{2}{y_{\max}}\right]$$

If $z = 0$ then $x = \frac{1}{y_{\max}}$ and the distribution is uniform, if $z = -\frac{2}{y_{\max}}$ then x is 0 and the pdf has the shape of an isosceles triangle. If $z = \frac{2}{y_{\max}}$ then $x = \frac{2}{y_{\max}}$ and the pdf has the shape of an inverse triangle. If $z \in \left(-\frac{2}{y_{\max}}, 0\right)$ then the distribution has the shape of a house, if $z \in \left(0, \frac{2}{y_{\max}}\right)$ the distribution has the shape of a (triangular) trough (see Figures 3.1 and 3.2). The larger z, the higher is inequality (in terms of $\bar{E} - \underline{E}$ for any given cutoff) and an increase in z amounts to a mean-preserving spread of the income distribution. Note that average income is constant $\forall z \in \left[-\frac{2}{y_{\max}}, \frac{2}{y_{\max}}\right]$, $E(y) = \frac{y_{\max}}{2}$.
Using (3.10), the pdf can be rewritten as

$$\begin{aligned} f(y) &= \frac{1}{y_{\max}} + \frac{z}{2} - \frac{2z}{y_{\max}} y \qquad \text{if } y \in \left[0, \frac{y_{\max}}{2}\right] \\ f(y) &= \frac{1}{y_{\max}} - \frac{3z}{2} + \frac{2z}{y_{\max}} y \qquad \text{if } y \in \left[\frac{y_{\max}}{2}, y_{\max}\right] \end{aligned}$$

If the cutoff $\hat{y}$ is in the interval $\left[0, \frac{y_{\max}}{2}\right]$, we have that

$$\underline{E}(\hat{y}) = \frac{\int_0^{\hat{y}} \left(\frac{1}{y_{\max}} + \frac{z}{2} - \frac{2z}{y_{\max}} y\right) y dy}{F(\hat{y})} = \frac{\int_0^{\hat{y}} \left(\frac{1}{y_{\max}} + \frac{z}{2} - \frac{2z}{y_{\max}} y\right) y dy}{\int_0^{\hat{y}} \left(\frac{1}{y_{\max}} + \frac{z}{2} - \frac{2z}{y_{\max}} y\right) dy}$$

$$= \frac{\frac{\hat{y}^2}{2y_{\max}} + \frac{\hat{y}^2 z}{4} - \frac{2z\hat{y}^3}{3y_{\max}}}{\frac{\hat{y}(1 - z\hat{y})}{y_{\max}} + \frac{\hat{y}z}{2}} = \frac{6\hat{y} + 3\hat{y}zy_{\max} - 8z\hat{y}^2}{12 - 12z\hat{y} + 6y_{\max}z}$$

If the cutoff is above $\frac{y_{\max}}{2}$ we need to calculate $\underline{E}(\hat{y})$ differently: The easiest way is to calculate

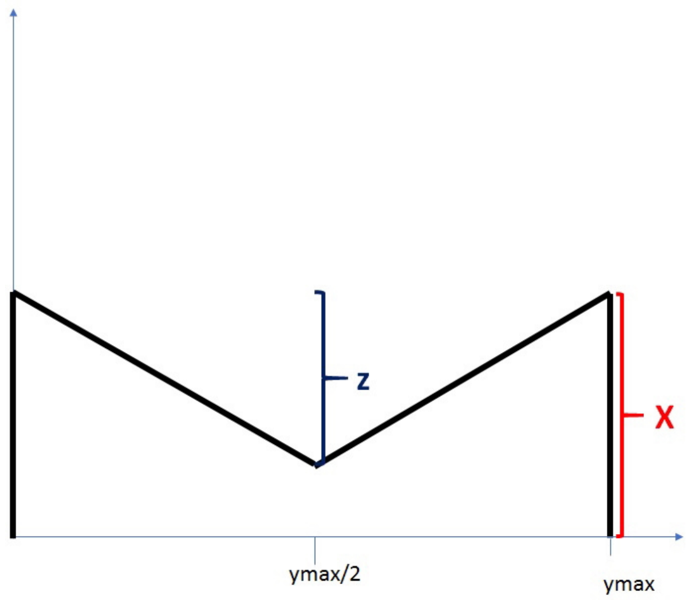

Figure 3.2: Trough distribution

$\bar{E}\left(\hat{y}\right)$ first

$$\bar{E}(\hat{y}) = \frac{\int\limits_{\hat{y}}^{y_{\max}} \left(\frac{1}{y_{\max}} - \frac{3z}{2} + \frac{2z}{y_{\max}}y\right) ydy}{1 - F(\hat{y})} = \frac{\int\limits_{\hat{y}}^{y_{\max}} \left(\frac{1}{y_{\max}} - \frac{3z}{2} + \frac{2z}{y_{\max}}y\right) ydy}{\int\limits_{\hat{y}}^{y_{\max}} \left(\frac{1}{y_{\max}} - \frac{3z}{2} + \frac{2z}{y_{\max}}y\right) dy}$$

$$= \frac{\frac{y_{\max}}{2} - \frac{1}{12}zy_{\max}^2 - \frac{\hat{y}^2}{2y_{\max}} + \frac{3}{4}z\hat{y}^2 - \frac{2}{3}\frac{z}{y_{\max}}\hat{y}^3}{1 - \frac{1}{2}zy_{\max} - \frac{\hat{y}}{y_{\max}} + \frac{3z\hat{y}}{2} - \frac{z\hat{y}^2}{y_{\max}}}$$

and then calculate $\underline{E}(\hat{y})$ via the formula

$$E = F\left(\hat{y}\right)\underline{E}(\hat{y}) + (1 - F(\hat{y}))\bar{E}(\hat{y})$$

(noting that $E = \frac{y_{\max}}{2}$), which gives

$$\underline{E}(\hat{y}) = \frac{\frac{1}{12}zy_{\max}^2 + \frac{\hat{y}^2}{2y_{\max}} - \frac{3}{4}z\hat{y}^2 + \frac{2}{3}\frac{z}{y_{\max}}\hat{y}^3}{\frac{1}{2}zy_{\max} + \frac{\hat{y}}{y_{\max}} - \frac{3z\hat{y}}{2} + \frac{z\hat{y}^2}{y_{\max}}}$$

$$= \frac{zy_{\max}^2 + \frac{6\hat{y}^2}{y_{\max}} - 9z\hat{y}^2 + \frac{8z\hat{y}^3}{y_{\max}}}{6zy_{\max} + \frac{12\hat{y}}{y_{\max}} - 18z\hat{y} + \frac{12\hat{y}^2z}{y_{\max}}}.$$

