Sorting equilibria with misperceptions

May 1, 2015

Abstract

Models where people are provided with a technology to sort into two different groups according to income, i.e. a sorting fee b, have the well-known feature of multiplicity of equilibria. Specifically, for any sorting fee b there will exist a monotonic partition of the set of incomes in the economy (characterized by a cutoff-income $\hat{y}$) which represents a binary sorting equilibrium. I examine what happens if, once people are sorted into the two groups, their perception of the average income in the other group (and hence of the overall income distribution) changes, due to anchoring inadjustment or similar perceptional biases.

I examine the existence and uniqueness properties of biased binary sorting equilibria, i.e. partitions of the set of incomes such that people, given their perceptional bias of the average income in the other group, prefer staying in their group rather than changing to the other group. As a further refinement, I introduce a consistency requirement, which states that people in one group should not think that people who are not joining their group are making a mistake. This additional requirement yields uniqueness for a large class of utility functions.

Furthermore, I examine how the direction of the bias affects existence and uniqueness of binary biased sorting equilibria and profits of a monopolist providing the sorting technology.

PRELIMINARY DRAFT, PLEASE DO NOT CITE!

1 A theoretical model of economic segregation

Let income in an economy be distributed according to $F(y)$, and suppose y is distributed continuously on $Y = [0, y_{\max}]$ where $y_{\max} < \infty$. Assume furthermore that $F(y) \in C([0, y_{\max}])$ and strictly monotonic. Suppose a person with income y_i gets utility $U_i = y_i E(y_i | y_i \in S_i)$, where S_i is individual i's reference group. Suppose that if there is no economic segregation, everybody's reference group is a representative sample of the whole population, such that $U_i = y_i E(y)$. However, a person with income y_i can pay b to join club S_b and get utility

$$U(y_i, E[y|y \in S_b]) - b$$

or refrain from paying b and get

$$U(y_i, E[y|y \notin S_b])$$

where S_b is the set of incomes y_i of people who have paid b. Let $U(.,.)$ be continuous, strictly increasing in both arguments and strictly supermodular, so

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

Then I can define the following:

Definition 1 *A binary sorting equilibrium is a subset S_b of Y and a sorting fee b such that all people in S_b pay b and interact only with other people in S_b and people not in S_b pay nothing and interact only with people $\notin S_b$, where b is such that*

$$U(y_i, E[y|y \in S_b]) - U(y_i, E[y|y \notin S_b]) \geq b \quad \forall y_i \in S_b \tag{IR1}$$

$$U(y_i, E[y|y \in S_b]) - U(y_i, E[y|y \notin S_b]) \leq b \quad \forall y_i \notin S_b \tag{IR2}$$

Theorem 2 *If $U(.,.)$ is increasing in both arguments and strictly supermodular then there can only be monotone sorting equilibria if $b \neq 0$.*

Proof. Suppose w.l.o.g. $b > 0$, $y_1 \in S_b$ and $y_2 \notin S_b$ but $y_2 > y_1$. Then I must have

$$U(y_1, E[y|y \in S_b]) - U(y_1, E[y|y \notin S_b]) \geq b \tag{IR1}$$

$$\Rightarrow U(y_1, E[y|y \in S_b]) > U(y_1, E[y|y \notin S_b])$$

$$\Rightarrow E[y|y \in S_b] > E[y|y \notin S_b] \tag{1}$$

and

$$U(y_2, E[y|y \in S_b]) - U(y_2, E[y|y \notin S_b]) \leq b \tag{IR2}$$

(IR1) and (IR2) imply

$$\begin{aligned} \Longrightarrow \quad & U(y_1, E[y|y \in S_b]) - U(y_1, E[y|y \notin S_b]) \\ \geq \quad & U(y_2, E[y|y \in S_b]) - U(y_2, E[y|y \notin S_b]) \end{aligned}$$

But together with (1) this is a contradition to U being strictly supermodular.

■

Hence, I can rewrite the definition of a binary sorting equilibrium in terms of a cutoff $\hat{y}$:

Definition 3 *A binary sorting equilibrium is a cutoff $\hat{y}$ and corresponding sorting fee b such that all people with income $\leq \hat{y}$ interact only with other people below $\hat{y}$ and people above $\hat{y}$ pay sorting fee b and interact only with people above $\hat{y}$, where*

$$U(\hat{y}, [E[y|y \geq \hat{y}]) - U(\hat{y}, E[y|y \leq \hat{y}]) = b$$

For any $\hat{y}$ there is a (unique) b such that people with y_i above $\hat{y}$ want to join S_b, people below don't. Equivalently, for any $b \in [U(0, E(y)) - U(0,0), U(y_{\max}, y_{\max}) - U(y_{\max}, E)]$ there is a cutoff $\hat{y} \geq 0$ (not necessarily unique) such that people with y_i above $\hat{y}$ want to join S_b, people below don't. If $U(\hat{y}, [E[y|y \geq \hat{y}]) - U(\hat{y}, E[y|y \leq \hat{y}])$ is strictly increasing in $\hat{y}$ then $\hat{y}(b)$ is unique.

1.1 Sorting with misperceptions

Now suppose that people, once they are in their group, i.e. either inside or outside S_b, become biased about the overall income distribution and hence about the conditional expectation of y in the other group. For a start, let me just assume that people in group S_b perceive the average in their group correctly, but they have a wrong perception about the average in the other group, $E_g[y|y \notin S_b] \neq E[y|y \notin S_b]$ and the people not in group S_b perceive the average in their group correctly, but they misperceive the average in group S_b, $E_{ng}[y|y \in S_b] \neq E[y|y \in S_b]$. Then I can define the following:

Definition 4 *A* ***binary biased sorting equilibrium*** *is a subset S_b of Y, a sorting fee b, and biased perceptions $E_g[y|y \notin S_b]$ and $E_{ng}[y|y \in S_b]$ satisfying*

$$U(y_i, E[y|y \in S_b]) - U(y_i, E_g[y|y \notin S_b]) \geq b \quad \forall y_i \in S_b \tag{IR1}$$

$$U(y_i, E_{ng}[y|y \in S_b]) - U(y_i, E[y|y \notin S_b]) \leq b \quad \forall y_i \notin S_b \tag{IR2}$$

A ***binary biased sorting equilibrium with consistency*** *additionally satisfies*

$$U(y_i, E[y|y \in S_b]) - U(y_i, E_g[y|y \notin S_b]) \leq b \quad \forall y_i \notin S_b \tag{CR1}$$

$$U(y_i, E_{ng}[y|y \in S_b]) - U(y_i, E[y|y \notin S_b]) \geq b \quad \forall y_i \in S_b \tag{CR2}$$

The last two inequalities constitute what I call the *consistency requirement*: The people who are in S_b think that the people who do not join are correct not to do so, while the people not in S_b think the people in S_b are correct in joining.

