Economic Inequality II: Implications for Redistributive Taxation

February 1, 2017

These notes are based on a presentation by Professor Maximilian Kasy (Harvard, Department of Economics) for the section on Economic Inequality in the Mechanism Design for Social Good Reading Group meeting as well as notes from Professor Kasy's online textbook Inequality Research, [2]. This document is written by members of the reading group and not the presenter.

1 Introduction

- We want to choose some policy/institutional rules/design of platform in order to maximize some objective social welfare function (SWF) subject to social, technological, and informational constraints.
- An overarching question here is: what do we actually want to optimize? We will discuss this in the first part of the talk. In the second part, we will go through an example of optimal redistributive taxation.
- We tap into discussions on the theory of justice. Some references include:
  - *A Theory of Justice*, J. Rawls, [3].
  - *Inequality Reexamined*, A. Sen, [6].
  - *Theories of Distributive Justice*, J.E. Roemer, [4].
    - Part 1: Survey of economic literature on social choice
    - Part 2: Survey of philosophical literature

2 Normative Individualism

- We can capture a lot of theories of justice within the framework of normative individualism.
- Common assumptions for most theories of justice are:
  - We can make normative statements about society based on statements about individual welfare.
  - We define social welfare as a function of individuals' welfare. SWF $= F(v_1, \cdots, v_n)$, where v_i are welfare functions for each individual $i \in \{1, \cdots, n\}$.
- This leads to some important questions:

- Who is to be included among the individuals? All citizens, all residents, all humans on earth? What about all future generations, animals, etc?
- How do we measure individual welfare v_i?
  - Opportunities or outcomes?
  - Utility, resources, capabilities?
- How do we aggregate to SWF? For instance, how much do we care about:
  - Individuals A vs B, C vs. D?
  - Millionaires vs. homeless people?
  - Sick vs. healthy people?
  - Groups that were victims of historic injustice?
- **Takeaway:** You are always assuming an answer to these questions when formulating an objective for a mechanism design problem.

- This framework does not cover everything. For instance:
  - Libertarian: "Market outcomes are just, no matter what the welfare consequences for individuals."
  - Fascist: "What counts is the greatness of the nation."
  - Environmental concerns: "Preservation of the environment is a value, beyond its consequences for humans."

3 Individual Welfare

- There are different approaches to measuring individual welfare:
  - **Utilitarian approach:**
    - This is the dominant approach in economics. Here, welfare is maximum utility over all options in a choice set.
    - Formally, we have a choice set C_i, utility function $u_i(x)$ for $x \in C_i$, and we have realized welfare
    $$v_i = \max_{x \in C_i} u_i(x).$$
    - Utility has a double role:
      - We can use it to determine choices: i.e., individuals choose their utility maximizing x.
      - It can be used as a normative yardstick: i.e., welfare is realized through utility.
    - Policies do not change $u_i(\cdot)$, but they change C_i, which changes v_i.
    - There are issues with this utilitarian approach:
      - Preferences do not exist in a pre-social vacuum. Constraints and utility are often jointly determined (e.g., parental aspirations, gender norms, etc).
      - People might not always act according to their preferences. We may need to separate positive and normative roles of utility. This is studied through behavioral economics.

        - It is also not immediately evident how to compare utility across people. (E.g., we cannot observe people making choices over others' utilities, but only their own. That is, we cannot use observational content and revealed preference theory.)
    - **Capabilities approach:**
        - Proposed by Sen in "Inequality Reexamined."
        - It says to think about choice sets directly, without reference to utility. i.e., evaluate C_i directly, without reference to u_i.
        - The focus is on "capabilities to function" (such as contribute to the work force) subject to all constraints faced by individuals such as legal, economic, political, social norms, and so on.
        - It makes a distinction between choices individuals make versus options that are available to them. (E.g., religious fasting vs. starving, respectively.)
    - **Opportunities approach:**
        - This is a more empirical or pragmatic approach.
        - Proposed by Roemer in "Equality of Opportunity."
        - Define a list of observable factors called circumstances such as parental background, race, gender, and so on.
        - Inequality predicted by these factors is "inequality of opportunity." The rest is "inequality of effort."
        - There are shortcomings with this approach:
            - Picking out the list of factors is not an easy task.
            - Likewise, it is not simple to distinguish between circumstances and effort.
            - What about things that are neither circumstance nor effort such as luck? (E.g., sudden shocks to income.) It's not evident that this is necessarily accounted for under this model.

