An economics-based rationale for the Rawlsian social welfare program

By ODED STARK

Reprinted from: JOHN A. BISHOP and JUAN GABRIEL RODRÍGUEZ, SERIES EDITORS

RESEARCH ON ECONOMIC INEQUALITY
VOLUME 28

CHAPTER 7

AN ECONOMICS-BASED RATIONALE FOR THE RAWLSIAN SOCIAL WELFARE PROGRAM

Oded Stark

ABSTRACT

We show that a social planner who seeks to allocate a given sum in order to reduce efficiently the social stress of a population, as measured by the aggregate relative deprivation of the population, pursues a disbursement procedure that is identical to the procedure adhered to by a Rawlsian social planner who seeks to allocate the same sum in order to maximize the Rawlsian maximin-based social welfare function. Thus, the constrained minimization of aggregate relative deprivation constitutes an economics-based rationale for the philosophy-based constrained maximization of the Rawlsian social welfare function.

Keywords: Rawlsian social welfare function; Aggregate relative deprivation (ARD); Social stress; An algorithm of cost-effective policy response to ARD; Congruence of the algorithm with the Rawlsian social welfare program

JEL Classification: A13; D04; D63; H53; P51

1. INTRODUCTION

In an extensive review of John Rawls's *A Theory of Justice*, which was published 45 years ago, Kenneth Arrow raised a number of concerns. A strong common denominator of Arrow's criticisms of assumptions, certain aspects, and implications

Inequality, Redistribution and Mobility
Research on Economic Inequality, Volume 28, 179–186

ISSN: 1049-2585/doi:10.1108/10.1108/S1049-258520200000028007

for economic policy of the *Theory*, is the lack of an economics-based foundation for Rawls's *Theory*. One of many examples of that is: "My critical stance is derived from a particular tradition of thought: that of welfare economics" (Arrow, 1973, p. 246). In a second review of *A Theory of Justice*, also published 45 years ago, by Scott Gordon, the core of the criticism of the *Theory* once again was that as a prescription for achieving optimal allocation, the *Theory* is not based on economics ground rules. In the long time since the publication of Rawls's book and the critical reviews referred to above, it has not been shown that the allocation advocated by Rawls is, in fact, a mirror image of an allocation protocol that emanates from economics-based algorithms which, in themselves, arise from bricks and mortar principles of welfarism and utilitarianism.

The purpose of this chapter is to provide such a link. We unravel a novel congruence: the manner in which a social planner will allocate a given sum in order to minimize the social stress of a population, as measured by the aggregate relative deprivation, ARD, of the population, is identical to the manner in which a Rawlsian social planner will allocate the same sum in order to maximize the Rawlsian maximin-based social welfare function. Thus, the constrained minimization of ARD can be conceived as an economics-based rationale for the philosophy-based constrained maximization of the Rawlsian social welfare function. The equivalence of the two optimization procedures is illuminating. In spite of extensive attention, different interpretations, and controversies in economics and beyond related to the approach of Rawls to social welfare, a rigorous demonstration from an economics-based stance of the optimality of the Rawlsian-guided procedure of allocating or disbursing a given sum has not been provided.[1]

2. THE RAWLSIAN SOCIAL WELFARE FUNCTION

The Rawlsian approach to social welfare, built on the foundation of the "veil of ignorance,"[2] measures the welfare of a society by the wellbeing of the worst-off individual (the maximin criterion). Rawls argues that if individuals were to select the concept of justice by which a society is to be regulated without knowing their position in that society - the "veil of ignorance" - they would choose principles that involve the least undesirable condition for the worst-off member over utilitarian principles. This hypothetical contract is the basis of the Rawlsian society, and of the Rawlsian maximin social welfare function. An individual who is positioned behind the "veil of ignorance" and who does not know what particular income he will end up having, will rationally choose "conservatively" (as if he were highly risk averse), thus be inclined to "vote" for a Rawlsian social welfare function; the prospect of ending up being the worst-off member of the population looms large.