Theorem 5 *If U is increasing in both arguments and strictly supermodular, only monotone biased sorting equilibria with consistency are possible if $b \neq 0$.*

Proof. Suppose w.l.o.g. $b > 0$, $y_1 \in S_b$ and $y_2 \notin S_b$ but $y_2 > y_1$. Then

$$\begin{aligned} U(y_2, E_{ng}[y|y &\in S_b]) - U(y_2, E[y|y \notin S_b]) \leq b && \text{(IR2)} \\ U(y_1, E[y|y &\in S_b]) - U(y_1, E_g[y|y \notin S_b]) \geq b && \text{(IR1)} \\ &\Longrightarrow E[y|y \in S_b] > E_g[y|y \notin S_b] \end{aligned}$$

$$\begin{aligned} U(y_2, E[y|y &\in S_b]) - U(y_2, E_g[y|y \notin S_b]) \leq b && \text{(CR1)} \\ U(y_1, E_{ng}[y|y &\in S_b]) - U(y_1, E[y|y \notin S_b]) \geq b && \text{(CR2)} \\ &\Longrightarrow E_{ng}[y|y \in S_b] > E[y|y \notin S_b] \end{aligned}$$

But as $y_2 > y_1$, both pairs of equations cannot hold at the same time if U is strictly supermodular! ■

Hence, also with bias, the consistency requirement ensures that there cannot be other equilibria than the cutoff, two-group case. What is also necessary for this argument is my definition of the bias: the bias[1] of a person does not depend on the person's own income directly, just on which group people are in, so the bias is constant across groups. Note that without the consistency requirement, non-monotonic sorting equilibria could be possible.

Having established this, I can rewrite the definition of a biased sorting equilibrium with consistency in terms of a cutoff $\hat{y}$ and a corresponding misperception of the other group's average:

Definition 6 *A* ***biased sorting equilibrium with consistency*** *is a cutoff $\hat{y}^* \in Y$, a sorting fee b and biases $E_g[y|y \leq \hat{y}]$ and $E_{ng}[y|y \geq \hat{y}]$ such that*

$$U(y, E[y|y \geq \hat{y}^*]) - U(y, E_g[y|y \leq \hat{y}^*] \geq b \quad \forall y > \hat{y} \qquad \text{(IR1)}$$

$$U(y, E_{ng}[y|y \geq \hat{y}^*]) - U(y, E[y|y \leq \hat{y}^*]) \leq b \quad \forall y \leq \hat{y} \qquad \text{(IR2)}$$

$$\begin{aligned} U(y, E[y|y &\geq \hat{y}^*]) - U(y, E_g[y|y \leq \hat{y}^*]) \leq b \quad \forall y > \hat{y} && \textit{(CR1)} \\ U(y, E_{ng}[y|y &\geq \hat{y}^*]) - U(y, E[y|y \leq \hat{y}^*]) \geq b \quad \forall y \leq \hat{y} && \textit{(CR2)} \end{aligned}$$

As long as $E[y|y \geq \hat{y}^*] > E_g[y|y \leq \hat{y}^*]$ and $E_{ng}[y|y \geq \hat{y}^*] > E[y|y \leq \hat{y}^*]$ (which is a restriction on the biases that is quite sensible and that I want to assume from here on) these inequalities can be simplified to

$$U(\hat{y}^*, E[y|y \geq \hat{y}^*]) - U(\hat{y}^*, E_g[y|y \leq \hat{y}^*] \geq b \qquad \text{(IR1)}$$

$$U(\hat{y}^*, E_{ng}[y|y \geq \hat{y}^*]) - U(\hat{y}^*, E[y|y \leq \hat{y}^*]) \leq b \qquad \text{(IR2)}$$

$$\begin{aligned} U(\hat{y}^*, E[y|y &\geq \hat{y}^*]) - U(\hat{y}^*, E_g[y|y \leq \hat{y}^*]) \leq b && \textit{(CR1)} \\ U(\hat{y}^*, E_{ng}[y|y &\geq \hat{y}^*]) - U(\hat{y}^*, E[y|y \leq \hat{y}^*]) \geq b && \textit{(CR2)} \end{aligned}$$

$$\begin{aligned} \Rightarrow \quad & U(\hat{y}^*, E_{ng}[y|y \geq \hat{y}^*]) - U(\hat{y}^*, E[y|y \leq \hat{y}^*]) \\ = \quad & U(\hat{y}^*, E[y|y \geq \hat{y}^*]) - U(\hat{y}^*, E_g[y|y \leq \hat{y}^*]) = b \end{aligned}$$

[1] When I write 'bias' here, I mean in fact people's wrong perception of the average income in the other group.

Hence, a biased sorting equilibrium with consistency is at a point $\hat{y}^*$ where the perceived benefits of sorting of the rich are equal to the perceived benefits of sorting of the poor.

Now to say more, e.g. whether a biased sorting equilibrium exists or whether it is unique, I need to make further assumptions on the bias. In general, the bias could take on any form, but I want to look at a particular kind of bias: As people live in their segregated communities, they see mostly people who have income y similar to their own (i.e. people from their own group). They do meet people from the other group, but not that often. 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, they suffer from some form of "anchoring inadjustment", in the sense that the poor think that the average in the rich 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** the average y in the high group and people above the cutoff **overestimate** the average y in the low group. Note that this will lead to both groups underestimating 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.

Formally, the bias is modelled as follows:

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

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

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

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

$\beta, \gamma \in [0, 1]$ parameterizes the "naivity" of agents and if β resp. γ is 0 agents have no misperceptions. Furthermore, I assume that the misperceptions (the anchoring inadjustment) 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).[2] It is straightforward to see that $\bar{E}_{ng} \leq \bar{E}$ and $\underline{E}_g > \underline{E}$.

Now having defined the biases in this way, I can examine whether such a biased sorting equilibrium will always exist. Remember, the sorting equilibrium is at $\hat{y}^*$ where

$$\begin{aligned} U(\hat{y}^*, E[y|y &\geq \hat{y}^*]) - U(\hat{y}^*, E_g[y|y \leq \hat{y}^*]) = \\ U(\hat{y}^*, E_{ng}[y|y &\geq \hat{y}^*]) - U(\hat{y}^*, E[y|y \leq \hat{y}^*])(= b) \end{aligned}$$

[2] This is also a sensible assumption when you consider the fact that this implies that the rich are correct at $\hat{y} = 0$ and the poor are correct at $\hat{y} = y_{\max}$, which seems to be reasonable given that at $\hat{y} = 0$ the rich basically know that everyone in the poor group has income $\hat{y} = 0$ and hence it would be very strange to assume that they misperceive the average in the poor group.

1.1.1 Existence

Will such a cutoff $\hat{y}^*$ at which the perceived benefits of sorting for the rich and the poor are equal always exist?

Let me look at the limits of $U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E}_g)$ and $U(\hat{y}, \bar{E}_{ng}) - U(\hat{y}, \underline{E})$ as $\hat{y}$ goes to 0 and $y_{\max}$. For notational simplicity I will write F instead of $F(\hat{y})$ here.