4 Aggregation

- How do we aggregate individual v_i's into the SWF?
- One answer is to use welfare weights ω_i:
    - Given SWF $= F(v_1, \cdots, v_n)$, define

$$\omega_i = \frac{\partial}{\partial v_i} F(v_1, \cdots, v_n).$$

    - These welfare weights measure how much we care about increasing the welfare of each individual.
    - Then, we can look at what happens due to small change of some policy:

$$d\text{SWF} = \sum_i \omega_i \cdot dv_i.$$

- It is a challenge how to pick these welfare weights as there is no "objective" way. There are questions like, what do we think the welfare weights should depend on? Income, health, etc? One solution to this is the following thought experiment:
- *The Veil of Ignorance:* Proposed by Rawls in "A Theory of Justice."
- This is the choice of welfare weights as a decision problem, formalizing impartiality.
- Imagine: You know nothing about yourself, the next morning you wake up as any $i \in \{1, \cdots, n\}$, and you have to pick between social arrangements and policies. By what criterion would you pick?
- Rawls' answer is: Faced with fundamental uncertainty, you want to insure yourself as much as possible. You want to mitigate the worst possible outcome. Thus, evaluate arrangements based on welfare of the person worst off (called maximim):

$$F(v_1, \cdots, v_n) = \min_i v_i.$$

5 Redistributive Taxation

- This is one important policy tool for dealing with economic inequality. Income taxation is an important special case, although the ideas here could be applied to similar problems including wealth taxes, inheritance taxes, and so on.
- There is a field that aims to derive "optimal taxes" including optimal redistributive income taxes. Here, we look at the distributive impact of changing taxes on individuals' welfare.
- Some important references are:
  - *Using Elasticities to derive optimal income tax rates*, E. Saez, [5].
  - *Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods*, R. Chetty, [1].
- The key assumptions in this section are:
  - We evaluate individual welfare in terms of utility,
  - We take welfare weights as given, and
  - We impose government budget constraint.
- Note that this last point suggests that we need to talk about government revenues, which adds to the complexity of the problem. We consider small changes in tax rates subject to government budget constraints and study the distributive impact of this.
- When are taxes optimal?
  - Optimality: No feasible change improves social welfare. That is, there is no effect on social welfare for any feasible small change. We care about finding policies that maximize social welfare.
  - Effect of change considers:

    - Individual welfare: If we measure individual welfare by using utility and assume that individuals are maximizing utility subject to the present constraints then we can ignore the effect of behavioral responses to policy changes, a result known as *the Envelope Theorem*. Details in [2].
    - Social welfare: The envelope theorem allows us to evaluate the effect of a policy change on any individual. To get from this to social welfare, we use the welfare weights above and sum up individual welfare using these weights. Recall again that these weights are not scientific and the trade-offs involve moral judgments as discussed above.

- We have already stated that we impose government budget constraints. In addition to individual and social welfare, we also have to think about what impact changing taxation has on government revenues. There are two ways to deal with this, which are mathematically equivalent:
  - Only consider tax changes that do not change total revenue.
  - Assume there is a marginal value λ of additional revenues where λ is set to be on the same scale as the welfare weights.
- Related to this is considering behavioral responses to tax changes as these will impact government revenues. For instance, rich individuals might end up exploiting loopholes, moving their wealth to the Cayman Islands, and so on.
- Similarly, we also have to take into account the effect of changing taxes on prices and wages, which can further complicate the problem. Here, we assume that prices and wages are not affected by changing taxes, but there is a vast amount of empirical literature studying these effects.