Using these expressions, I can show the following:

Proposition 3.21 *If $z \in \left[-\frac{2}{y_{\max}}, 0\right]$, the monopolist's optimal cutoff is in the interval $\left[0, \frac{y_{\max}}{2}\right]$, if $z = 0$ the monopolist's optimal cutoff is $\hat{y}^* = \frac{y_{\max}}{2}$ and if $z \in \left[0, \frac{2}{y_{\max}}\right]$, the monopolist's optimal cutoff is in the interval $\left[\frac{y_{\max}}{2}, y_{\max}\right]$.*

Proof. The monopolist's profit at cutoff $\hat{y}$ is given by

$$\Pi(\hat{y}) = \hat{y}(E - \underline{E})$$

Using the expressions for $\underline{E}$ from above, I find that

$$\Pi(\hat{y}) = \hat{y}\left(\frac{y_{\max}}{2} - \frac{6\hat{y} + 3\hat{y}zy_{\max} - 8z\hat{y}^2}{12 - 12z\hat{y} + 6y_{\max}z}\right) \tag{3.11}$$

if $\hat{y} \in \left[0, \frac{y_{\max}}{2}\right]$ and

$$\Pi(\hat{y}) = \hat{y}\left(\frac{y_{\max}}{2} - \frac{zy_{\max}^2 + \frac{6\hat{y}^2}{y_{\max}} - 9z\hat{y}^2 + \frac{8z\hat{y}^3}{y_{\max}}}{6zy_{\max} + \frac{12\hat{y}}{y_{\max}} - 18z\hat{y} + \frac{12\hat{y}^2 z}{y_{\max}}}\right) \tag{3.12}$$

if $\hat{y} \in \left[\frac{y_{\max}}{2}, y_{\max}\right]$. It can be calculated (numerically) that (3.11) has a local and global maximum in $\left[0, \frac{y_{\max}}{2}\right]$ when $z < 0$, while (3.12) has a local and global maximum in $\left[\frac{y_{\max}}{2}, y_{\max}\right]$ when $z > 0$. ■

Proposition 3.22 *The monopolist's profit-maximizing cutoff $\hat{y}^*$ is increasing in z for all $z \in \left[-\frac{2}{y_{\max}}, \frac{2}{y_{\max}}\right]$.*

Proof. Note that

$$\frac{d\hat{y}^*}{dz} = \frac{-\frac{\partial \underline{E}(\hat{y}^*,z)}{\partial z} - \hat{y}^* \frac{\partial^2 \underline{E}(\hat{y}^*,z)}{\partial \hat{y}\partial z}}{\hat{y}^* \frac{\partial^2 \underline{E}(\hat{y},z)}{(\partial \hat{y})^2} + 2\frac{\partial \underline{E}(\hat{y}^*,z)}{\partial \hat{y}}}$$

and hence according to Lemma 3.1, sufficient conditions for

$$\frac{d\hat{y}^*}{dz} > 0$$

are

$$\frac{\partial^2 \underline{E}(\hat{y},z)}{\partial \hat{y}\partial z} \leq 0 \quad \text{and} \quad \frac{\partial^2 \underline{E}(\hat{y},z)}{(\partial \hat{y})^2} \geq 0.$$

Show $\frac{\partial^2 \underline{E}(\hat{y},z)}{(\partial \hat{y})^2} \geq 0$ if $z < 0$:
If $z < 0$ we know that the monopolist's optimal cutoff is in the interval $[0, \frac{y_{\max}}{2})$. Setting $y_{\max} = 1$ again for simplicity of notation we have that

$$\underline{E}(\hat{y}, z) = \frac{6\hat{y} + 3\hat{y}z - 8z\hat{y}^2}{12 - 12z\hat{y} + 6z}$$

$$\begin{aligned}\frac{\partial \underline{E}(\hat{y},z)}{\partial \hat{y}} &= \frac{(6 + 3z - 16z\hat{y})(12 - 12z\hat{y} + 6z) + (6\hat{y} + 3\hat{y}z - 8z\hat{y}^2)12z}{(12 - 12z\hat{y} + 6z)^2} \\ &= 6\frac{12 + 12z - 32z\hat{y} - 16z^2\hat{y} + 16z^2\hat{y}^2 + 3z^2}{(12 - 12z\hat{y} + 6z)^2}\end{aligned}$$

Therefore

$$\frac{\partial^2 \underline{E}(\hat{y},z)}{(\partial \hat{y})^2} = 6.\frac{(-32z - 16z^2 + 32z^2\hat{y})(12 - 12z\hat{y} + 6z) + 24z(12 + 12z - 32z\hat{y} - 16z^2\hat{y} + 16z^2\hat{y}^2 + 3z^2)}{(12 - 12z\hat{y} + 6z)^3}$$

It is immediate to see that

$$12 - 12z\hat{y} + 6z > 0$$

for all $z < 0$, i.e. $z \in [-2, 0]$, therefore it suffices to examine the numerator of this expression. The numerator can be rewritten as

$$36(-16z - 16z^2 - 4z^3)$$

which is always positive if $z < 0$. I have therefore demonstrated that $\frac{\partial^2 \underline{E}(\hat{y},z)}{(\partial \hat{y})^2} > 0$ if $z < 0$.

Show $\frac{\partial^2 \underline{E}(\hat{y},z)}{\partial \hat{y} \partial z} \leq 0$ if $z < 0$:
Given that

$$\frac{\partial \underline{E}(\hat{y},z)}{\partial \hat{y}} = 6\frac{12 + 12z - 32z\hat{y} - 16z^2\hat{y} + 16z^2\hat{y}^2 + 3z^2}{(12 - 12z\hat{y} + 6z)^2}$$

I can calculate

$$\frac{\partial^2 \underline{E}(\hat{y},z)}{\partial \hat{y} \partial z} = 6\frac{\left[\begin{array}{c}(12 - 32\hat{y} - 32z\hat{y} + 32z\hat{y}^2 + 6z)(12 - 12z\hat{y} + 6z) \\ -2(-12\hat{y} + 6)(12 + 12z - 32z\hat{y} - 16z^2\hat{y} + 16z^2\hat{y}^2 + 3z^2)\end{array}\right]}{(12 - 12z\hat{y} + 6z)^3}$$