For population N consisting of n individuals whose incomes are represented by the ordered vector $x=(x_1,...,x_n)$, where $x_1 \leq x_2 \leq ... \leq x_n$, the Rawlsian social welfare function, $SWF_R(x)$, is

$$SWF_R(x) = \min_{i\in\{1,...,n\}} \{u(x_i)\},$$

where $u(x_i)$ is the utility function of individual i. This utility depends positively on individual i's income, x_i. In his writings, Rawls referred to primary goods which include basic rights and liberties, and income and wealth. It is the economists, with their strong interest in conceptualizing and measuring social welfare and income distribution, who, when reviewing the Rawlsian stance, singled out income for analysis.

3. AGGREGATE RELATIVE DEPRIVATION

We quantify the social stress of a population by the sum of the levels of social stress experienced by the individuals who constitute the population. As in Stark (2013), we measure the social stress of an individual by his relative deprivation. Also in line with the definition of relative deprivation in Stark (2013), we resort to income-based comparisons, namely an individual feels relatively deprived when others in his comparison group earn more than he does. To concentrate on essentials, we assume that the comparison group of each individual consists of all members of his population. Thus, we measure the social stress of an individual by the extra income units that others in the population have, we sum up these excesses, and we divide the sum by the size of the population. This approach, inspired by the pioneering two-volume work of Stouffer et al. (1949), tracks the seminal work of Runciman (1966) and its articulation by Yitzhaki (1979), Hey and Lambert (1980), Ebert and Moyes (2000), Bossert and D'Ambrosio (2006), and Stark, Bielawski, and Falniowski (2017). Adding together the levels of relative deprivation experienced by all the individuals belonging to a given population yields the aggregate relative deprivation (ARD) of the population. We refer to this sum as the social stress of the population.

For population N characterized in the preceding section we define the relative deprivation of individual i, RD_i, whose income is x_i as

$$RD_i \equiv \begin{cases} \dfrac{1}{n}\sum\limits_{j=i+1}^{n}(x_j - x_i) & \text{for } \ i = 1,\dots,n-1, \\ 0 & \text{for } \ i = n. \end{cases} \tag{1}$$

Multiplying and dividing the formula in the first line of (1) by the number of the individuals whose incomes are higher than the income of individual i yields an equivalent measure: RD_i is the fraction of those in the population whose incomes are higher than the income of individual i times their mean excess income. Formally, let $F(x_i)$ be the fraction of those in population N whose incomes are smaller than or equal to x_i. The relative deprivation of an individual whose income is x_i is

$$RD_i = [1 - F(x_i)] \cdot E(x - x_i \mid x > x_i). \tag{2}$$

To obtain (2), we multiply $\frac{1}{n}$ in (1) by the number of the individuals whose incomes are higher than x_i, and we divide $\sum\limits_{j=i+1}^{n}(x_j - x_i)$ in (1) by this same number. We then obtain two ratios: the first is the fraction of the population whose incomes are higher than the income of individual *i*, namely $[1-F(x_i)]$; the second is the mean excess income, namely $E(x-x_i \mid x>x_i)$. This representation of RD_i is used in the construction of the algorithm in the next section.

The aggregate relative deprivation of population *N*, ARD^N, is the sum of the levels of relative deprivation experienced by the individuals belonging to *N*,

$$ARD^N = \sum_{i=1}^{n-1} RD_i = \frac{1}{n}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left(x_j - x_i\right).$$

Remark. An alternative rewrite of (1) provides a novel interpretation of RD_i. We denote by $\tilde{x}_i \equiv \frac{1}{n-i}\sum_{j=i+1}^{n} x_j$ the average income of the individuals whose incomes are higher than the income of individual *i*. Then

$$RD_i = \frac{n-i}{n}\left[\frac{1}{n-i}\sum_{j=i+1}^{n}(x_j - x_i)\right] = \frac{n-i}{n}\left(\frac{\sum_{j=i+1}^{n} x_j}{n-i} - x_i\right) = \frac{n-i}{n}(\tilde{x}_i - x_i),$$

namely the relative deprivation of individual *i* is the product of two terms: the relative distance of his rank from the top-ranked individual (the rank of the latter individual, whose income is the highest in *N*, is 1), and the distance of individual *i*'s income from the mean income of the individuals whose incomes are higher than the income of individual *i*.