As $\hat{y} \to 0$, clearly $\bar{E} \to E$ and $\underline{E} \to 0$, and the poor perceived benefits of sorting converge to

$$\lim_{\hat{y}\to 0} \bar{E}_{ng} = \lim_{\hat{y}\to 0} (\beta(1 - F(\hat{y}))\hat{y} + (1 - \beta(1 - F(\hat{y}))\bar{E})$$

$$= \beta.0 + (1 - \beta)\bar{E} = (1 - \beta)E < \lim_{\hat{y}\to 0} \bar{E} = E$$

The rich perceived benefits of sorting become

$$\lim_{\hat{y}\to 0} \underline{E}_g = \lim_{\hat{y}\to 0} (\gamma F(\hat{y})\hat{y} + (1 - \gamma F(\hat{y}))\underline{E}) = 0 = \lim_{\hat{y}\to 0} \underline{E}$$

As $\hat{y} \to y_{\max}$, clearly $\bar{E} \to y_{\max}$ and $\underline{E} \to E$, and the poor perceived benefits of sorting converge to

$$\lim_{\hat{y}\to y_{\max}} \bar{E}_{ng} = \lim_{\hat{y}\to y_{\max}} (\beta(1 - F(\hat{y}))\hat{y} + (1 - \beta(1 - F(\hat{y}))\bar{E})$$

$$= y_{\max} = \lim_{\hat{y}\to y_{\max}} \bar{E}$$

The rich perceived benefits of sorting become

$$\lim_{\hat{y}\to y_{\max}} \underline{E}_g = \lim_{\hat{y}\to y_{\max}} (\gamma F(\hat{y})\hat{y} + (1 - \gamma F(\hat{y}))\underline{E})$$

$$= \gamma y_{\max} + (1 - \gamma)E < \lim_{\hat{y}\to y_{\max}} \underline{E} = E$$

To summarize, I find that $U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E}_g)$ and $U(\hat{y}, \bar{E}_{ng}) - U(\hat{y}, \underline{E})$ are always $\leq U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E})$. For $\hat{y} \to 0$, I have that $U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E}_g) \to U(0, E) - U(0, 0)$ and $U(\hat{y}, \bar{E}_{ng}) - U(\hat{y}, \underline{E}) \to U(0, (1-\beta)E) - U(0, 0)$, where the latter is strictly smaller than the former. Hence, the perceived benefits of the rich are larger than the perceived benefits of the poor as $\hat{y}$ goes to 0. For $\hat{y} \to y_{\max}$ I have that $U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E}_g) \to U(y_{\max}, y_{\max}) - U(y_{\max}, \gamma y_{\max} + (1-\gamma)E)$ and $U(\hat{y}, \bar{E}_{ng}) - U(\hat{y}, \underline{E}) \to U(y_{\max}, y_{\max}) - U(y_{\max}, E)$, where the latter is strictly larger than the former. Hence, the poor perceived benefits are larger than the rich ones as $\hat{y}$ goes to $y_{\max}$. As all functions here are continuous, there must be a $\hat{y}^*$ in $]0, y_{\max}[$ such that the perceived benefits of sorting of the rich are exactly equal to the perceived benefits of the poor, $U(\hat{y}^*, \bar{E}) - U(\hat{y}^*, \underline{E}_g) = U(\hat{y}^*, \bar{E}_{ng}) - U(\hat{y}^*, \underline{E})$.

Theorem 7 *If $U(.,.)$ is strictly increasing in both arguments and supermodular, a biased sorting equilibrium with consistency and bias $\underline{E}_g$ and $\bar{E}_{ng}$ as defined always exists.*

Remark 8 *To guarantee existance of a biased binary sorting equilibrium with consistency, the bias obviously need not be defined in exactly the way I have defined it. Sufficient conditions on the bias* $\underline{E}_g$ *and* $\bar{E}_{ng}$ *are that both* $\underline{E}_g$ *and* $\bar{E}_{ng}$ *are continuous functions of* $\hat{y}$ *and that* $\lim_{\hat{y}\to y_{\max}} \bar{E}_{ng} = \lim_{\hat{y}\to y_{\max}} \bar{E}$ *and* $\lim_{\hat{y}\to 0} \underline{E}_g = \lim_{\hat{y}\to 0} \underline{E}$ *and* $\bar{E}_{ng} < \bar{E}$ *for all* $\hat{y} \neq y_{\max}$ *and* $\underline{E}_g > \underline{E}$ *for all* $\hat{y} \neq 0$ *(in fact the latter two assumptions can be further relaxed to* $\lim_{\hat{y}\to 0} \bar{E}_{ng} < \lim_{\hat{y}\to 0} \bar{E}$ *and* $\lim_{\hat{y}\to y_{\max}} \underline{E}_g > \lim_{\hat{y}\to y_{\max}} \underline{E}$*)*

From the observations on the limits of the differences for $\hat{y} \to 0$ I can also conclude the following:

Theorem 9 *If* $U(.,.)$ *is strictly increasing in both arguments and supermodular, any non-degenerate binary biased sorting equilibrium with consistency and bias* $\underline{E}_g$ *and* $\bar{E}_{ng}$ *is always bounded away from zero.*

Note that of course there always exists the "degenerate" sorting equilibrium at zero, where b and $\hat{y}$ are zero and everybody is in the same group, with no bias.

1.1.2 Uniqueness of equilibrium

A sufficient condition for uniqueness of the biased sorting equilibrium with consistency is the condition that the biases should be such that the difference between the correct benefits of sorting and the perceived benefits of sorting for the rich, $U(\hat{y}, \underline{E}_g) - U(\hat{y}, \underline{E})$, monotonically increases, while the difference between the correct benefits of sorting and the perceived ones of the poor, $U(\hat{y}, \bar{E}) - U(\hat{y}, \bar{E}_{ng})$, monotonically decreases.

To find a necessary condition I look at the functions $U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E}_g)$ and $U(\hat{y}, \bar{E}_{ng}) - U(\hat{y}, \underline{E})$. Do they cut only once? To ensure that they cut only once, I need to have a single crossing condition satisfied: At any $\hat{y}^*$ where the two lines cut, the slope of the poor perception needs to be higher than the slope of the rich perception:

$$\forall \hat{y}^* \text{ such that } U(\hat{y}^*, \bar{E}) - U(\hat{y}^*, \underline{E}_g) = U(\hat{y}^*, \bar{E}_{ng}) - U(\hat{y}^*, \underline{E})$$

I need

$$\frac{\partial U(\hat{y}^*, \bar{E})}{\partial \hat{y}} + \frac{\partial U(\hat{y}^*, \bar{E})}{\partial \bar{E}} \frac{\partial \bar{E}}{\partial \hat{y}} - \frac{\partial U(\hat{y}^*, \underline{E}_g)}{\partial \hat{y}} - \frac{\partial U(\hat{y}^*, \underline{E}_g)}{\partial \underline{E}_g} \frac{\partial \underline{E}_g}{\partial \hat{y}} <$$

$$\frac{\partial U(\hat{y}^*, \bar{E}_{ng})}{\partial \hat{y}} + \frac{\partial U(\hat{y}^*, \bar{E}_{ng})}{\partial \bar{E}_{ng}} \frac{\partial \bar{E}_{ng}}{\partial \hat{y}} - \frac{\partial U(\hat{y}^*, \underline{E})}{\partial \hat{y}} - \frac{\partial U(\hat{y}^*, \underline{E})}{\partial \underline{E}} \frac{\partial \underline{E}}{\partial \hat{y}}$$

$$\Leftrightarrow \frac{\partial U(\hat{y}^*, \bar{E})}{\partial \hat{y}} - \frac{\partial U(\hat{y}^*, \bar{E}_{ng})}{\partial \hat{y}} + \frac{\partial U(\hat{y}^*, \underline{E})}{\partial \underline{E}} \frac{\partial \underline{E}}{\partial \hat{y}} + \frac{\partial U(\hat{y}^*, \bar{E})}{\partial \bar{E}} \frac{\partial \bar{E}}{\partial \hat{y}} <$$

$$\frac{\partial U(\hat{y}^*, \underline{E}_g)}{\partial \hat{y}} - \frac{\partial U(\hat{y}^*, \underline{E})}{\partial \hat{y}} + \frac{\partial U(\hat{y}^*, \underline{E}_g)}{\partial \underline{E}_g} \frac{\partial \underline{E}_g}{\partial \hat{y}} + \frac{\partial U(\hat{y}^*, \bar{E}_{ng})}{\partial \bar{E}_{ng}} \frac{\partial \bar{E}_{ng}}{\partial \hat{y}}$$