Again it suffices to examine the numerator, which can be rewritten as

$$\begin{aligned}&36[(12 - 32\hat{y} - 32z\hat{y} + 32z\hat{y}^2 + 6z)(2 - 2z\hat{y} + z) \\ &-(-4\hat{y} + 2)(12 + 12z - 32z\hat{y} - 16z^2\hat{y} + 16z^2\hat{y}^2 + 3z^2)]\end{aligned}$$

and simplified to

$$36\hat{y}(-16 - 8z)$$

which is always negative if $z \in [-2, 0]$. I have therefore demonstrated that $\frac{\partial^2 \underline{E}(\hat{y},z)}{\partial \hat{y} \partial z} < 0$ if $z < 0$.
As both sufficient conditions hold, we have that $\frac{d\hat{y}^*}{dz} > 0$ if $z < 0$.
Show $\frac{\partial^2 \underline{E}(\hat{y},z)}{(\partial \hat{y})^2} \geq 0$ if $z > 0$:
If $z > 0$ we know the monopolist's optimal cutoff lies above $\frac{y_{\max}}{2}$ and therefore (again setting $y_{\max} = 1$)

$$\underline{E}(\hat{y},z) = \frac{z + 6\hat{y}^2 - 9z\hat{y}^2 + 8z\hat{y}^3}{6z + 12\hat{y} - 18z\hat{y} + 12\hat{y}^2 z}$$

Therefore we have

$$\begin{aligned}\frac{\partial \underline{E}(\hat{y},z)}{\partial \hat{y}} &= \frac{6}{(6z + 12\hat{y} - 18z\hat{y} + 12\hat{y}^2 z)^2}[-2z + 3z^2 + 12\hat{y}^2 + 12zy - 22z^2\hat{y} + 51z^2\hat{y}^2 \\ &\quad -36z\hat{y}^2 - 48z^2\hat{y}^3 + 16z^2\hat{y}^4 + 32z\hat{y}^3]\end{aligned}$$

and

$$\begin{aligned}\frac{\partial^2 \underline{E}(\hat{y},z)}{(\partial \hat{y})^2} &= \frac{36}{(6z + 12\hat{y} - 18z\hat{y} + 12\hat{y}^2 z)^3} \times \\ &\quad [(24\hat{y} + 12z - 22z^2 + 102z^2\hat{y} - 72z\hat{y} - 144z^2\hat{y}^2 + 64z^2\hat{y}^3 + 96z\hat{y}^2)(z + 2\hat{y} - 3z\hat{y} + 2z\hat{y}^2) \\ &\quad +(-4 + 6z - 8z\hat{y})(-2z + 3z^2 + 12\hat{y}^2 + 12zy - 22z^2\hat{y} \\ &\quad +51z^2\hat{y}^2 - 36z\hat{y}^2 - 48z^2\hat{y}^3 + 16z^2\hat{y}^4 + 32z\hat{y}^3)]\end{aligned}$$

which can be shown to be positive $\forall y \in [0.5, 1]$ and $\forall z > 0$.
Show $\frac{\partial^2 \underline{E}(\hat{y},z)}{\partial \hat{y} \partial z} \leq 0$ if $z > 0$:

$$\frac{\partial \underline{E}(\hat{y},z)}{\partial z} = \frac{6(2\hat{y} + 4\hat{y}^4 - 6\hat{y}^2)}{(6z + 12\hat{y} - 18z\hat{y} + 12\hat{y}^2 z)^2}$$

and hence

$$\begin{aligned}\frac{\partial^2 \underline{E}(\hat{y},z)}{\partial z \partial y} &= \frac{36}{(6z + 12\hat{y} - 18z\hat{y} + 12\hat{y}^2 z)^3}[(2 + 16\hat{y}^3 - 12\hat{y})(z + 2\hat{y} - 3z\hat{y} + 2z\hat{y}^2) \\ &\quad -2(2\hat{y} + 4\hat{y}^4 - 6\hat{y}^2)(2 - 3z + 4z\hat{y})]\end{aligned}$$

Unfortunately, this is not always negative. In fact for high $\hat{y}$ it can be seen from simulations that it is positive for all z. The intuition for this is that the shape of the distribution is that of a trough in this case, and as z increases the trough becomes deeper. This means that there is a lot of mass higher up in the income distribution, and as the cutoff moves towards there, average income in the poor group increases due to this. This means that one of the sufficient conditions doesn't hold in the case of $z > 0$, so we need to calculate the whole expression for $\frac{d\hat{y}}{dz}$ to prove that it is positive. Plugging all the derivatives into this expression yields indeed that $\frac{d\hat{y}^*}{dz} > 0$ for all z (numerically calculated - note that the maximum $\hat{y}^*$ is at 0.6427051, when $z = -2$). ■

We already know from Proposition 3.1 that an increase in inequality (resp. z) increases the monopolist's maximized profits. But what happens to total welfare, welfare of the richest and average welfare in the poor group?

Proposition 3.23 *Welfare from sorting at the monopolist's optimal cutoff is increasing in* z *if* $z \in \left[-\frac{2}{y_{\max}}, 0\right]$.

Proof. The derivative of total welfare with respect to z at the monopolist's optimal cutoff $\hat{y}^*$ amounts to

$$\frac{dTW(\hat{y}^*, z)}{dz} = \left(f(\underline{E}^2 - \bar{E}^2) + F2\underline{E}\frac{\partial \underline{E}}{\partial \hat{y}^*} + (1-F)2\bar{E}\frac{\partial \bar{E}}{\partial \hat{y}^*} \right) \frac{d\hat{y}^*}{dz}$$

$$+\frac{\partial F}{\partial z}(\underline{E}^2 - \bar{E}^2) + F2\underline{E}\frac{\partial \underline{E}}{\partial z} + (1-F)2\bar{E}\frac{\partial \bar{E}}{\partial z} - \frac{d\Pi(\hat{y}^*, z)}{dz}$$

where $\Pi(\hat{y}^*, z)$ is the monopolist's maximized profit and we know that

$$\frac{d\Pi(\hat{y}^*, z)}{dz} = -\hat{y}^* \frac{\partial \underline{E}}{\partial z} > 0$$

Hence, the above expression can be simplified to

$$\frac{dTW(\hat{y}^*, z)}{dz} = f(\bar{E} - \underline{E})(\bar{E} + \underline{E} - 2\hat{y}^*)\frac{d\hat{y}^*}{dz} +$$

$$+\frac{\partial F}{\partial z}(\underline{E}^2 - \bar{E}^2) + (F2\underline{E} + \hat{y}^*)\frac{\partial \underline{E}}{\partial z} + (1-F)2\bar{E}\frac{\partial \bar{E}}{\partial z}$$

