

Utilitarian Preference for Redistribution: A Concern with Max-Min

George A. Waters*
Department of Economics
Campus Box 4200
Illinois State University
Normal, IL 61761-4200

December 1, 2016

Abstract

This paper studies an optimal linear tax or basic income regime with an extensive margin in the supply of labor and a utilitarian welfare criterion. The presence of the extensive margin in labor supply breaks the connection between allocational efficiency and aggregate utility maximization. In a general framework, some redistribution maximizes aggregate utility, but a lesser degree than what Rawls' max-min criterion would require. Rawls' concept of the Original Position is useful, but the rationality of the max-min criterion is questionable.

Keywords: basic income, Original Position, max-min, redistribution

*gawater@ilstu.edu

1 Introduction

Though the thought of sacrificing the well-being of the least fortunate in a group is repugnant to most, the response to *The Lottery* [Jackson, 1948] being a prominent example, developing a rigorous method for determining a desirable degree of redistribution is not trivial. The introduction of heterogeneity into macroeconomic models raises important questions about the proper welfare criterion and the definition of efficiency. The goal of the present paper is to develop a simple model of redistributive taxation and examine the associated political economy issues, particularly Rawls' conceptions of the Original Position and max-min welfare criterion.

The present work examines an environment where wages are heterogeneous due to differences in worker productivity, and the government collects income taxes via a flat marginal rate and redistributes lump sum transfers to all agents, i.e. a negative income tax or basic income plan. There is an extensive margin in labor supply so the income tax may drive some low productivity workers out of the labor force.

The primary contribution of the paper is the interpretation of a formal comparison of optimal redistribution regime under utilitarian (Benthamite) welfare criteria and the Rawlsian max-min criterion with minimal assumptions about household preferences. Rawls [1971] advocates for the max-min criterion, where the outcome of the least fortunate member of society is maximized. He argues that the max-min criterion is the rational response of a person behind the "veil of ignorance" a key aspect of the Original Position¹, meaning she does not know the realization of personal circumstances when deciding fair and just social relations. In the present environment, the question becomes one of the optimal tax regime when households do not know their level of productivity.

While most of the preliminary results have counterparts in the existing literature, the most important result is Proposition 6, which shows that the degree of redistribution satisfying the max-min criterion is necessarily greater than the one that maximizes aggregate utility in an environment with minimal assumptions about preferences. The model formalizes the objection of some philosophers, such as Nagel (1973) and Scanlon (1973), to the max-min criterion that rational agents behind the veil might be willing to risk a bad outcome. Rawls' Original Position is a valuable concept, but his argument about the rationality of the max-min welfare criterion is suspect.

It is common to define welfare using a concave transformation across utilities, where the concavity represents a societal preference for equality above that given by the utilitarian criterion, for example Stiglitz (1985). With such a transformation, the max-min criterion coincides with welfare maximization in a limiting case with respect to the concavity of the transformation.

¹Rawls [1958] describes the concept. Harsanyi [1955] discusses a formal representation.

We eschew such transformations representing a desire for equality beyond what is implied by diminishing marginal utility. Whether society-wide agreement about a preference for equality is possible or meaningful is questionable, and the associated transformations are necessarily arbitrary. Though some undoubtedly have such a preference, subjective elements of a model limit its relevance. In the present environment without such a transformation, where welfare is simply the sum or expected value of the utilities across the population, the max-min criterion coincides with the tax rate at the peak of the Laffer Curve.

The basic income regime is not subject to some common criticisms of redistribution. Under the linear (flat rate) tax, workers do not have incentive to take lower wage jobs to avoid taxes. Also, the lump sum transfer is given to all so there is no distortion of labor supply decisions. For decreasing marginal utility, redistribution is desirable since the gain to low wage workers outweighs the loss for high wage workers.

The presence of an extensive margin in labor supply breaks the equivalence between welfare maximization and the usual allocational efficiency condition. In the present model, the usual allocational efficiency condition is obtained when the tax rate is zero and there is no redistribution as long as household choose a positive labor supply at any wage, a questionable assumption.