$$\Leftrightarrow \frac{\partial U(\hat{y}^*,\bar{E})}{\partial \hat{y}} - \frac{\partial U(\hat{y}^*,\bar{E}_{ng})}{\partial \hat{y}} + \frac{\partial U(\hat{y}^*,\underline{E})}{\partial \underline{E}}\frac{\partial \underline{E}}{\partial \hat{y}} + \frac{\partial U(\hat{y}^*,\bar{E})}{\partial \bar{E}}\frac{\partial \bar{E}}{\partial \hat{y}} <$$

$$\frac{\partial U(\hat{y}^*,\underline{E}_g)}{\partial \hat{y}} - \frac{\partial U(\hat{y}^*,\underline{E})}{\partial \hat{y}} + \frac{\partial U(\hat{y}^*,\underline{E}_g)}{\partial \underline{E}_g}\left(\frac{\partial \underline{E}}{\partial \hat{y}} + \gamma F\right) + \frac{\partial U(\hat{y}^*,\bar{E}_{ng})}{\partial \bar{E}_{ng}}\left(\frac{\partial \bar{E}}{\partial \hat{y}} + \beta(1-F)\right)$$

(Note that I have used here the fact that with my definition of the bias I have $\frac{\partial \underline{E}_g}{\partial \hat{y}} = \frac{\partial \underline{E}}{\partial \hat{y}} + \gamma F$ and $\frac{\partial \bar{E}_{ng}}{\partial \hat{y}} = \frac{\partial \bar{E}}{\partial \hat{y}} + \beta(1-F)$, something that I prove below.) This is all that can be said about a necessary condition for uniqueness of equilibrium. Sufficient conditions for this necessary condition to be satisfied are

$$\frac{\partial U(\hat{y}^*,\bar{E})}{\partial \hat{y}} - \frac{\partial U(\hat{y}^*,\bar{E}_{ng})}{\partial \hat{y}} \leq \frac{\partial U(\hat{y}^*,\underline{E}_g)}{\partial \hat{y}} - \frac{\partial U(\hat{y}^*,\underline{E})}{\partial \hat{y}}$$

$$\frac{\partial U(\hat{y}^*,\bar{E})}{\partial \bar{E}} \leq \frac{\partial U(\hat{y}^*,\bar{E}_{ng})}{\partial \bar{E}_{ng}}$$

$$\frac{\partial U(\hat{y}^*,\underline{E})}{\partial \underline{E}} \leq \frac{\partial U(\hat{y}^*,\underline{E}_g)}{\partial \underline{E}_g}$$

whenever

$$U(\hat{y}^*,\bar{E}) - U(\hat{y}^*,\underline{E}_g) = U(\hat{y}^*,\bar{E}_{ng}) - U(\hat{y}^*,\underline{E})$$

These three conditions are basically conditions on the second derivatives of U. One utility function for which this is satisfied is $U(y,x) = yx$: At the equilibrium,

$$\hat{y}^*[\bar{E} - \underline{E}_g] = \hat{y}^*[\bar{E}_{ng} - \underline{E}]$$

or

$$[\bar{E} - \underline{E}_g] = [\bar{E}_{ng} - \underline{E}]$$

Hence

$$\frac{\partial U(\hat{y}^*,\bar{E})}{\partial \hat{y}} - \frac{\partial U(\hat{y}^*,\bar{E}_{ng})}{\partial \hat{y}} \leq \frac{\partial U(\hat{y}^*,\underline{E}_g)}{\partial \hat{y}} - \frac{\partial U(\hat{y}^*,\underline{E})}{\partial \hat{y}}$$

becomes

$$\bar{E} - \bar{E}_{ng} \leq \underline{E}_g - \underline{E}$$

which is satisfied at $\hat{y}^*$ (with equality), and

$$\frac{\partial U(\hat{y}^*,\bar{E})}{\partial \bar{E}} \leq \frac{\partial U(\hat{y}^*,\bar{E}_{ng})}{\partial \bar{E}_{ng}}$$

and

$$\frac{\partial U(\hat{y}^*,\underline{E})}{\partial \underline{E}} \leq \frac{\partial U(\hat{y}^*,\underline{E}_g)}{\partial \underline{E}_g}$$

both become

$$\hat{y}^* \leq \hat{y}^*$$

which is trivially satisfied with equality. Hence I can conclude that with $U(y,x) = yx$ the biased sorting equilibrium is unique.

Calculating $\hat{y}^*$ For $U(y,x) = yx$ I can also easily calculate the cutoff $\hat{y}^*$ at which the rich perception of the sorting benefits, $\bar{E} - \underline{E}_g$, are equal to the poor perception of the benefits, $\bar{E}_{ng} - \underline{E}$ (and they are equidistant to $\bar{E} - \underline{E}$)

$$\bar{E} - \underline{E}_g = \bar{E}_{ng} - \underline{E}$$

$$\bar{E} - \bar{E}_{ng} = \underline{E}_g - \underline{E}$$

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

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

Clearly, this is not a solution for $\hat{y}^*$ yet, because $\hat{y}$ appears on both sides here. It is straight forward to see that $\hat{y}^*$ must be a fixed point of the RHS. First, existence can be established by looking at the limits of this equation: as $\hat{y} \to 0$ the LHS is 0 and the RHS becomes E, so the LHS is smaller than the RHS. As $\hat{y} \to y_{\max}$ the LHS is $y_{\max}$ and the RHS is again E, so now the LHS is larger than the RHS, so surely a $\hat{y}^*$ will exist, as both the RHS and the LHS are continuous. Uniqueness can be established by noting that the slope of the RHS at $\hat{y}^* = \frac{\beta(1-F)\bar{E}+\gamma F\underline{E}}{\beta(1-F)+\gamma F}$ is always zero (so the 45 degree line cuts the RHS always from below, and hence it can only cut it once), and it can be shown that the 45 degree line (i.e. the LHS) cuts the RHS exactly at its minimum if $\beta < \gamma$ and at its maximum if $\beta > \gamma$.[3]

- If $\beta = \gamma$, 2 simplifies to $\hat{y} = E$, so the unique binary biased sorting equilibrium is such that the cutoff is exactly at the mean.

Theorem 10 *If $U(y,x) = yx$, a unique biased sorting equilibrium with consistency and bias $\underline{E}_g$ and $\bar{E}_{ng}$ exists and the unique cutoff $\hat{y}^*$ is the fixed point of $\frac{\beta(1-F)\bar{E}+\gamma F\underline{E}}{\beta(1-F)+\gamma F}$. If $\beta = \gamma$, the unique cutoff is at $\hat{y}^* = E$.*

The relationship between naivety and the cutoff I can actually define the bias with only one 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. I can write

$$\hat{y}^* = \frac{a(1-F)\bar{E} + F\underline{E}}{a(1-F) + F} \tag{3}$$

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

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

[3] Note that I do not need to establish existence and uniqueness for this utility function again, as I have already shown above that it satisfies the sufficient conditions.

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

Hence, $\hat{y}^*$ 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!

