

Referee report on 'The complexity of economic policy: restricted local optima in tax policy design' by Gilles Saint-Paul

This paper considers a situation where the government does not implement the first best policy in one sweeping reform, but where it implements a series of small welfare improving reforms. These reforms continue until no (small) reform can improve welfare. This is called a restricted local optimum (RLO). Clearly, the order in which the government does the reforms, matters for the final outcome. The paper assumes that the government does this in a more or less random way. Indeed different realizations of this process leads to different results. There is history dependence. The paper then calculates the performance of the economy for different values of the number of instruments available (here the number of tax brackets) and for different values of the labor supply elasticity.

The most fascinating result is that as the number of instruments (tax brackets) increases, the performance of the economy goes down. In a first best world, this cannot happen. More instruments can only improve welfare. However, because of the history dependence problem, more instruments can make it harder to get to a welfare maximizing RLO and the economy ends up more often at inefficient RLO's.

The paper is based on simulations only. It does not try to create an analytical framework in which we can understand why more complexity leads to less efficient outcomes on average. To be honest, I would also not know how this can be done. Therefore, the current attempt to make this point is good enough for me. I would only like to see one point (as a robustness check). At the moment, the selection of the bracket to be reformed is rather arbitrary. However, with some sort of democratic process, it would not be weird to assume that the next reform is the one where the welfare function is steepest. Then at each moment in time the marginal reform with the biggest effect on welfare would be undertaken. We know that steepest ascent algorithms do not necessarily lead us to a global optimum. Hence the point of the paper that piece meal reforms do not lead to the first best outcome is still valid.

The question I have is: is it still the case with this reform selection process that more complexity leads to less desirable outcomes on average?