4. THE ALGORITHM OF MINIMIZING ARD

Consider population *N* of size *n* with an ordered income vector $x = (x_1,...,x_n)$, and let there be a social planner who has an amount *T* to be distributed at no cost among members of the population. We denote by Ω a subset of individuals from *N* whose incomes are the lowest. We analyze what happens when marginally, and by the same amount, we increase the incomes of the individuals in Ω, where a marginal increase refers to an increase such that the incomes of these individuals will not become higher than the income of any individual outside the set Ω.

First, suppose that the set Ω consists of just one individual out of the *n* members of the population, meaning that there is only one individual earning the lowest income; that is, $x_1 < x_i$ for $i = 2,...,n$. Suppose that the social planner appropriates a sum ε to increase the income of this lowest-earning individual (namely individual 1), where ε is small enough to satisfy our definition of a marginal increase in income; that is, $\varepsilon \le x_2 - x_1$. Using (2), this individual's relative deprivation decreases by $\frac{n-1}{n}\varepsilon$, because the mean excess income of the fraction of $\frac{n-1}{n}$ individuals earning more than him is reduced by the amount ε. At the same time, as this individual's income was, and continues to be, the lowest in the population, this expenditure does not increase the relative deprivation of any other individual belonging to N. Therefore, the change in the aggregate relative deprivation of the population is equal to the decrease in the relative deprivation of individual 1, namely

$$\Delta ARD^N = -\frac{n-1}{n}\varepsilon. \tag{3}$$

We next show that upon spending ε on a single individual, the term on the right-hand side of (3) is the highest marginal decrease in aggregate relative deprivation achievable. We do this by contradiction. Suppose that we were to increase by ε not the income of the lowest-earning individual, x_1, but, rather, the income of an individual earning $x_i > x_1$, where $i \in N$ and $i > 1$, such that $x_i + \varepsilon \leq x_{i+1}$, so as to abide by the condition of a marginal change. Then, the relative deprivation of individual i would decrease as a result of his income getting closer to the incomes of the individuals earning more than he does, but the relative deprivation of those individuals who earn less than individual i would increase. Namely when $\overline{n}_i$ ($\tilde{n}_i$) is the number of the individuals earning strictly more (less) than x_i, the change in the aggregate relative deprivation of the population would be

$$\Delta ARD^N = -\frac{\overline{n}_i}{n}\varepsilon + \frac{\tilde{n}_i}{n}\varepsilon = -\frac{\overline{n}_i - \tilde{n}_i}{n}\varepsilon, \tag{4}$$

because the mean excess income of the fraction of $\frac{\overline{n}_i}{n}$ individuals earning more than x_i would fall by the amount ε, yet, at the same time, the relative deprivation of each of the $\tilde{n}_i$ individuals earning less than x_i would increase by $\frac{\varepsilon}{n}$. Because $\tilde{n}_i \geq 1$ and $\overline{n}_i < n$, comparing (4) and (3) yields

$$\frac{\overline{n}_i - \tilde{n}_i}{n}\varepsilon < \frac{n-1}{n}\varepsilon. \tag{5}$$

Thus, channeling the transfer ε to an individual who is not the lowest income recipient in the population yields a lower decrease in aggregate relative deprivation than increasing by ε the income of the individual who earns the lowest income.

Second, we consider a population N in which there are several individuals who earn the same income which constitutes the lowest income in the population, namely the set Ω includes more than one individual. We denote by $|\Omega|$ the size of this set. Suppose again that the social planner appropriates the sum ε to increase the earnings of each member of the subset Ω by $\frac{\varepsilon}{|\Omega|}$. As in the case of a single individual who has the lowest income in the population, such a marginal transfer to each member of Ω does not change the relative deprivation of any of the individuals not belonging to Ω. Thus, the change in the aggregate relative deprivation in N arises only from a decrease of the relative deprivation sensed by the lowest-earning individuals in Ω whose incomes become closer to the incomes of the individuals earning more than they do. The fraction of the individuals in N