The effect of changing inequality on $\hat{y}$ Suppose that, after society is sorted into the two disjoint groups with cutoff $\hat{y}$, the poor group suddenly becomes poorer, so $\underline{E}$ decreases, but $F(\hat{y})$ stays constant (hence $\bar{E}$ stays the same). It is clear that the equilibrium cutoff will change, but how? From (3) it is immediate to see that the new cutoff will be lower than before. What happens to people's misperceptions? The misperceptions in the poor group will become more severe, while the misperceptions in the rich group become less severe (in terms of difference to the correct benefits of sorting). The perceived benefits of sorting of the rich and the poor must be equal again at the new cutoff: The perceived benefits of sorting of the rich are $\bar{E} - \underline{E}_g = \bar{E} - \gamma F\hat{y} - (1-\gamma F)\underline{E}$. Because both $\bar{E}$ and $\underline{E}_g$ go down if $\hat{y}$ decreases (and also the average income in the initial poor group), whether $\bar{E} - \underline{E}_g$ increases or decreases is ambiguous, and of course the same holds for $\bar{E}_{ng} - \underline{E}$. Even the correct benefits of sorting change in an ambiguous way. If $\bar{E} - \underline{E}$ is decreasing in $\hat{y}$ then they will for sure decrease, because both $\hat{y}$ and $\underline{E}$ decrease. If $\bar{E} - \underline{E}$ is increasing in $\hat{y}$, then the change is ambiguous: a lower $\hat{y}$ decreases the benefits of sorting, while a lower $\underline{E}$ increases them.

Suppose that the rich group gets richer on average, again ceteris paribus. Then it is straightforward to see from (3) that the new equilibrium cutoff will be higher.

Suppose that both things happen, so $\bar{E}$ increases, while $\underline{E}$ decreases (while F and $1-F$ stay constant). Then whether the new equilibrium cutoff is higher or lower than the old one depends on $a:$ if a is high, then the new equilibrium cutoff will be higher, if a is low, then the new equilibrium cutoff will be lower. If $a = 1$ then the cutoff is always E and hence will go down if E decreases due to this increase in inequality. E decreases if $F(E)$ high and $(1-F(E))$ is low, a feature that characterizes unequal distributions with positive skew.

1.1.3 Sorting under different types of misperceptions

After doing this analysis on existence and uniqueness of a binary biased sorting equilibrium if both groups *underestimate* the benefits of sorting, the natural question pops up on how this analysis would change if the biases were different. What if the biases are such that both groups overestimate the benefits of

sorting? Or one group underestimates them, the other overestimates them? In this section, I will examine all four possible combinations of biases and see what this implies for existence and uniqueness of equilibria. Note that I continue to assume that people perceive the average income in their own group correctly. It is just the average income in the other group that they are biased about.

For the following analysis, I will continue to assume that $U(.,.)$ is strictly increasing in both arguments, strictly supermodular and continuous. Furthermore, I leave the exact specification of the bias (i.e. $\underline{E}_g$ and $\bar{E}_{ng}$ as functions of $\hat{y}$) open, except for two very basic requirements, namely that $\underline{E}_g < \bar{E}$ and $\bar{E}_{ng} > \underline{E}$ for all $\hat{y}$.

The rich overestimate the poor and the poor overestimate the rich: This assumption means that the poor will overestimate the benefits of sorting and the rich will underestimate the benefits of sorting. I have

$$\underline{E}_g > \underline{E} \text{ and } \bar{E}_{ng} > \bar{E}$$

For a binary biased sorting equilibrium with cutoff $\hat{y}$, the following has to hold:

$$\exists\, b > 0 \text{ s.t.}$$

$$U(y_i, \bar{E}) - U(y_i, \underline{E}_g) \geq b \quad \forall y_i \geq \hat{y} \tag{IR1}$$

$$U(y_i, \bar{E}_{ng}) - U(y_i, \underline{E}) \leq b \quad \forall y_i \leq \hat{y} \tag{IR2}$$

This can be simplified to the following conditions (if $\underline{E}_g < \bar{E}$ and $\bar{E}_{ng} > \underline{E}$, a requirement which I have assumed above)

$$U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E}_g) \geq b \tag{IR1}$$

$$U(\hat{y}, \bar{E}_{ng}) - U(\hat{y}, \underline{E}) \leq b \tag{IR2}$$

Even without imposing the consistency requirement I see that this cannot hold because (IR1) and (IR2) would imply that

$$U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E}_g) \geq U(\hat{y}, \bar{E}_{ng}) - U(\hat{y}, \underline{E})$$

but this cannot hold because $U(\hat{y}, \bar{E}) < U(\hat{y}, \bar{E}_{ng})$ and $U(\hat{y}, \underline{E}_g) > U(\hat{y}, \underline{E})$!

Theorem 11 *If $\underline{E}_g > \underline{E}$ and $\bar{E}_{ng} > \bar{E}$ there is no biased binary monotonic sorting equilibrium. Also, there is no non-monotonic biased binary sorting equilibrium in this case.*

The rich underestimate the poor and the poor overestimate the rich This assumption means that both groups overestimate the benefits of sorting. I have

$$\underline{E}_g < \underline{E} \text{ and } \bar{E}_{ng} > \bar{E}$$

For a binary biased sorting equilibrium with cutoff $\hat{y}$, the following has to hold:

$$\exists\ b > 0 \text{ s.t.}$$

$$U(y_i, \bar{E}) - U(y_i, \underline{E}_g) \geq b \quad \forall y_i \geq \hat{y} \tag{IR1}$$

$$U(y_i, \bar{E}_{ng}) - U(y_i, \underline{E}) \leq b \quad \forall y_i \leq \hat{y} \tag{IR2}$$

This can be simplified to the following conditions if $\underline{E}_g < \bar{E}$ and $\bar{E}_{ng} > \underline{E}$ (which is automatically fulfilled in this case)

$$U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E}_g) \geq b \tag{IR1}$$

$$U(\hat{y}, \bar{E}_{ng}) - U(\hat{y}, \underline{E}) \leq b \tag{IR2}$$

Both (IR1) and (IR2) can hold at the same time.

Theorem 12 *A biased binary monotonic sorting equilibrium can exists in this case.*

Remark 13 *Also non-monotonic biased binary sorting equilibria can exist in this case, but the consistency requirement would rule these out.*

The consistency requirement yields

$$U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E}_g) \leq b \tag{CR1}$$

$$U(\hat{y}, \bar{E}_{ng}) - U(\hat{y}, \underline{E}) \geq b \tag{CR2}$$

and together this gives

$$U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E}_g) = U(\hat{y}, \bar{E}_{ng}) - U(\hat{y}, \underline{E})$$

as the equilibrium condition.

Theorem 14 *If $\underline{E}_g < \underline{E}$ and $\bar{E}_{ng} > \bar{E}$ a biased binary sorting equilibrium with consistency can exist and will be monotonic.*

Remark 15 *With this bias, b would be higher than in the unbiased case for the same partition. Or, put differently, in the unbiased case the same partition would happen with a lower b. With the same b, $\hat{y}$ would be higher in the unbiased model.*