Individually, the elements of the present framework are not novel, but the combination and the focus on the interpretation of the welfare criteria is. These results agree with those found in the standard optimal linear tax model, summarized in Piketty and Saez (2013), but the model here relies only on standard household utility preferences, not on diminishing marginal welfare across utilities on the part of the social planner². Optimal flat rate or linear tax regimes are studied in Stiglitz [1985] and Dixit and Sandmo [1977]. The extensive labor supply margin is included in the models in Mirlees [1971] and Diamond [1980]. Atkinson [1973] analyzes optimal taxation under the max-min criterion. Manski [2014] also considers simple tax regimes in a model with government production that affects consumption and productivity. He discusses informally some of the issues around efficiency in the present work.

The paper is organized as follows. Section 2 describes the elements of the model and some of the desirable properties of the basic income regime. Section 3 discusses the formalization of the welfare criteria. Section 4 reports the primary formal results of the paper related to optimal tax rate. Sections 5 and 6 introduce a particular calibration of the model to show the quantitative relevance of the results and their relation to the literature on income distribution. Section 7 concludes.

²Also, the model presented here has an explicit production sector, unlike some of the models described in Piketty and Saez [2013].

2 The Model

The key elements of the model are heterogeneous productivity and the extensive margin in labor supply. A unit mass of households indexed by i have heterogeneous productivity a_i with single-peaked density $a_i \sim f(a_i) \in (0, \infty)$ and associated distribution $F(a_i)$. Labor markets are competitive so real wages w_i are determined by productivity such that $w_i = a_i$. Labor is the only input in the model. Each household supplies quantity of labor h_i which produces output $a_i h_i$, the production function for a single input with constant returns to scale. Aggregate production Y is thereby

$$Y = \int_0^{\infty} a_i h(a_i) dF(a_i). \quad (1)$$

Households receive wage income, pay taxes at a marginal rate of τ and receive a lump sum transfer $T = \tau Y$. The household budget constraint is

$$c(a_i) = a_i h(a_i) (1 - \tau) + \tau Y \quad (2)$$

where $c(a_i)$ is household consumption. Both consumption and hours worked are functions of an individual's productivity a_i . Since there is a unit mass of agents, the values $c(a_i)$, $h(a_i)$, Y and T can be interpreted as per capita values. For positive tax rates $\tau > 0$, all households have a minimum income τY , which requires redistribution in that taxes reduce income for workers with sufficiently high productivity, implementing a basic income plan or negative income tax, advocated by Milton Friedman³ among others.

Households maximize utility $U^i(c(a_i), h(a_i))$ that is a function of consumption and leisure. We make minimal assumptions about the form of the utility function: $U_1^i > 0$, $U_{11}^i \leq 0$, $U_2^i \leq 0$, $U_{22}^i \geq 0$.

The pair $(c^*(a_i), h^*(a_i))$ satisfies the optimality condition determined by maximizing $U^i(c(a_i), h(a_i))$ subject to the household budget (2).

$$U_1^i(c^*(a_i), h^*(a_i)) a_i (1 - \tau) = -U_2^i(c^*(a_i), h^*(a_i)) \quad (3)$$

This condition shows the standard labor-leisure trade-off. Under a representative agent, allocational efficiency, where utility is maximized subject to the production constraints, is equivalent to the labor-leisure optimality condition without redistribution.

Lemma 1 *Given a representative agent with productivity \bar{a} , the efficiency condition where utility $U(c(\bar{a}), h(\bar{a}))$*

³Friedman [1962] presents the negative income tax as a fixed deduction from wages that are taxed, which is equivalent to the regime here for any given tax rate. See also <http://www.basicincome.org/bien/>.

is maximized subject to the production constraints $Y = \bar{a}h$ and $c = Y$, is equivalent to the household optimality condition (3) where $\tau = 0$.

Proof. The first order conditions for the Lagrangian

$$L = U(c(\bar{a}), h(\bar{a})) + \lambda(\bar{a}h - c)$$

are $U_1 = \lambda$ and $U_2 = -\lambda\bar{a}$. Solving out the Lagrange multiplier λ yields the household optimality condition (3) for $a_i = \bar{a}$ and $\tau = 0$. ■

However, for the above result to hold under heterogeneous productivity and wages requires an interior solution for all $a_i \in (0, \infty)^4$, which means that workers work for any positive wage. Hence, for the above condition (3) to apply in the neighborhood of $a_i = 0$, it must also be the case that the marginal disutility of labor is zero at that point, $U_2^i(c(0), 0) = 0$ and strictly negative $U_2^i(c(a_i), h(a_i)) < 0$ for $a_i > 0$.