Note that if $z < 0$ we know that $\hat{y}^* < \frac{y_{\max}}{2}$.
(Set $y_{\max} = 1$ again) We have

$$\bar{E} = \frac{6 - 6\hat{y}^2 - 3z\hat{y}^2 + 8z\hat{y}^3}{12 - 12\hat{y} - 6z\hat{y} + 12z\hat{y}^2}$$

and

$$\frac{\partial \bar{E}}{\partial z} = \frac{6}{(12 - 12\hat{y} - 6z\hat{y} + 12z\hat{y}^2)^2}\hat{y}(6 - 18\hat{y} - 4\hat{y}^3 + 16\hat{y}^2)$$

Furthermore

$$\frac{\partial \underline{E}}{\partial z} = \frac{6}{(12 - 6z + 12z\hat{y})^2}\hat{y}(-4\hat{y}^2)$$

and note that

$$F(\hat{y}, z) = \hat{y} + \frac{z\hat{y}}{2} - z\hat{y}^2$$

and hence

$$\frac{\partial F}{\partial z} = \frac{\hat{y}}{2} - \hat{y}^2 > 0 \quad \forall \hat{y} \in [0, 0.5]$$

Note that $\frac{\partial \underline{E}}{\partial z} < 0$ but $\frac{\partial F}{\partial z} > 0$ because $\hat{y}^* < \frac{y_{\max}}{2}$. As I have shown above that $\frac{d\hat{y}^*}{dz} > 0 \; \forall z$, sufficient conditions for $\frac{dTW(\hat{y}^*,z)}{dz} > 0$ are that

$$\bar{E} + \underline{E} - 2\hat{y}^* > 0 \tag{3.13}$$

and

$$\frac{\partial F}{\partial z}(\underline{E}^2 - \bar{E}^2) + (F2\underline{E} + \hat{y}^*)\frac{\partial \underline{E}}{\partial z} + (1-F)2\bar{E}\frac{\partial \bar{E}}{\partial z} > 0 \tag{3.14}$$

It is easy to check that condition (3.13) always holds in this case. After some algebra, it can be seen from numerical calculations that also (3.14) holds. Hence, $\frac{dTW(\hat{y}^*,z)}{dz} > 0$ if $z < 0$. ■

Proposition 3.24 *Welfare from sorting at the monopolist's optimal cutoff is decreasing in z if* $z \in \left[0, \frac{2}{y_{\max}}\right]$.

Proof. Note that if $z > 0$ we know that $\hat{y} > \frac{y_{\max}}{2}$. As above we have

$$\frac{dTW(\hat{y}^*, z)}{dz} = f(\bar{E} - \underline{E})(\bar{E} + \underline{E} - 2\hat{y}^*)\frac{d\hat{y}^*}{dz} +$$

$$+\frac{\partial F}{\partial z}(\underline{E}^2 - \bar{E}^2) + (F2\underline{E} + \hat{y}^*)\frac{\partial \underline{E}}{\partial z} + (1-F)2\bar{E}\frac{\partial \bar{E}}{\partial z}$$

Sufficient conditions for $\frac{dTW(\hat{y}^*,z)}{dz} < 0$ are that

$$\bar{E} + \underline{E} - 2\hat{y}^* < 0 \tag{3.15}$$

and

$$\frac{\partial F}{\partial z}(\underline{E}^2 - \bar{E}^2) + (F2\underline{E} + \hat{y}^*)\frac{\partial \underline{E}}{\partial z} + (1-F)2\bar{E}\frac{\partial \bar{E}}{\partial z} < 0. \tag{3.16}$$

Note that in this case we have (again setting $y_{\max} = 1$) that

$$\bar{E} = \frac{6 - z - 6\hat{y}^2 + 9z\hat{y}^2 - 8z\hat{y}^3}{12 - 6z - 12\hat{y} + 18z\hat{y} - 12z\hat{y}^2}$$

and

$$\underline{E} = \frac{z + 6\hat{y}^2 - 9z\hat{y}^2 + 8z\hat{y}^3}{6z + 12\hat{y} - 18z\hat{y} + 12\hat{y}^2 z}.$$

Hence

$$\frac{\partial \underline{E}}{\partial z} = \frac{6}{(6z + 12\hat{y} - 18z\hat{y} + 12\hat{y}^2 z)^2}(2\hat{y} + 4\hat{y}^4 - 6\hat{y}^2)$$

and

$$\frac{\partial \bar{E}}{\partial z} = \frac{6}{(12 - 6z - 12\hat{y} + 18z\hat{y} - 12z\hat{y}^2)^2}(4 - 16\hat{y} - 16\hat{y}^3 + 24\hat{y}^2 + 16\hat{y}^4).$$

Furthermore, note that

$$F(\hat{y}, z) = \frac{1}{2}z + \hat{y} - \frac{3z\hat{y}}{2} + z\hat{y}^2$$

and hence

$$\frac{\partial F}{\partial z} = \frac{1}{2} - \frac{3\hat{y}}{2} + \hat{y}^2.$$

Plugging in these expressions, it can easily be shown that (3.15) is always negative. However, concerning (3.16), there is a small range of $\hat{y} > 0.5$ for which this expression is positive. Indeed it can be shown (in numerical simulations) that the whole expression $\frac{dTW(\hat{y}^*,z)}{dz}$ is positive for all z for small $\hat{y}^* > 0.5$. However, note that the monopolist's optimal cutoff increases with z as well, and this increase moves $\hat{y}^*$ out of the area for which total welfare increases with z also for very small z. In fact, for all $z > 0$ it can be shown (again numerically) that $\hat{y}^*$ is greater than the small range of $\hat{y}$ for which $\frac{dTW(\hat{y}^*,z)}{dz}$ would be positive. Hence, total welfare from sorting at the monopolist's optimal cutoff decreases with z if $z > 0$. ■

Proposition 3.25 *Welfare of the richest from sorting at the monopolist's optimum is increasing in z for low rates of inequality and decreasing in z for high rates of inequality.*

Proof.

$$W_{y_{\max}}(\hat{y}^*, z) = (y_{\max} - \hat{y}^*)\bar{E}(\hat{y}^*, z) + \hat{y}^* \underline{E}(\hat{y}^*, z)$$

The monopolist's optimal cutoff satisfies the FOC and hence

$$E - \underline{E}(\hat{y}^*, z) = \hat{y}^* \frac{\partial \underline{E}(\hat{y}^*, z)}{\partial \hat{y}} \tag{3.17}$$