To conclude, a biased binary sorting equilibrium with and without consistency CAN exist in this case, but under which circumstances will it exist? Analogous to the original case analyzed above, a sufficient condition for existence is that the rich are correct at 0 and the poor are correct at $y_{\max}$ and the biases $\underline{E}_g$ and $\bar{E}_{ng}$ are continuous functions of $\hat{y}$. (This is the case of Figure 1) Then there will be at least one $\hat{y}$ at which $U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E}_g) = U(\hat{y}, \bar{E}_{ng}) - U(\hat{y}, \underline{E})$ and hence there will exist at least one binary biased sorting equilibrium with

consistency. If the biases are such that the perceived benefits of sorting converge to the truth monotonically (for $\hat{y}$ to 0 resp. $y_{\max}$) then the equilibrium will be unique. In any case, at $\hat{y} = 0$ the rich will be correct and the poor will overestimate the benefits of sorting, so for $\hat{y}$ smaller than the smallest $\hat{y}^*$ where the perceived benefits of sorting are equal there can be no sorting equilibrium (because the poor would always be willing to pay more than the rich). After this first crossing, a biased sorting equilibrium exists for all $\hat{y}$ such that $U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E}_g) \geq U(\hat{y}, \bar{E}_{ng}) - U(\hat{y}, \underline{E})$ (if there are further crossings, then this condition will not hold for all $\hat{y}$ greater than the first $\hat{y}^*$). Clearly, consistency in addition is only satisfied where this inequality holds with equality. As we get to $\hat{y} = y_{\max}$ we know that $U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E}_g) > U(\hat{y}, \bar{E}_{ng}) - U(\hat{y}, \underline{E})$, so as $\hat{y} \to y_{\max}$ consistency is for sure not satisfied, but a binary biased sorting equilibrium exists for those $\hat{y}$. Note that the reasoning doesn't work in the exact opposite direction, i.e. if the poor are correct at 0 and the rich are correct at $y_{\max}$. The reason is that the rich will automatically be correct at 0 because as the poor group converges to a group where everybody just has income $\hat{y} = 0$ the rich cannot be wrong about the average in that group, and the same reasoning explains why the poor will also be correct at $y_{\max}$.

So really it is the more logical thing to assume that the bias works such that the rich are correct at 0 and the poor at $y_{\max}$.

Theorem 16 *If f $\underline{E}_g < \underline{E}$ and $\bar{E}_{ng} > \bar{E}$ and $\underline{E}_g$ and $\bar{E}_{ng}$ are continuous functions of $\hat{y}$ and such that the rich are correct at 0 and the poor at $y_{\max}$, a binary biased sorting equilibrium with consistency exists. If in addition $\underline{E}_g$ and $\bar{E}_{ng}$ are such that the perceived benefits of sorting monotonically decrease for the poor and increase for the rich as $\hat{y}$ increases, the equilibrium is unique.*

The rich underestimate the poor, the poor underestimate the rich This assumption implies that the rich overestimate the benefits of sorting and the poor underestimate the benefits of sorting. I have

$$\underline{E}_g < \underline{E} \text{ and } \bar{E}_{ng} < \bar{E}$$

For a binary biased sorting equilibrium with cutoff $\hat{y}$, the following has to hold:

$$\exists\, b > 0 \text{ s.t.}$$

$$U(y_i, \bar{E}) - U(y_i, \underline{E}_g) \geq b \quad \forall y_i \geq \hat{y} \tag{IR1}$$

$$U(y_i, \bar{E}_{ng}) - U(y_i, \underline{E}) \leq b \quad \forall y_i \leq \hat{y} \tag{IR2}$$

This can be simplified to the following conditions if $\underline{E}_g < \bar{E}$ and $\bar{E}_{ng} > \underline{E}$ (the first one is automatically fulfilled in this case)

$$U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E}_g) \geq b \tag{IR1}$$

$$U(\hat{y}, \bar{E}_{ng}) - U(\hat{y}, \underline{E}) \leq b \tag{IR2}$$

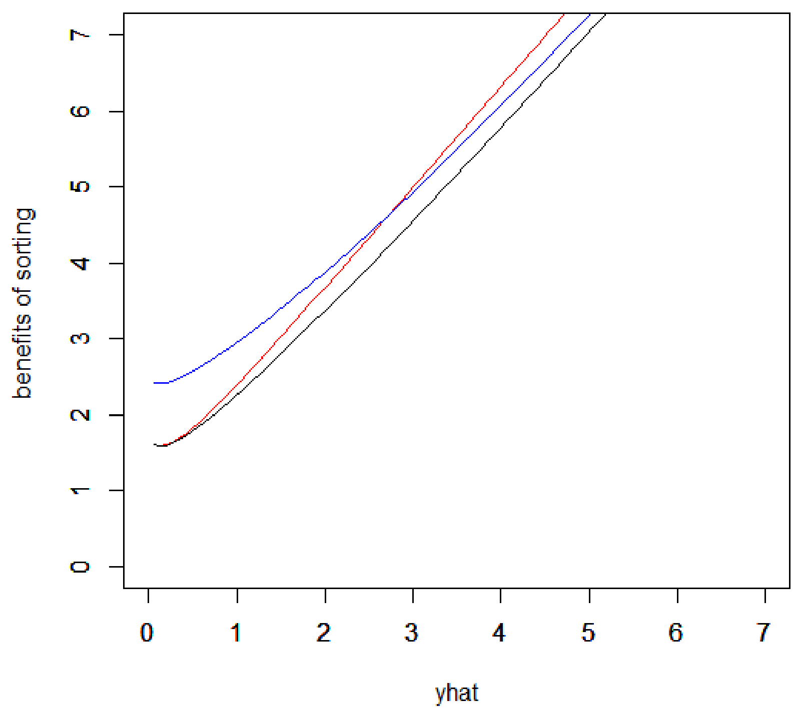

Figure 1: Both groups overestimate the benefits of sorting, the rich are correct at $\hat{y} = 0$, the poor are correct at $\hat{y} = y_{\max}$

Note that with this bias the rich overestimate the benefits of sorting and the poor underestimate them. (IR1) and (IR2) combined lead to

$$U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E}_g) \geq U(\hat{y}, \bar{E}_{ng}) - U(\hat{y}, \underline{E})$$

which always holds if the bias is like this, and note that the inequality will even be strict in this case.

Theorem 17 *If $\underline{E}_g < \underline{E}$ and $\bar{E}_{ng} < \bar{E}$, a monotonic binary biased sorting equilibrium exists. Note that non-monotonic sorting equilibria are not possible in this case.*

The consistency requirement yields

$$U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E}_g) \leq b \tag{CR1}$$

$$U(\hat{y}, \bar{E}_{ng}) - U(\hat{y}, \underline{E}) \geq b \tag{CR2}$$

Together, these four conditions yield

$$U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E}_g) = U(\hat{y}, \bar{E}_{ng}) - U(\hat{y}, \underline{E})$$

but this cannot hold, because as I noted, the inequality will be strict in this case.

Theorem 18 *If $\underline{E}_g < \underline{E}$ and $\bar{E}_{ng} < \bar{E}$, a binary biased sorting equilibrium with consistency does not exist.*

The rich overestimate the poor and the poor underestimate the rich (original case) This assumption means that both groups underestimate the benefits of sorting. I have

$$\underline{E}_g > \underline{E} \text{ and } \bar{E}_{ng} < \bar{E}$$

For a binary biased sorting equilibrium with cutoff $\hat{y}$, the following has to hold:

$$\exists\ b > 0 \text{ s.t.}$$

$$U(y_i, \bar{E}) - U(y_i, \underline{E}_g) \geq b \quad \forall y_i \geq \hat{y} \tag{IR1}$$

$$U(y_i, \bar{E}_{ng}) - U(y_i, \underline{E}) \leq b \quad \forall y_i \leq \hat{y} \tag{IR2}$$

This can be simplified to the following conditions if $\underline{E}_g < \bar{E}$ and $\bar{E}_{ng} > \underline{E}$

$$U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E}_g) \geq b \tag{IR1}$$

$$U(\hat{y}, \bar{E}_{ng}) - U(\hat{y}, \underline{E}) \leq b \tag{IR2}$$

Note that both groups underestimate the benefits of sorting in this case. Both (IR1) and (IR2) can hold at the same time.