Remark 2 A necessary condition for efficiency with heterogeneous wages is a positive labor supply for any positive wage.

Naturally, the assumption that worker would never drop out of the labor market is questionable. If that assumption does not hold there is an extensive margin in labor supply. Working negative hours makes little sense, so let the labor supply function be

$$h(a_i) = \max(0, h^*(a_i)), \tag{4}$$

which, along with the budget constraint (2), determines the consumption function $c(a_i)$. The bound on hours implies that there is a productivity threshold below which workers drop out of the labor market and their only income is the transfer from the government.

Definition 3 The productivity threshold \tilde{a} is such that $h^*(a_i) \leq 0$ for all $a_i \leq \tilde{a}$.

There are a number of additional reasons to focus on a tax and redistribution system of the present form. The transfer applies to all agents so there is no need for means-testing, which is costly and raises the potential for some to misrepresent their level of productivity. It also avoids problems of envy found in the optimal tax literature and most prominently in the model with two labor types in Stiglitz [1982]⁵. This problem of envy does not arise in the flat-tax, lump-sum redistribution regime.

⁴Having an interior solution is not a sufficient condition for welfare to be maximized at $\tau = 0$, but Stiglitz (2009) shows that it is the case for a lognormal distribution of wages in the absence of public goods.

⁵See Stiglitz (1987) and Piketty and Saez (2013) for a full discussion.

Proposition 4 Household utility $U^i(c(a_i), h(a_i))$ is non-decreasing in a_i for any tax rate τ .

Proof. See Appendix A. ■

The tax-transfer regime under consideration here does not lead to envy among high productivity (wage) workers, and the lump sum transfer does not distort labor markets as noted above. Furthermore, it does not require much sophistication on the part of the government to implement. Hence, this tax-transfer specification is one of the least objectionable redistribution schemes.

Not to say that taxes do not have adverse effects. On the upward sloping portion of the labor supply curve the threshold \tilde{a} is increasing in τ , so as the tax rate increases workers drop out of the labor market.

3 Welfare

Welfare is defined without an explicit preference for equality. Aggregate utility is one obvious choice of welfare criterion to judge the effect of redistribution, following most of the optimal tax literature including Mirlees [1971], Manski [2012] and Werning [2007]. The Rawlsian approach to welfare is treated separately.

Aggregate utility U^e is specified by aggregating household utility over the distribution of productivity $F(a_i)$ equivalent to expected utility.

$$\begin{aligned} U^e &= \int_0^\infty U^i(c(a_i), h(a_i)) dF(a_i) \\ &= \int_0^{\tilde{a}} U^i(\tau Y, 0) f(a_i) + \int_{\tilde{a}}^\infty U^i(c(a_i), h(a_i)) dF(a_i) \end{aligned} \quad (5)$$

The second line expresses separately the utility of those with productivity below \tilde{a} who do not work and who consume only the transfer τY .

Unlike many models in this literature, we do not include an arbitrary weighting function representing preferences the government or social planner. Such models introduce a concave function $G(\cdot)$ and define welfare as $\int_0^\infty G(U^i) dF(a_i)$. Concavity of the weighting function $G(\cdot)$ indicates diminishing marginal social welfare across the utilities of the population with the limiting case $G'(0) = \infty$ representing the Rawlsian preference to maximize the utility of the least productive agent. While such an approach is mathematically elegant, the weighting function is arbitrary, and the results that follow could be questioned.

One could also consider aggregate utility U^e to be the utility of a representative agent behind the "veil of ignorance," meaning he does not know his own draw from the distribution $F(a_i)$. Rawls uses this concept in his development of principles of justice from the "Original Position", which would be agreed upon by citizens behind the veil. Werning [2007] considers aggregate utility to be a representation of insurance for

households against poor outcomes.

Rawls' concept is a "thick veil", where agents have no knowledge of the structure of the economy, which leads him to argue in favor of the stronger max-min criterion, which requires that the utility of the least advantaged agent is maximized. In the present context with the extensive margin in labor supply, the max-min principle is satisfied by the tax rate at the peak of the Laffer curve⁶, which relates the transfers (government revenue) $T(\tau)$ and the tax rate τ . The transfer is also the only income of the least productive worker at $a \approx 0$, so the tax rate that maximizes $T(\tau)$ satisfies the max-min principle.