The derivative of $W_{y_{\max}}(\hat{y}^*)$ with respect to z is:

$$\frac{dW_{y_{\max}}(\hat{y}^*, z)}{dz} = (y_{\max} - \hat{y}^*)\left(\frac{\partial \bar{E}}{\partial z}\right) + \hat{y}^* \frac{\partial \underline{E}}{\partial z} + \left[(y_{\max} - \hat{y}^*)\frac{\partial \bar{E}}{\partial \hat{y}} - (\bar{E} - \underline{E}) + \hat{y}^* \frac{\partial \underline{E}}{\partial \hat{y}}\right] \frac{d\hat{y}^*}{dz}$$

Using (3.17) this becomes

$$\frac{dW_{y_{\max}}(\hat{y}^*, z)}{dz} = (y_{\max} - \hat{y}^*)\left(\frac{\partial \bar{E}}{\partial z}\right) + \hat{y}^* \frac{\partial \underline{E}}{\partial z} + \left[(y_{\max} - \hat{y}^*)\frac{\partial \bar{E}}{\partial \hat{y}} - (\bar{E} - E)\right] \frac{d\hat{y}^*}{dz}$$

Hence, sufficient conditions for $W_{y_{\max}}(\hat{y}^*)$ to be decreasing in z are that

$$(y_{\max} - \hat{y}^*)\left(\frac{\partial \bar{E}}{\partial z}\right) + \hat{y}^* \frac{\partial \underline{E}}{\partial z} < 0$$

and

$$\left[(y_{\max} - \hat{y}^*)\frac{\partial \bar{E}}{\partial \hat{y}} - (\bar{E} - E)\right] < 0$$

(because we already know that $\frac{d\hat{y}^*}{dz} > 0$).
For $z > 0$ it can be shown (numerically) that both terms are negative for $\hat{y}^* \in [0.5, 1]$ and all $z \in [0, 0.5]$. Hence, $\frac{dW_{y_{\max}}(\hat{y}^*, z)}{dz}$ is negative for all $\hat{y}$ in $[0.5, 1]$ for all z and thus $W_{y_{\max}}(\hat{y}^*, z)$ is decreasing in z: As inequality increases, welfare of the richest in society from sorting at the monopolist's optimal cutoff goes down.
For $z < 0$ these sufficient conditions don't hold. In fact it can be shown (numerically) that except for very small $z < -1.9$, welfare of the richest in society from sorting at the monopolist's optimal cutoff increases due to an increase in inequality. ■
This last proposition helps in understanding the effect of an increase in inequality on the rich in the presence of sorting: as inequality increases, the monopolist increases the cutoff due to an increase in inequality, because the amount by which she can raise the sorting fee is higher than her loss of "customers" (= members of the rich group, who pay the fee). The increase in the cutoff benefits the rich group, but the increase in the sorting fee harms them. For low rates of inequality, the former effect is higher than the latter, hence welfare of the rich increases with inequality, but if inequality becomes too high (which, because it is in the form of a mean-preserving spread, means that there are more rich people as well as more poor) membership of their exclusive group becomes too expensive and the second effect dominates, leading to a negative relationship between inequality and welfare of the rich.

Proposition 3.26 *Average welfare in the poor group from sorting at the monopolist's optimum decreases in inequality.*

Proof. Average welfare in the poor group amounts to $\underline{E}^2$ (note that they don't have to pay b). We know that

$$\frac{d\underline{E}(\hat{y}^*, z)}{dz} = \frac{\partial \underline{E}}{\partial \hat{y}^*}\frac{d\hat{y}^*}{dz} + \frac{\partial \underline{E}}{\partial z} \tag{3.18}$$

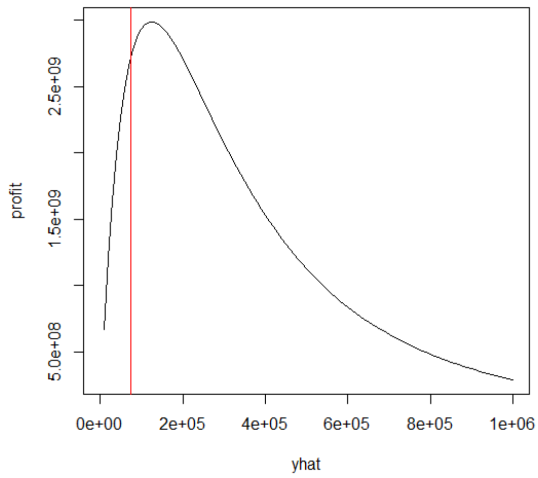

Figure 3.3: Monopolist profit as a function of the cutoff $\hat{y}$ (black) and median (red) if $\mu = 10.85$ and $\sigma = 0.85$

Plugging in the expressions derived above, it is straightforward to show that (3.18) is negative for all z and all $\hat{y}^* > 0.5$. The intuition for this result is that, even though an increase in $\hat{y}$ actually benefits the poor group (because they get to interact with richer people on average), this is not enough to counteract the negative effect of an increasing mass of poor people with zero income in their group. The overall effect of an increase in inequality is thus negative. ■

3.7.5 Lognormal distribution

If we calibrate μ and σ in the lognormal distribution to match the first and second moment of the US household distribution, we get $\mu \approx 10.85$ and $\sigma \approx 0.85$. I will thus often refer to these parameters in this section when comparing monopolist profits and welfare of sorting.

From numerical simulations, it can be concluded that the profit maximization problem of the monopolist always has a unique solution and the optimal cutoff is always increasing in σ. For instance, for $\mu = 10.85$ and $\sigma = 0.85$ the monopolist's profit as a function of the cutoff looks as in Figure 3.3. The vertical line marks the median of the underlying income distribution and therefore demonstrates that the optimal partition for the monopolist is such that the cutoff is above median (in fact it is even above average). If σ declines, the optimal cutoff goes down and eventually will be below median income. If σ increases, the opposite happens: the optimal cutoff increases.

