

Referee Report for: Guilt aversion and Redistributive Politics

1. Overview

The paper addresses the long standing question of why we observe different levels of redistribution in representative democracies (notably in the EU and in the US) even between countries with roughly similar economic fundamentals.

The paper's answer is that voters are not completely selfish. In line with the large literature on social preferences, the author assumes that they weight their own income against the desire to obtain a more fair distribution of wealth. This is particularly important in cases in which part of individuals' income is determined by the natural or social lottery.

A brief look at the paper reveals that the author has spent a fairly large amount of time on it and surely he cares about his argument and results. However, I must say that I don't think the paper is publishable in the present form. Also, I believe that the modifications that are necessary to make it publishable would be so many and so deep to mutate the paper into a different object. That's why I'm not asking for a revision, but rather encourage the author to write another version of the paper. What follow are my constructive comments the author may consider in case he decides to follow my advice.

I am mostly concerned with the way the content of the paper is exposed, rather than the content itself, although I also have some issues concerning the content (see below).

The *first* problem is: the mathematics *is* cumbersome. The paper presents a main proposition (Proposition 1) and a second proposition which is an extension of the first (Proposition 3). What follows is an incomplete list of the symbols that are used:

$$e_{it}, e_{it}^{im}, e_{it}^{em}, u_{it}, c_{it}, a_i, \tau_i, \epsilon_i, \hat{u}_{it}, F_i, \sigma_\epsilon, \sigma_a, L, \Delta, \Omega_t, G_t, \Phi, \bar{a}$$

Way too many for such a meagre result. Because of such heavy notation, the paper's proofs are almost impossible to read. It would have taken me at least a week just to make sense of the math. To have an idea, just look at the main proposition

Proposition 1 Consider the tax sequence $\{\tau_t\}_{t=0}^{t=\infty}$ and assume that $\tau^f - \tau^s \geq 2\sqrt{\frac{\Phi}{2\bar{a}-a_{med}}}$, if $\tau_0 \in]\tau^f - \delta; \tau^f + \delta[$ then $\lim_{t \rightarrow \infty} \tau_t = \tau^f$, otherwise $\lim_{t \rightarrow \infty} \tau_t = \frac{1}{2} \left(\tau^f + \tau^s - \sqrt{(\tau^f - \tau^s)^2 - \frac{2\Phi\tau^s}{\Delta}} \right)$ where $\delta = \frac{\tau^f - \tau^s - \sqrt{(\tau^f - \tau^s)^2 - \frac{2\Phi\tau^s}{\Delta}}}{2}$, $\tau^f = 1 - \sqrt{\inf \left(1 - \frac{L}{2}, 1 \right)}$, $L = \frac{\sigma_a^2}{\sigma_\epsilon^2}$, $\tau^s = \frac{2(\bar{a}-a_{med})}{2\bar{a}-a_{med}}$ and $\Delta = \bar{a} - a_{med}$.

The *second* problem is that the author seems to be unsure about his own argument, and so backs *any* statement with two or three quotations. For a paper that totals 24 pages, there are 99(!) bibliographical entries. Most of the quotes are just redundant. Consider the following note (page 15)

^s As stated by Bisin and Verdier (2001), internalization of cultural practices can occur through socialization inside the family and in the society via imitation and learning.

Bisin and Verdier may well have said that, but so did many other authors, who on this issue are more authoritative than Bisin. For example Robert Boyd and Peter Richerson.

Part of the problem is that the author violates a standard rule in academic papers in economics, according to which the model is kept separated from the literature. It is a common practice to write an introductory section in which the existing literature is summarized, which shows where the contribution of the paper lies. In the following sections the author just makes his own assumptions and derives his results, without bothering the reader with continuous references to the existing literature. In the present paper, references to other papers keep popping up almost at any page. This makes very painful to get through the paper.

2. Content

As for the content, I think that the author sets too an ambitious goal for his paper. For example, I cannot see the point of all the references to Kahneman dual-process decision-making. What the author does in the paper is pretty standard nowadays in economics: instead of maximizing money, economic agents maximize a combination of money plus fairness. Authors who make this claim (Fehr & Schmidt on inequality aversion, Charness and Rabin on fairness, Dufwenberg on guilt aversion and so on) do not make general claims about the dual process of decision making (or any other sophisticated psychological construct), and I think the author should follow their example. The use of the psychological jargon induces the author to make questionable statements such as:

According to our analysis, the high redistribution European style welfare state is then characterized by an emotional contagion whereas the low redistribution American style welfare system is characterized by a rational control: $\tau^s < \tau^{US} < \tau^{EU} = \tau^f$. As a consequence, in the European style

The point is not whether the Europeans are more emotional than the Americans. The point is that the type of model the author discusses does not imply such an interpretation. Why not simply state that people have different preferences which are partly shaped by their past experiences? This is what the mathematical side of the paper says, and I see no point in making further, and highly debatable, statements concerning emotions vs. reason.

As I said, I did not check the math mostly because I think it is far too involved for what it delivers. But a second reason is that there are points in which it raises a few doubts that are hard to clear because of the poor way in which proofs are organized. I shall make three examples.

Consider *first* equation (11). It defines the experience of guilt for an adult, on the basis of his experience as a young man

$$G_t = \frac{(\tau_t^f - \tau_t)}{[\tau_{t-1}^f - \tau_{t-1}^*]^2}.$$

At the denominator we have the difference between the fair redistribution at time $t - 1$ (τ_{t-1}^f , that is when the individual was young) and the redistribution chosen at $t - 1$ (τ_{t-1}^*). (Notice that τ_{t-1}^* is defined a couple of pages later than it is first used in equation 11.)

But wouldn't this imply that if at $t - 1$ the most equitable redistribution was implemented (i.e. $\tau_{t-1}^f = \tau_{t-1}^*$) then guilt would be infinite for any $\tau_t^f \neq \tau_t^*$. And what is the meaning of an infinite guilt. Maybe I am missing something (the meaning of the square brackets?) but if this is the case it must be something which is not clearly stated in the text.

Second: Figure 3 represents today's fair redistribution in terms of yesterday's. What strikes me is that the relationship is not monotone. I would expect that the more equitable is today's society, the more it will be tomorrow (notice that this does *not* imply that the fairness of the redistribution is bound to increase to infinity, as the steady state may well be finite even if the relationship between yesterday and today is strictly increasing.) Instead, in the picture the fairness of today's redistribution increases with yesterday's up to point τ^{EU} , and then decreases. So if you live in a country that today is more fair than the average EU country (say Finland) then tomorrow you will want to have less fairness than a person who lived in a more unfair country (say Greece). This sounds paradoxical to me, but the author makes no comment on this point. I

suspect that the reason for this paradox is in the special parametric formula the author uses for individual utility, but then again it would be very difficult to tell.

Finally, Proposition 1 seems to suggest that there is a dynamic process of adjustment over time that (depending on initial conditions) either converges to the EU or to the US level of fairness. However, in the Appendix I only see a calculation of the steady states, not a proof that these steady states are stable. How can the author rule out a limit cycle?

3. Bottom Line

The author has read a lot on his topic and done quite a lot of work on the model. I think it would be a waste just to throw all this away. However, he should try to write a different, less ambitious paper, with a mathematical model which is much more reader-friendly.