The present work considers the role of redistribution on welfare for both the max-min criterion and aggregate utility. The following derivative shows the effect of redistributive taxation on aggregate utility. The calculation involves multiple applications of Leibniz Rule, since the threshold \tilde{a} depends on the tax rate τ . Also note that $F'(a) = f(a)$.

$$\begin{aligned} \frac{dU^e}{d\tau} = & U_1^i(\tau Y, 0) \left(Y + \tau \frac{dY}{d\tau} \right) F(\tilde{a}) + U^i(\tau Y, 0) f(a) \frac{d\tilde{a}}{d\tau} \\ & + \int_{\tilde{a}}^{\infty} \left[U_1^i(c(a_i), h(a_i)) \frac{dc_i}{d\tau} + U_2^i(c(a_i), h(a_i)) \frac{dh_i}{d\tau} \right] dF(a_i) - U^i(c(\tilde{a}), h(\tilde{a})) f(a) \frac{d\tilde{a}}{d\tau} \end{aligned}$$

The effects of an increase in the marginal tax rate τ expressed above are intuitive. The $\frac{d\tilde{a}}{d\tau}$ show the effect of workers exiting the workforce, the first term shows the increased utility of those receiving only the transfer, and the integral term shows the change in utility for those in the workforce.

The definition of \tilde{a} (Definition 3) implies that for $a_i < \tilde{a}$, it is the case that $h(a_i) = 0$ and $c(a_i) = \tau Y$. Hence, the $\frac{d\tilde{a}}{d\tau}$ terms cancel and the above expression simplifies.

$$\frac{dU^e}{d\tau} = U_1^i(\tau Y, 0) \left(Y + \tau \frac{dY}{d\tau} \right) F(\tilde{a}) + \int_{\tilde{a}}^{\infty} \left[U_1^i(c(a_i), h(a_i)) \frac{dc_i}{d\tau} + U_2^i(c(a_i), h(a_i)) \frac{dh_i}{d\tau} \right] dF(a_i)$$

Using the household optimality condition (3) and differentiating the household budget constraint (2) with respect to τ and substituting yields an alternative form for $\frac{dU^e}{d\tau}$.

$$\frac{dU^e}{d\tau} = U_1^i(\tau Y, 0) \left(Y + \tau \frac{dY}{d\tau} \right) F(\tilde{a}) + \int_{\tilde{a}}^{\infty} U_1^i(c(a_i), h(a_i)) \left[Y - a_i h(a_i) + \tau \frac{dY}{d\tau} \right] dF(a_i) \quad (6)$$

4 Redistribution

Some degree of redistribution maximizes aggregate utility given minimal assumptions. However, the aggregate utility maximizing tax rate is less than the rate that satisfies the Rawlsian welfare criterion.

⁶Piketty and Saez [2013] note that this concept can be traced to Jules Dupuit.

Another form of the derivative $\frac{dU^e}{d\tau}$ is obtained by combining the two terms in the relation (6), using the fact that the first term does not depend on a_i for $a_i \in (0, \tilde{a})$.

$$\frac{dU^e}{d\tau} = \int_0^\infty U_1^i(c(a_i), h(a_i)) \left[Y - a_i h(a_i) + \tau \frac{dY}{d\tau} \right] dF(a_i) \quad (7)$$

A mathematical trivial but important special case is that of constant marginal cost $U_1^i(\cdot) = \bar{c}$, where the above reduces to $\frac{dU^e}{d\tau} = \bar{c} \tau \frac{dY}{d\tau}$. Output is decreasing in the tax rate, so $\frac{dU^e}{d\tau} \leq 0$. Further, at the zero tax rate it is the case that $\frac{dU^e}{d\tau}|_{\tau=0} = 0$ under constant marginal utility so the introduction of redistribution does not improve welfare. Hence, any argument in favor of redistribution depends on decreasing marginal utility. Of course, assuming constant marginal utility across all income levels is highly questionable.

For realistic specifications of the utility function, some degree of redistribution is desirable. Using the equilibrium condition⁷ $Y = \int_0^\infty c(a_i) dF(a_i)$ gives the following.