5.1 Optimal Taxation

- To study the effects on social welfare:
  - Consider small changes $d\tau$ of some tax on the income distribution α, where τ is the tax rate.
  - The effect on social welfare is then:
$$d\text{SWF} = \sum_i \omega_i \cdot EV_i.$$
where the ω_i are the weights (the value of additional dollar for person i) and EV_i is equivalent variation. (See Chapter 9 of [2] for details.)
  - By the envelope theorem: EV_i is the mechanical effect on i's budget, holding all choices constant. For instance $EV_i = -x_i \cdot d\tau$ for τ on x_i.
- The effect on government budget G is:
  - The sum of the mechanical effect and behavioral effect. For instance for a tax rate τ on x_i,
$$dG = \sum_i x_i \cdot d\tau + dx_i \cdot \tau.$$

- Note: Estimating dx_i, which is the change in behavioral effect, is difficult.

- Top income taxes:
  - Suppose welfare weights are very small for the very rich, relative to average. That is ω_i goes to 0 as income goes to infinity. Then, optimal top income taxes also maximize revenue, so this is a simpler problem.
  - We have tax rate τ for incomes above threshold $\bar{y}$. Government revenues from top bracket, per tax payer is:

$$G(\tau) = \tau \cdot (E[Y \mid Y \geq \bar{y}] - \bar{y}).$$

  - Mechanical and behavioral effect:

$$\partial_\tau G(\tau) = (E[Y \mid Y \geq \bar{y}] - \bar{y}) + \tau \cdot E[\partial_\tau Y \mid Y \geq \bar{y}] = 0$$

    the first one is the mechanical impact and the second part is the behavioral impact.
  - Remember the property of the Pareto distribution that:

$$P(Y > y \mid Y \geq \bar{y}) = (\bar{y}/y)^\alpha$$
$$E[Y \mid Y \geq \bar{y}] = \frac{\alpha}{\alpha - 1} \cdot \bar{y}$$

    Here, α measures how fat the tail of our income distribution is. We also have first order condition $\partial_\tau G(\tau) = 0$ and elasticity notation,

$$E[\partial_\tau Y \mid Y \geq \bar{y}] = -\frac{\eta}{1 - \tau} E[Y \mid Y \geq \bar{y}].$$

    where η measures how responsive reported income is to changes in the tax rate.
  - Using these, after some algebra, we get that,

$$\tau^* = \frac{1}{1 + \alpha \cdot \eta}.$$

  - Question: How do optimal tax rates depend on the amount of income inequality and on the elasticity of taxable income? We can plug in $\alpha = 2$ and $\eta = 0.25$, which are both considered reasonable estimates in the literature to obtain an optimal tax rate of $\tau^* = 0.67$.
  - In general, optimal tax rates are larger when there is more unequal distribution of income (i.e., small α and when taxable incomes are less responsive to changes in tax rates (i.e., fewer tax loopholes and better tax reinforcement).

6 Takeaways

- The main takeaway with all of this is that, as a designer, if you want to solve a social good problem, care must be taken in defining the objective function. E.g., how to measure welfare, tradeoffs, who to include, what to optimize for, and so on.

- Within redistributive taxation, for instance, we can basically rationalize almost anything using appropriately chosen welfare weights. The choice of the objective function really matters! We can also reverse engineer weights from tax schedules and use this to tell who has social power.
- Within redistributive taxation, many intellectually interesting technical questions have already been answered. However, we want to also address that there are many aspects of this that are not technical and there are more normative choices that are being made, so we have to get back to this discussion.
- Some references for redistributive taxation:
  - Saez (2001), Using elasticities to derive optimal income tax rates. (Jumpstarted the literature again in 2001 after dying down in the 70s.)
  - Chetty (2009), Sufficient statistics for welfare analysis: a bridge between structural and reduced-form methods. (Survey paper)

References

[1] Raj Chetty. Sufficient statistics for welfare analysis: a bridge between structural and reduced-form methods. *Annual Review of Economics*, 1:451–488, 2009.

[2] Maximilian Kasy. Inequality research. `http://inequalityresearch.net/`. Online; accessed February 2017.

[3] John Rawls. *A Theory of Justice*. Belknap Press, 2005.

[4] John E. Roemer. *Theories of Distributive Justice*. Harvard University Press, 1998.

[5] Emmanuel Saez. Using elasticities to derive optimal income tax rates. *Review of Economic Studies*, 68:205–229, 2001.

[6] Amartya Sen. *Inequality Reexamined*. Clarendon Press, Oxford, 1992.