who earn more than members of the set Ω is equal to $\frac{n-|\Omega|}{n}$, and the mean excess income of each individual who receives the transfer is reduced by $\frac{\varepsilon}{|\Omega|}$. Therefore, each of the members of Ω experiences a decrease in his relative deprivation equal to $\frac{n-|\Omega|}{n}\frac{\varepsilon}{|\Omega|}$. With no individual in N experiencing an increase in his relative deprivation, this expenditure yields the following change in the aggregate relative deprivation

$$\Delta ARD^{N} = -|\Omega|\frac{n-|\Omega|}{n}\frac{\varepsilon}{|\Omega|} = -\frac{n-|\Omega|}{n}\varepsilon. \tag{6}$$

As in the case of the set Ω consisting of a single individual, this is obviously the optimal use of ε for any subset of individuals in the population.

Drawing on the preceding reasoning, and assuming that the marginal efficiency cost of redistribution is zero, the cost-effective response to the lowering of social stress, as measured by aggregate relative deprivation, can be represented in the form of an algorithm, as follows.

Algorithm

1. Include in the set Ω all the individuals who earn the lowest income in the population.
2. Proceed to increase simultaneously the incomes of the members of the set Ω, until (i) the amount T is exhausted, or (ii) the incomes of the members of the set Ω reach the income of the lowest-earning individual(s) who is (are) not a member (members) of this set, in which case expand the set Ω by including him (them) in Ω. If condition (i) is met, then the procedure is completed. If condition (ii) is met, start from step 1 once again. Notice that the incomes of the pre-expansion members of the set Ω should be increased from the level already reached, that is, from the level equal to the income(s) of the individual(s) newly included.

It is easy to ascertain the optimality of the protocol of the algorithm: at each step, we increase the incomes of those individuals who earn the least, so the decrease in the aggregate relative deprivation of the population is most effective, and no one experiences an increase of their relative deprivation in the process. Increasing the incomes from below is ratcheted up through the hierarchy of the individuals, and it ceases when the funds available for reducing the aggregate relative deprivation are exhausted.

5. CONGRUENCE OF THE ALGORITHM WITH THE RAWLSIAN SOCIAL WELFARE PROGRAM

Let there be a population of five individuals whose incomes are (1, 2, 3, 4, 5). Let there be two social planners who are constrained not to reduce the incomes of members of the population.

Consider first a Rawlsian social planner who seeks to increase social welfare by adhering to the maximin principle, and who has at his disposal three units of income. This planner will allocate the first unit of income to the individual with the lowest utility, that is, to the individual whose income is 1; the income vector will then become (2, 2, 3, 4, 5). Thereafter, the Rawlsian social planner will reach out to the now worst off, namely to the two individuals whose incomes are 2 each, and increase the incomes of each of these two individuals to 3, thereby obtain an income vector (3, 3, 3, 4, 5). Clearly, as the allocation proceeded, the identity of the worst-off individuals changed (first it was the individual whose income was initially 1, then these were the two individuals whose incomes were initially 1 and 2). As is easily seen, the principle guiding the allocation of the income available for disbursement is to attend to the individuals from the bottom up.

A social planner who seeks to *minimize aggregate relative deprivation*, ARD, will allocate the three units of income in exactly the same way; this was proved in the algorithm presented in the preceding section. Given that proof, it follows then that the constrained minimization of ARD can be construed as an economics-based rationale for administering the philosophy-based Rawlsian maximin-based social welfare program.

Naturally, there is a difference between the Rawlsian procedure and the ARD protocol in that the *reasons* for proceeding from the bottom up are not the same. This difference notwithstanding, a Rawlsian social planner with a "policy budget" of three units of income allocates these units in the very same way as a social planner who applies the ARD minimization protocol.

The configuration presented in this section is not limiting; the same procedure applies to larger populations and to larger sums available for allocation.