If income is lognormally distributed with $\mu = 10.85$ and $\sigma = 0.85$, welfare as a function of the cutoff looks as in Figure 3.4. The optimal cutoff is above the median (and it can easily be seen that total welfare with sorting at this cutoff is higher than total welfare without sorting, which is the red line in the graph, $E(y)^2$). Note however, that for smaller σ total welfare as a function of the cutoff looks differently. Figure 3.5 shows total welfare as a function of the cutoff for $\sigma = 0.4$. Welfare is first declining in $\hat{y}$ and then increases again until it becomes flat and converges to the welfare of no sorting, $E(y)^2$. Therefore, no sorting is more efficient than sorting. Only once σ increases above 0.65 does the shape change and a unique optimum > 0 appears (see Figure 3.6 for the case where $\sigma = 0.7$). As σ increases further from then on, the welfare-maximizing cutoff increases. As the above analysis shows, monopolist profit-maximization and welfare maximization are not necessarily opposed goals in the case of the lognormal - indeed the optimal cutoffs in both cases are very close to each other and move in

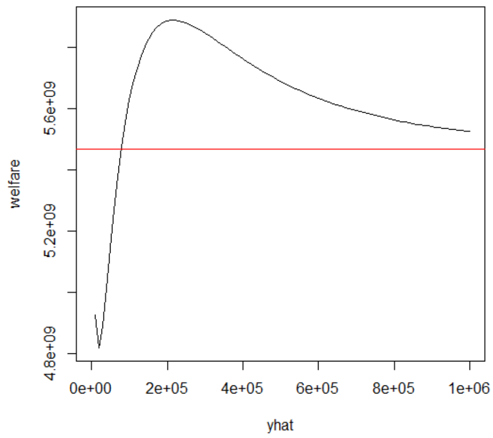

Figure 3.4: Welfare from sorting at cutoff $\hat{y}$ (black) and welfare without sorting (red) if $\mu = 10.85$ and $\sigma = 0.85$

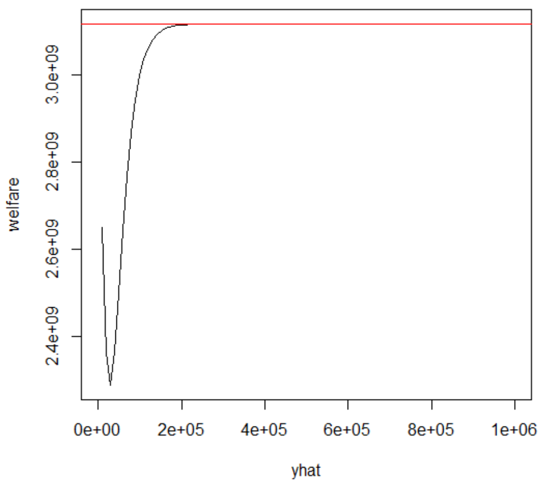

Figure 3.5: Welfare from sorting at cutoff $\hat{y}$ (black) and welfare without sorting (red) if $\mu = 10.85$ and $\sigma = 0.4$

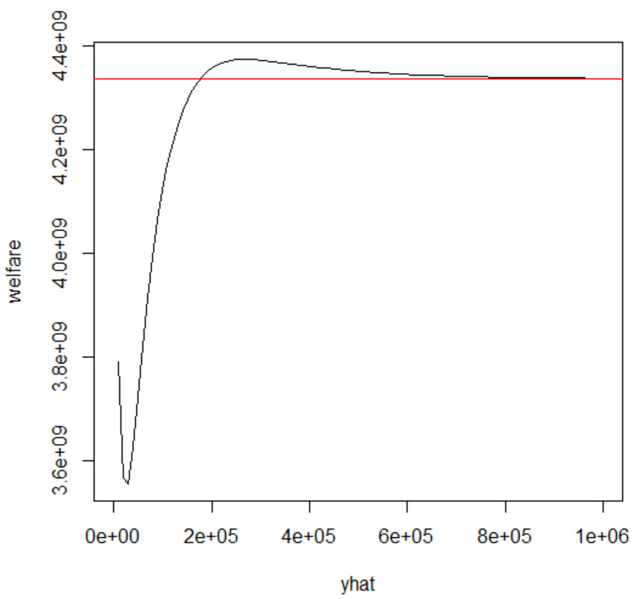

Figure 3.6: Welfare from sorting at cutoff $\hat{y}$ (black) and welfare without sorting (red) if $\mu = 10.85$ and $\sigma = 0.7$

the same direction as inequality increases (in the form of a median-preserving spread) for low rates of σ. However, as can be easily demonstrated in simulations, the goals diverge for very high σ. As σ increases above 8, the monopolist's optimal cutoff becomes much higher than the optimal cutoff for welfare, and welfare at the optimal cutoff starts to decline.

3.7.6 Proof that for the atom distribution no sorting is more efficient than perfect sorting, i.e. that it has $CV \leq 1$

Welfare from perfect sorting is given by

$$\frac{E(y^2)}{2}.$$

We can calculate that

$$E(y^2) = \int_0^1 y^2 f(y) dy + z = \int_0^1 y^2 (1 - 2z) dy + z = \frac{1}{3} + \frac{z}{3}$$

Therefore we see that, as described in Section 3.5, $E(y^2)$ (the total surplus of perfect sorting) and welfare of perfect sorting (which is just half of it) are increasing in inequality z. However, welfare of perfect sorting is smaller than welfare of no sorting for all z:

$$\frac{E(y^2)}{2} = \frac{1}{6} + \frac{z}{6} \leq \frac{1}{4} \iff z \leq 0.5$$

Another way to see this is to calculate the coefficient of variation:

$$CV = \frac{\sqrt{Var(y)}}{E(y)} = \frac{\sqrt{\frac{1}{3} + \frac{z}{3} - \frac{1}{4}}}{\frac{1}{2}} = 2\sqrt{\frac{1}{12} + \frac{z}{3}}$$

It is straightforward to see that $CV \leq 1$ $\forall z \in [0, 0.5]$ and that it reaches its maximum of 1 where $z = 0.5$. If $z = 0.5$, perfect sorting would yield the same welfare than no sorting if the sorting fee is set at $\frac{1}{2}$, such that the total surplus is split in half. However, the sorting fee is not uniquely determined in this case and a profit-maximizing monopolist would set it as high as possible, which would be 1 in this case, such that total welfare is 0 and the monopolist gets all the surplus from sorting (which is 0.5) for herself.