Proposition 5 *Given that marginal utility is decreasing in $a_i \in (0, \infty)$, $U_{11}^i(c(a_i), h(a_i)) < 0$, consumption is increasing in productivity, $c'(a_i) > 0$ and there is an extensive labor supply margin $\tilde{a} > 0$ then some degree of redistribution increases aggregate utility such that $\frac{dU^e}{d\tau}|_{\tau=0} > 0$.*

Proof. See Appendix A. ■

For redistributive taxation to have any benefit, the marginal utilities and mass of agents whose productivity (and wage) places them below mean income must be greater than those above. Even with the restrictions of a fixed marginal tax and lump sum transfer for all, redistribution is beneficial, as in Mirlees [1971]. The assumption that consumption is increasing in productivity is equivalent to the assumption that it is a normal good. It might be violated for an extremely backward bending labor supply curve at higher levels of productivity, but the case for redistribution would be even stronger here. Workers on the backward bending portion of the labor supply curve would supply more labor at a higher tax rate.

The existence of an extensive margin in labor supply $\tilde{a} > 0$ means that the result in Lemma 1 where efficiency is achieved without redistribution does not apply. This result clouds all arguments that rely on such a definition of efficiency. While some concept of efficiency properly defined may be useful, the role of heterogeneity is crucial.

Now that it has been established that a positive tax rate maximizes aggregate utility, the primary result that the Rawls' max-min criterion calls for an even greater degree of redistribution can be shown.

Proposition 6 *Given that the unique tax rate that maximizes aggregate utility U^e is τ^* , and that the rate*

⁷This condition can be recovered by integrating the household budget constraint as well.

that satisfies the max-min criterion is $\bar{\tau}$, then

$$\tau^* < \bar{\tau}.$$

Proof. The household optimality condition (3) and labor supply specification (4) implies that if $\tau = 1$, $h(a_i) = 0$ for all a_i . Hence, for τ sufficiently large, output Y is decreasing with τ , and the Laffer curve has its usual shape where $T(\tau)$ has a single maximum and $T(0) = T(1) = 0$.

The proof proceeds by contradiction. Assume $\tau^* > \bar{\tau}$. Since, both functions of the tax rate $U^e(\tau)$ and $T(\tau)$ have a unique maximum on $\tau \in [0, 1]$, it follows that $\frac{dU^e}{d\tau}|_{\tau=\bar{\tau}} > 0$ and $\frac{dT}{d\tau}|_{\tau=\bar{\tau}} = 0$. The latter equality implies $(Y + \tau \frac{dY}{d\tau})|_{\tau=\bar{\tau}} = 0$. Combining these relations with the expression (7) yields

$$U_1^i(c(a_i), h(a_i)) \left[Y - a_i h(a_i) + \tau \frac{dY}{d\tau} \right] dF(a_i) |_{\tau=\bar{\tau}} > 0,$$

but this can be written

$$-U_1^i(c(a_i), h(a_i)) a_i h(a_i) dF(a_i) |_{\tau=\bar{\tau}} + U_1^i(c(a_i), h(a_i)) \left[Y + \tau \frac{dY}{d\tau} \right] dF(a_i) |_{\tau=\bar{\tau}} > 0.$$

Since $(Y + \tau \frac{dY}{d\tau})|_{\tau=\bar{\tau}} = 0$, $U_1^i(\cdot) > 0$ and all the other terms in the first integral are positive, the above relation cannot be true. Therefore, it must be the case that $\tau^* < \bar{\tau}$ as required. ■

In an economic environment with minimal restrictions, the aggregate utility maximizing degree of redistribution is less than that determined by the max-min principle⁸. Alternatively, satisfying the max-min principle would require a reduction in aggregate utility as compared to the optimum. The proof makes the reason clear. Considering the expression for aggregate utility (5), the max-min principle requires maximizing the first integral over $[0, \tilde{a}]$ while ignoring the second integral over $[\tilde{a}, \infty]$, meaning the utility of the least productive is maximized without regard to the effect on the rest of the population.

The results that the Rawlsian marginal tax is revenue maximizing but not welfare maximizing agree with results from the standard optimal linear tax model in Picketty and Saez (2013) that includes social welfare weighting function. Proposition 6 also parallels results in Werning [2007], which derives conditions for Pareto efficient income tax regimes. The resulting income distribution for such a regime must stochastically dominate the distribution that satisfies the Rawlsian (max-min) criterion for social welfare.