6. CONCLUSION

Even if a translation into mainstream economics of the voluminous work of Rawls can ingeniously take a one-line equation (as in Section 2 above), a task that has not as yet been met is to provide an economics-based rationale for the income allocation protocol chartered by the Rawlsian social welfare function. In policy formation, public choice, and public economics there is a menu of social welfare functions to choose from (including utilitarian, Bernoulli-Nash, and Rawlsian), and unless a particular condition holds under which the social planners of all stripes see eye to eye, each social welfare function can give rise to a different rule of income allocation and, when enacted, lead to different income distributions.[3] In view of the general importance of choosing which social function to follow, it is necessary to have in place the underlying economics rationales. Hence the inquiry undertaken here.

It is an intriguing finding that the pursuit of the Rawlsian social welfare program is equivalent to a cost-effective treatment of social stress. Because the direction taken in this chapter can be reversed, the reported congruence can be interpreted as the Rawlsian program providing a social welfare rationale, as well as a procedure, for addressing social stress.

NOTES

1. Notable examples of criticisms of and controversies related to the approach of Rawls to social welfare are Harsanyi (1975), and Sen (2009); see also the "Symposium on The Idea of Justice" in the 2011, Vol. 5 issue of the *Indian Journal of Human Development*.

2. "[N]o one knows his place in society, his class position or social status; nor does he know his fortune in the distribution of natural assets and abilities, his intelligence and strength, and the like." (Rawls, 1999, p. 118).

3. Stark, Falniowski, and Jakubek (2017) provide a condition under which the utilitarian, Rawlsian, and Bernoulli-Nash social planners come up with the same optimal income distribution when a tax and transfer procedure is subject to a deadweight loss. Stark, Falniowski, and Jakubek (2017) further show that when the individuals' utility functions exhibit a sufficiently high concern at having a low relative income, the optimal tax policies of all the social planners align. They characterize the consensus optimal income distribution - which is a distribution of equal incomes - and find that the intensity of the individuals' concern at having a low relative income crowds out the preferences over income distribution harbored by particular social planners.

REFERENCES

Arrow, K. J. (1973). Some ordinalist-utilitarian notes on Rawls's Theory of Justice by John Rawls. *The Journal of Philosophy*, *70*(9), 245–263.

Bossert, W., & D'Ambrosio, C. (2006). Reference groups and individual deprivation. *Economics Letters*, *90*(3), 421–426.

Ebert, U., & Moyes, P. (2000). An axiomatic characterization of Yitzhaki's index of individual deprivation. *Economics Letters*, *68*(3), 263–270.

Gordon, S. (1973). John Rawls's difference principle, utilitarianism, and the optimum degree of inequality. *The Journal of Philosophy*, *70*(9), 275–280.

Harsanyi, J. (1975). Can the Maximin Principle serve as the basis for morality? A critique of John Rawls's theory. *American Political Science Review*, *69*, 594–606.

Hey, J. D., & Lambert, P. J. (1980). Relative deprivation and the Gini coefficient: Comment. *Quarterly Journal of Economics*, *95*(3), 567–573.

Rawls, J. (1999). *A theory of justice*. Cambridge, MA: Harvard University Press.

Runciman, W. G. (1966). *Relative deprivation and social justice*. Berkeley: University of California Press.

Sen, A. (2009). *The idea of justice*. London: Penguin Books.

Stark, O. (2013). Stressful integration. *European Economic Review*, *63*, 1–9.

Stark, O., Bielawski, J., & Falniowski, F. (2017). A class of generalized measures of relative deprivation. *Economics Letters*, *160*, 105–110.

Stark, O., Falniowski, F., & Jakubek, M. (2017). Consensus income distribution. *Review of Income and Wealth*, *63*(4), 899–911.

Stouffer, S. A., Suchman, E. A., DeVinney, L. C., Star, S. A., & Williams, R. M., Jr. (1949). *The American soldier: Adjustment during army life*, Vol. I. Stouffer, S. A., Lumsdaine, A. A., Lumsdaine, M. H., Williams, R. M., Jr., Smith, B. M., Janis, I. L., Star, S. A., & Cottrell, L. S., Jr. *The American soldier: Combat and its aftermath*, Vol. II. Studies in Social Psychology in World War II. Princeton, NJ: Princeton University Press.

Yitzhaki, S. (1979). Relative deprivation and the Gini coefficient. *Quarterly Journal of Economics*, *93*(2), 321–324.