3.7.7 Proof that the house distribution is NBUE

In order to prove that the house distribution is NBUE, I need to show that

$$\bar{E} - E - \hat{y} < 0 \quad \forall \hat{y}, \ \ \forall z \in [-2, 2].$$

If $\hat{y} \leq 0.5$ we have that

$$\bar{E} - E - \hat{y} = \frac{6 - 6\hat{y}^2 - 3z\hat{y}^2 + 8z\hat{y}^3}{12 - 12\hat{y} - 6z\hat{y} + 12z\hat{y}^2} - \frac{1}{2} - \hat{y}$$

$$= \frac{-6\hat{y} + 6\hat{y}^2 + z(-4\hat{y}^3 + 3\hat{y} - 3\hat{y}^2)}{12 - 12\hat{y} - 6z\hat{y} + 12z\hat{y}^2}.$$

The denominator is always positive, so we just need to analyze the numerator: $-6\hat{y} + 6\hat{y}^2$ is always negative, and $-4\hat{y}^3 + 3\hat{y} - 3\hat{y}^2$ is positive for $\hat{y} \leq 0.5$, hence if z is negative, the whole expression is negative for sure. If z is positive, then the numerator reaches its maximum at $z = 2$, where it becomes $-6\hat{y} + 6\hat{y}^2 - 8\hat{y}^3 + 6\hat{y} - 6\hat{y}^2 = -8\hat{y}^3$ which is always negative. Hence, $\bar{E} - E - \hat{y} < 0$ if $\hat{y} \leq 0.5$.
If $\hat{y} \geq 0.5$ we have that

$$\bar{E} - E - \hat{y} = \frac{6 - z - 6\hat{y}^2 + 9z\hat{y}^2 - 8z\hat{y}^3}{12 - 6z - 12\hat{y} + 18z\hat{y} - 12z\hat{y}^2} - \frac{1}{2} - \hat{y}$$

$$= \frac{-6\hat{y} + 6\hat{y}^2 + z(2 + 4\hat{y}^3 - 3\hat{y} - 3\hat{y}^2)}{12 - 6z - 12\hat{y} + 18z\hat{y} - 12z\hat{y}^2}.$$

The denominator is again positive, and the first term of the numerator, $-6\hat{y} + 6\hat{y}^2$ is always negative. $2 + 4\hat{y}^3 - 3\hat{y} - 3\hat{y}^2$ reaches its minimum at $\frac{1}{4} + \sqrt{\frac{5}{16}}$ where it is negative, and hence $z(2 + 4\hat{y}^3 - 3\hat{y} - 3\hat{y}^2)$ is positive if $z < 0$, and maximal at $z = -2$. Combined with $-6\hat{y} + 6y^2$ evaluated at $\frac{1}{4} + \sqrt{\frac{5}{16}}$ the total expression is negative. $2 + 4\hat{y}^3 - 3\hat{y} - 3\hat{y}^2$ reaches its maximum at 0.5 where it is positive and hence $z(2 + 4\hat{y}^3 - 3\hat{y} - 3\hat{y}^2)$ is maximal at $z = 2$. Again combined with $-6\hat{y} + 6y^2$ evaluated at 0.5 the whole expression is negative. Hence $\bar{E} - E - \hat{y} < 0$ if $\hat{y} > 0.5$, and thus the house distribution is NBUE for all z.

3.7.8 Proof that for the house distribution no sorting is more efficient than perfect sorting, i.e. that it has $CV \leq 1$

Total surplus from perfect sorting is

$$E(y^2) = \int_0^{0.5} y^2(1 + \frac{z}{2} - 2zy)dy + \int_{0.5}^{1} y^2(1 - \frac{3z}{2} + 2zy)dy$$

$$= \frac{1}{3} + \frac{z}{48}$$

Hence, welfare from perfect sorting is

$$\frac{1}{6} + \frac{z}{96}$$

and no sorting yields higher welfare than perfect sorting iff

$$\frac{1}{6} + \frac{z}{96} < \frac{1}{4} \iff 24 > 16 + z$$

which is satisfied for all $z \in [-2, 2]$.
Equivalently, the coefficient of variation is

$$CV = \frac{\sqrt{\frac{1}{3} + \frac{z}{48} - \frac{1}{4}}}{\frac{1}{2}} = 2\sqrt{\frac{1}{12} + \frac{z}{48}}$$

which is strictly smaller than 1 for all $z \in [-2, 2]$.

Bibliography

[1] Abdulkadiroğlu, A., Angrist, J., & Pathak, P. (2014). The elite illusion: Achievement effects at Boston and New York exam schools. Econometrica, 82(1), 137-196.

[2] Alesina, A., & Angeletos, G. M. (2005). Fairness and Redistribution. American Economic Review, 960-980.

[3] Algan, Y., Do, Q. A., Dalvit, N., Le Chapelain, A., & Zenou, Y. (2015). How Social Networks Shape Our Beliefs: A Natural Experiment among Future French Politicians. (mimeo)

[4] Andreasch, M., Fessler, P., Mooslechner, P., Schürz, M. (2013) Fakten zur Vermögensverteilung in Österreich, BMASK (Ed.), Sozialbericht 2011-2012 (Vienna 2013) 247-265.

[5] Ashok, V., Kuziemko, I., & Washington, E. (2015). Support for Redistribution in an Age of Rising Inequality:: New stylized facts and some tentative explanations. Brookings papers on economic activity, (1), 2.

[6] Bagwell, L. S., & Bernheim, B. D. (1996). Veblen effects in a theory of conspicuous consumption. The American Economic Review, 349-373.

[7] Bartels, L. M. (2009). Economic inequality and political representation. The unsustainable American state, 167-96.

[8] Becker, G. S. (1974). A theory of marriage. In Economics of the family: Marriage, children, and human capital (pp. 299-351). University of Chicago Press.

[9] Benabou, R., & Ok, E. A. (2001). Social Mobility and the Demand for Redistribution: The Poum Hypothesis. The Quarterly Journal of Economics, 116(2), 447-487.

[10] Bishop, B. (2008). The big sort. Why the Clustering of America is Tearing Us Apart. NY: Houghton Mifflin.

[11] Boisjoly, J., Duncan, G. J., Kremer, M., Levy, D. M., & Eccles, J. (2006). Empathy or antipathy? The impact of diversity. The American economic review, 96(5), 1890-1905.

[12] Bonica, A., McCarty, N., Poole, K. T., & Rosenthal, H. (2013). Why hasn't democracy slowed rising inequality?. The Journal of Economic Perspectives, 27(3), 103-123.

[13] Borge, L., & Rattsoe, E. J. (2004). Income Distribution and Tax Structure: Empirical Test of the Meltzer Richard Hypothesis. European Economic Review, 48, 805-826.

[14] Burns, J., Corno, L., & La Ferrara, E. (2013). Does interaction affect racial prejudice and cooperation? Evidence from randomly assigned peers in South Africa. (mimeo)

[15] Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States (No. w19843).

[16] Corneo, G., & Grüner, H. P. (2000). Social limits to redistribution. The American Economic Review, 90(5), 1491-1507.

[17] Cowell, F. A. (2000). Measurement of inequality. Handbook of income distribution, 1, 87-166.

[18] Cowell, F. (2011). Measuring inequality. OUP Oxford.