Theorem 19 *A biased binary monotonic sorting equilibrium can exists in this case.*

Remark 20 *Also non-monotonic biased binary sorting equilibria can exist in this case, but the consistency requirement would rule these out.*

The consistency requirement yields

$$U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E}_g) \leq b \tag{CR1}$$

$$U(\hat{y}, \bar{E}_{ng}) - U(\hat{y}, \underline{E}) \geq b \tag{CR2}$$

and together this gives

$$U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E}_g) = U(\hat{y}, \bar{E}_{ng}) - U(\hat{y}, \underline{E})$$

as the equilibrium condition.

Theorem 21 *If $\underline{E}_g > \underline{E}$ and $\bar{E}_{ng} < \bar{E}$ a biased binary sorting equilibrium with consistency can exist and will be monotonic.*

Remark 22 *With this bias, b would be lower than in the unbiased case for the same partition. Or, put differently, in the unbiased case the same partition would happen with a higher b. With the same b, $\hat{y}$ would be lower in the unbiased model.*

To conclude, a biased binary sorting equilibrium with and without consistency CAN exist, but under which circumstances will it exist? A sufficient condition for existence is that the rich are correct at 0 and the poor are correct at $y_{\max}$ and the biases $\underline{E}_g$ and $\bar{E}_{ng}$ are continuous functions of $\hat{y}$. (This is the case of Figure 2) Then there will be at least one $\hat{y}$ at which $U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E}_g) = U(\hat{y}, \bar{E}_{ng}) - U(\hat{y}, \underline{E})$ and hence there will exist at least one binary biased sorting equilibrium with consistency. If the biase are such that the perceived benefits of sorting converge to the truth monotonically (for $\hat{y}$ to 0 resp. $y_{\max}$) then the equilibrium will be unique. In any case, at $\hat{y} = 0$ the rich will be correct and the poor will underestimate the benefits of sorting, so for all $\hat{y}$ smaller than the smallest $\hat{y}^*$ where the perceived benefits of sorting are equal those $\hat{y}$ will be sorting equilibria (because the rich would always be willing to pay more than the poor), but consistency will not be satisfied. After this first crossing, only all $\hat{y}$ such that $U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E}_g) \geq U(\hat{y}, \bar{E}_{ng}) - U(\hat{y}, \underline{E})$ are sorting equilibria (so if there is only a single crossing then no $\hat{y} > \hat{y}^*$ is a biased sorting equilibrium). Clearly, consistency in addition is only satisfied where this inequality holds with equality. As we get to $\hat{y} = y_{\max}$ we know that $U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E}_g) < U(\hat{y}, \bar{E}_{ng}) - U(\hat{y}, \underline{E})$, so as $\hat{y} \to y_{\max}$ this is for sure no sorting equilibrium. Note that the reasoning doesn't work in the exact opposite direction if the poor are correct at 0 and the rich are correct at $y_{\max}$, because the rich will automatically be correct at 0 because as the poor group converges to a group where everybody just has income $\hat{y} = 0$ the rich cannot be wrong

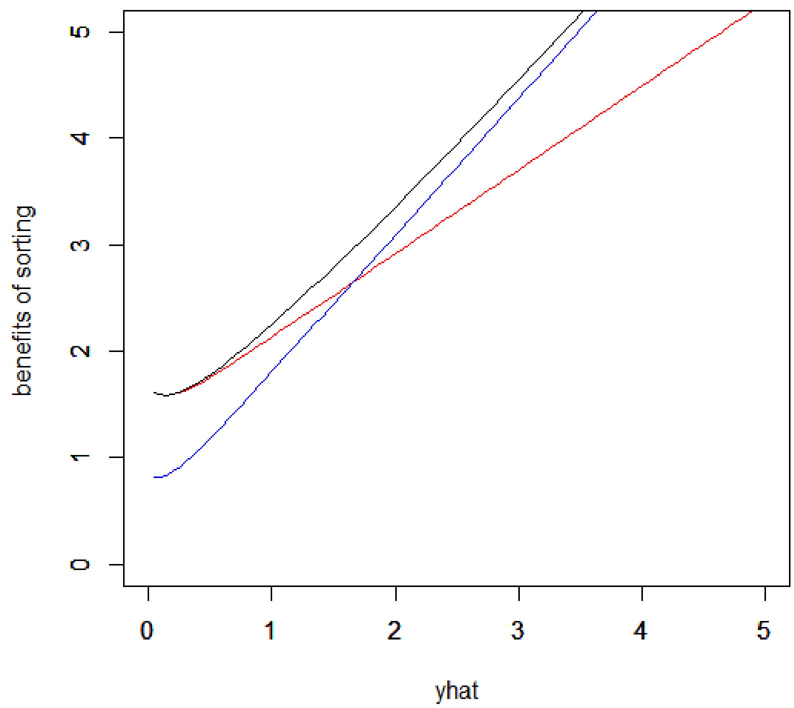

Figure 2: Both underestimate the benefits of sorting, rich are correct at 0, poor at $y_{\max}$

about the average in that group, and the same reasoning explains why the poor will also be correct at $y_{\max}$.

So really it is the more logical thing to assume that the bias works such that the rich are correct at 0 and the poor at $y_{\max}$.

Theorem 23 *If $\underline{E}_g > \underline{E}$ and $\bar{E}_{ng} < \bar{E}$ and $\underline{E}_g$ and $\bar{E}_{ng}$ are continuous functions of $\hat{y}$ and such that the rich are correct at 0 and the poor at $y_{\max}$, a binary biased sorting equilibrium with consistency exists. If in addition $\underline{E}_g$ and $\bar{E}_{ng}$ are such that the perceived benefits of sorting monotonically increase for the poor and decrease for the rich as $\hat{y}$ increases, the equilibrium is unique.*

The rich underestimate the poor, the poor underestimate the rich This means the rich overestimate the benefits of sorting, the poor underestimate the benefits of sorting. I have

$$\underline{E}_g < \underline{E} \text{ and } \bar{E}_{ng} < \bar{E}$$

For a binary biased sorting equilibrium with cutoff $\hat{y}$, the following has to hold:

$$\exists\, b > 0 \text{ s.t.}$$

$$U(y_i, \bar{E}) - U(y_i, \underline{E}_g) \geq b \quad \forall y_i \geq \hat{y} \tag{IR1}$$

$$U(y_i, \bar{E}_{ng}) - U(y_i, \underline{E}) \leq b \quad \forall y_i \leq \hat{y} \tag{IR2}$$

This can be simplified to the following conditions if $\underline{E}_g < \bar{E}$ and $\bar{E}_{ng} > \underline{E}$ (the first one is automatically fulfilled in this case)

$$U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E}_g) \geq b \tag{IR1}$$

$$U(\hat{y}, \bar{E}_{ng}) - U(\hat{y}, \underline{E}) \leq b \tag{IR2}$$

Note that with this bias the rich overestimate the benefits of sorting and the poor underestimate them. (IR1) and (IR2) combined lead to

$$U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E}_g) \geq U(\hat{y}, \bar{E}_{ng}) - U(\hat{y}, \underline{E}) \tag{4}$$

which always holds if the bias is like this, and note that the inequality will even be strict in this case.