[19] Cruces, G., Perez-Truglia, R., & Tetaz, M. (2013). Biased perceptions of income distribution and preferences for redistribution: Evidence from a survey experiment. Journal of Public Economics, 98, 100-112.

[20] Dalton, H. (1920). The measurement of the inequality of incomes. The Economic Journal, 30(119), 348-361.

[21] Damiano, E., & Li, H. (2007). Price discrimination and efficient matching. Economic Theory, 30(2), 243-263.

[22] Dustmann, C., & Preston, I. (2001). Attitudes to ethnic minorities, ethnic context and location decisions. The Economic Journal, 111(470), 353-373.

[23] Epple, D., & Romano, R. E. (1998). Competition between private and public schools, vouchers, and peer-group effects. American Economic Review, 33-62.

[24] Fernandez, R., & Rogerson, R. (2003). Equity and resources: An analysis of education finance systems. Journal of Political Economy, 111(4), 858-897.

[25] Evans, G. & Tilley, J. (2011). Private schools and public divisions: The influence of private schooling on social attitudes and political choices. in Alison Park et al (eds.), British Social Attitudes: The 28th Report. London: Sage.

[26] Forman, B., & Koch, C. (2012). Geographic Segregation: The Role of Income Inequality. Communities and Banking.

[27] Georgiadis, A., & Manning, A. (2012). Spend it like Beckham? Inequality and redistribution in the UK, 1983–2004. Public choice, 151(3-4), 537-563.

[28] Golub, B., & Jackson, M. O. (2012). How Homophily Affects the Speed of Learning and Best-Response Dynamics. The Quarterly Journal of Economics, 127(3), 1287-1338.

[29] Gouveia, M., & Masia, N. A. (1998). Does the median voter model explain the size of government?: Evidence from the states. Public Choice, 97(1-2), 159-177.

[30] Hoppe, H. C., Moldovanu, B., & Sela, A. (2009). The theory of assortative matching based on costly signals. The Review of Economic Studies, 76(1), 253-281.

[31] Karadja, M., Mollerstrom, J., & Seim, D. (2014). Richer (and Holier) than Thou? The Effect of Relative Income Improvements on Demand for Redistribution (No. 1050). George Mason University, Interdisciplinary Center for Economic Science.

[32] Kenworthy, L., & McCall, L. (2008). Inequality, public opinion and redistribution. Socio-Economic Review, 6(1), 35-68.

[33] Kiatpongsan, S., & Norton, M. I. (2014). How much (more) should CEOs make? A universal desire for more equal pay. Perspectives on Psychological Science, 9(6), 587-593.

[34] Kuziemko, I., Norton, M. I., Saez, E., & Stantcheva, S. (2015). How elastic are preferences for redistribution? Evidence from randomized survey experiments. The American Economic Review, 105(4), 1478-1508.

[35] Larcinese, V. (2005). Electoral competition and redistribution with rationally informed voters. Contributions in Economic Analysis & Policy, 4(1).

[36] Levy, G. and Razin, R. (2016). The Coevolution of Segregation, Polarised Beliefs and Discrimination: The Case of Private vs. State Education (mimeo)

[37] Levy, G., & Razin, R. (2015). Preferences over equality in the presence of costly income sorting. American Economic Journal: Microeconomics, 7(2), 308-337.

[38] Matakos, K. & Xefteris, D. (2016) Divide and Rule: Redistribution in a Model with Differentiated Candidates. forthcoming in Economic Theory

[39] Meltzer, A. H., & Richard, S. F. (1981). A rational theory of the size of government. The Journal of Political Economy, 914-927.

[40] Meltzer, A. H., & Richard, S. F. (1983). Tests of a rational theory of the size of government. Public Choice, 41(3), 403-418.

[41] Milanovic, B. (2000). The median-voter hypothesis, income inequality, and income redistribution: an empirical test with the required data. European Journal of Political Economy, 16(3), 367-410.

[42] Norton, M. I., & Ariely, D. (2011). Building a better America—One wealth quintile at a time. Perspectives on Psychological Science, 6(1), 9-12.

[43] Norton, M. I., Neal, D. T., Govan, C. L., Ariely, D., & Holland, E. (2014). The Not-So-Common-Wealth of Australia: Evidence for a Cross-Cultural Desire for a More Equal Distribution of Wealth. Analyses of Social Issues and Public Policy, 14(1), 339-351.

[44] Pesendorfer, W. (1995). Design innovation and fashion cycles. The American Economic Review, 771-792.

[45] Piketty, T. (1995). Social mobility and redistributive politics. The Quarterly journal of economics, 551-584.

[46] Piketty, T. (2014). Capital in the Twenty-first Century. Cambridge Massachusetts: Belknap of Harvard University Press

[47] Piketty, T., Saez, E., & Stantcheva, S. (2014). Optimal taxation of top labor incomes: A tale of three elasticities. American Economic Journal: Economic Policy, 6(1), 230-271.

[48] Rayo, L. (2013). Monopolistic Signal Provision. The BE Journal of Theoretical Economics, 13(1), 27-58.

[49] Reardon, S. F., & Bischoff, K. (2011). Income Inequality and Income Segregation. American Journal of Sociology, 116(4), 1092-1153.

[50] Rodriguez, F. (1999). Inequality, redistribution and rent-seeking. Department of Economics, University of Maryland, 1999b.(mimeo).

[51] Rothschild, M. & Stiglitz, J. (1970). Increasing Risk: I. A Definition. Journal of Economic Theory, 2, 225-243.

[52] Saez, E., & Zucman, G. (2016). Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data. The Quarterly Journal of Economics, 131(2), 519-578.

[53] Savage, M. (2015). Social class in the 21st century. Penguin UK.

[54] Scervini, F. (2012). Empirics of the median voter: democracy, redistribution and the role of the middle class. The Journal of Economic Inequality, 10(4), 529-550.

[55] Veblen, T. (1899). The theory of the leisure class: An economic study in the evolution of institutions. Macmillan.

[56] Watson, T. (2009). Inequality and the measurement of residential segregation by income in American neighborhoods. Review of Income and Wealth, 55(3), 820-844.

[57] Wenglinsky, H. (2007). Are private high schools better academically than public high schools?[electronic resource]. Center on Education Policy, 2007.

[58] Windsteiger, L. (2017a). The Redistributive Consequences of Segregation.

[59] Windsteiger, L. (2017b). Sorting in the Presence of Misperceptions.

[60] Windsteiger, L. (2017c). Monopolistic Supply of Sorting, Inequality and Welfare.