Theorem 24 *If $\underline{E}_g < \underline{E}$ and $\bar{E}_{ng} < \bar{E}$, a binary biased sorting equilibrium exists. Note that non-monotonic sorting equilibria are not possible in this case.*

The consistency requirement yields

$$U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E}_g) \leq b \tag{CR1}$$

$$U(\hat{y}, \bar{E}_{ng}) - U(\hat{y}, \underline{E}) \geq b \tag{CR2}$$

Together, these four conditions yield

$$U(\hat{y}, \bar{E}) - U(\hat{y}, \underline{E}_g) = U(\hat{y}, \bar{E}_{ng}) - U(\hat{y}, \underline{E})$$

but this cannot hold, because, as I noted, inequality (4) will be strict in this case.

Theorem 25 *If $\underline{E}_g < \underline{E}$ and $\bar{E}_{ng} < \bar{E}$, a binary biased sorting equilibrium with consistency does not exist.*

Conclusion 26 *There are two cases in which a binary biased sorting equilibrium with consistecy exists: the poor underestimate the rich and the rich underestimate the poor (original case) and the rich underestimate the poor and the poor overestimate the rich (see above). If the rich underestimate the poor and the poor underestimate the rich, then every $\hat{y}$ constitutes a sorting equilibrium, but consistency is never satisfied. If the poor overestimate the rich and the rich overestimate the poor then no biased sorting equilibrium can exist, because the poor would at any cutoff be willing to pay more for the sorting than the rich.*

2 Monopolist profit without bias

So far, I have not mentioned who would offer the sorting technology? Who offers people to pay a certain fee b to join a club and then ensures that they will interact only with other people from the club? We could for instance imagine that this technology is offered by a monopolist.

If people are unbiased, any cutoff $\hat{y}$ and corresponding sorting fee b will constitute a sorting equilibrium. The monopolist can choose the sorting fee b and hence the cutoff $\hat{y}$ to maximize her profit:

$$\max b \int_{\hat{y}}^{y_{\max}} f(y)dy = \max b(1 - F(\hat{y}) = \max_{\hat{y}} \hat{y}(\bar{E} - \underline{E})(1 - F(\hat{y}))$$

Clearly, if $\hat{y} = 0$ then b is zero and hence the monopolist's profit will be zero. If $\hat{y} = y_{\max}$ then $1 - F(\hat{y})$ is zero and the profit will be zero again. For any interior $\hat{y}$, the monopolist's profit will be strictly positive. The FOC delivers candidates for a maximum:

$$\hat{y}f(\hat{y})\left(\frac{\bar{E} - \hat{y}}{1 - F} - \frac{\hat{y} - \underline{E}}{F}\right)(1 - F) - \hat{y}(\bar{E} - \underline{E})f(\hat{y}) + (\bar{E} - \underline{E})(1 - F) = 0$$

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

Note that the median can never be a solution to this, as the RHS would be zero then, but the LHS would not!

From the above finding that the monopolist's profit is zero at the boundaries and positive for any interior $\hat{y}$, we know that the profit maximizing $\hat{y}$ will be among those candidates for a local maximum given by the FOC, if there is more than one then the monopolist simply picks the one that delivers the highest profit. If the profit function is concave, the maximizer is unique.

3 Monopolist profit with bias

3.1 Poor underestimate rich, rich overestimate poor

If we require a sorting equilibrium with consistency, then the monopolist cannot choose $\hat{y}$, but has to fix the sorting fee such that $\hat{y} = \hat{y}^*$, and the corresponding profits are

$$b(1 - F(\hat{y}^*)) = \hat{y}^*(\bar{E} - \underline{E}_g)(1 - F(\hat{y}^*)) = \hat{y}^*(\bar{E}_{ng} - \underline{E})(1 - F(\hat{y}^*))$$

Clearly in this case the profits are smaller than in the unbiased case, for two reasons: First, for any cutoff the corresponding sorting fee is smaller (both groups underestimate the benefits of sorting) and second, the monopolist cannot choose the profit maximizing $\hat{y}$, but has to stick with $\hat{y}^*$.

If we drop the consistency requirement, then the monopolist can choose any $\hat{y}$ up to $\hat{y}^*$ (she cannot go beyond $\hat{y}^*$ because the poor would value sorting more than the rich beyond that point). Still, the profits will obviously be lower than in the unbiased case.

3.2 Poor overestimate rich, rich underestimate poor

If we require a sorting equilibrium with consistency, then the monopolist cannot choose $\hat{y}$, but has to fix the sorting fee such that $\hat{y} = \hat{y}^*$, and the corresponding profits are

$$b(1 - F(\hat{y}^*)) = \hat{y}^*(\bar{E} - \underline{E}_g)(1 - F(\hat{y}^*)) = \hat{y}^*(\bar{E}_{ng} - \underline{E})(1 - F(\hat{y}^*))$$

If we drop the consistency requirement, the monopolist can choose any $\hat{y} \geq \hat{y}^*$. In this case, the profits could theoretically be higher than in the unbiased case. For any given cutoff, the corresponding sorting fee is higher than in the unbiased case (both groups overestimate the benefits of sorting), so if the monopolist could freely choose $\hat{y}$, her profits would surely be higher than in the unbiased case. However, the monopolist is restricted in her choice of $\hat{y}$ to either $\hat{y} = \hat{y}^*$ or $\hat{y} \geq \hat{y}^*$, so it is not clear whether profits will be higher or lower than in the unbiased case.

3.3 Poor underestimate rich, rich underestimate poor

In this case a biased sorting equilibrium with consistency does not exist, but any $\hat{y}$ constitutes a biased sorting equilibrium without consistency. As the poor underestimate the benefits of sorting and the rich overestimate the benefits of sorting, there always exists a sorting fee that the rich are willing to pay to separate from the poor, while the poor are not interested in doing this. Clearly, the monopolist's profits will be higher than in the unbiased case here, because she can choose any $\hat{y}$ just like in the unbiased case, but the corresponding sorting fee is higher, because the rich overestimate the benefits of sorting and the monopolist will choose the maximum possible sorting fee.

4 Appendix

- With the way I define the bias in the main case I aways have

  $$\bar{E}_{ng} > \underline{E}$$

  and

  $$\bar{E} > \underline{E}_g$$

  (which I need to have for my proofs).

- Note: What is $\frac{\partial \bar{E}_{ng}}{\partial \hat{y}}$ and $\frac{\partial \underline{E}_g}{\partial \hat{y}}$?

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

$$\frac{\partial \bar{E}_{ng}}{\partial \hat{y}} = \beta(-f(\hat{y}))\hat{y} + \beta(1-F) + (1-\beta(1-F))\frac{\partial \underline{E}}{\partial \hat{y}} + \beta f \bar{E}$$

$$= \frac{f(\hat{y})(\bar{E}-\hat{y})}{1-F} + \beta(1-F) = \frac{\partial \bar{E}}{\partial \hat{y}} + \beta(1-F) > \frac{\partial \bar{E}}{\partial \hat{y}}$$

and

$$\underline{E}_g = \gamma F \hat{y} + (1-\gamma f)\underline{E}$$

$$\frac{\partial \underline{E}_g}{\partial \hat{y}} = \gamma F \hat{y} + \gamma F - \gamma f \underline{E} + (1-\gamma F)\frac{\partial \underline{E}}{\partial \hat{y}}$$

$$= \frac{f(\hat{y})(\hat{y}-\underline{E})}{F} + \gamma F = \frac{\partial \underline{E}}{\partial \hat{y}} + \gamma F > \frac{\partial \underline{E}}{\partial \hat{y}}$$