Proposition 6 formalizes the objection to the max-min criterion that, even for someone deciding behind the veil of ignorance, a person might be willing to risk a poor outcome if the probability is sufficiently

⁸In related, though more restrictive, models, Stiglitz [1987] and Picketty and Saez [2013] demonstrate a simple formula for level of taxation that maximizes aggregate utility.

small. Rawls [1971] argues that such a risk would not be rational but is not wholly convincing, as argued by Nagel [1973] and Scanlon [1973]. The view embodied in the present model is that the Original Position is a useful concept for evaluating policy and societal norms. However, the max-min criterion is interesting as a benchmark, but relying on it in arguments about optimal policy is suspect.

A defender of the max-min criterion would argue that the veil in the above model is not sufficiently thick in that agents have knowledge about their preferences and the structure of production in the economy. Though the assumptions about the utility function are quite general, it is true that the production sector is stylized. However, there is no obvious extension that would overturn the result in Proposition 6, though this is an area for future research. More importantly, if the goal is to provide a rigorous framework for fairness, redistribution can be justified on utilitarian grounds with a thin veil, as in the model presented here, and there is no need to discuss the thickness of the veil, which is inherently vague and unlikely to be convincing for many.

5 Conclusion

There is no welfare criterion that is above criticism. It is more important to select a criterion that is reasonable across a wide range of preferences, than one that fits closely with a particular viewpoint. The utilitarian criterion (maximizing aggregate utility) does not capture all aspects of well-being, but has the advantage of simplicity. Diminishing marginal utility is not arbitrary. In identifying an appropriate welfare criterion, Rawls' concept of the Original Position is useful in that it provides a foundation for the meaning of aggregate utility as expected utility behind the veil. However, his assertion that the max-min criterion is rational is suspect. Under minimal assumptions about household preferences the includes an extensive margin for labor supply, the marginal tax rate that satisfies the max-min criterion is higher than that which maximizes aggregate utility. The max-min criterion ignores the possibility that a rational agent behind the veil might be willing to risk an outcome worse than the maximum possible transfer under the basic income policy regime. Furthermore, with a heterogeneous population, the aggregate utility maximizing tax-transfer regime does not correspond to common formulations of allocational efficiency. Researchers studying heterogeneous populations should continue to focus on aggregate utility, making sure that any efficiency concept is rigorously defined.

Appendix A: Proofs

Proof of Proposition 4:

Proof. From the derivative

$$\frac{dU_i}{da_i} = U_1^i(c^*(a_i), h^*(a_i)) c'(a_i) + U_2^i(c^*(a_i), h^*(a_i)) h'(a_i)$$

and the household optimization condition (3), one can derive the following.

$$\frac{dU_i}{da_i} = U_1^i(c^*(a_i), h^*(a_i)) h(a_i) (1 - \tau)$$

The above expression is non-negative so U_i is non-decreasing in a_i for any $a_i > \tilde{a}$. For $a_i \leq \tilde{a}$, household do not work and consume only the transfer, so $\frac{dU_i}{da_i} = 0$ for them. ■

Proof of Proposition 5:

Proof. Let the level of productivity \hat{a} be such that $c(\hat{a}) = Y$. From the equilibrium condition above, we can write

$$\int_{\tilde{a}}^{\hat{a}} [Y - c(a_i)] dF(a_i) + \int_{\hat{a}}^{\infty} [Y - c(a_i)] dF(a_i) = \int_{\tilde{a}}^{\infty} [Y - c(a_i)] dF(a_i) = 0$$

so that the first term is positive and the second term is negative. Since marginal utility U_1^i is decreasing in a_i ,

$$\int_{\tilde{a}}^{\hat{a}} U_1^i(c(a_i), h(a_i)) [Y - c(a_i)] dF(a_i) > U_1^i(c(\hat{a}), h(\hat{a})) \int_{\tilde{a}}^{\hat{a}} [Y - c(a_i)] dF(a_i)$$

and

$$\int_{\hat{a}}^{\infty} U_1^i(c(a_i), h(a_i)) [Y - c(a_i)] dF(a_i) > U_1^i(c(\hat{a}), h(\hat{a})) \int_{\hat{a}}^{\infty} [Y - c(a_i)] dF(a_i).$$

Therefore, adding inequalities and using established identities yields the desired result, noting that $U_1^i(0, 0) Y F(\tilde{a}) \geq 0$.

$$\begin{aligned} \frac{dU^e}{d\tau} |_{\tau=0} &= U_1^i(0, 0) Y F(\tilde{a}) + \int_{\tilde{a}}^{\infty} U_1^i(c(a_i), h(a_i)) [Y - c(a_i)] dF(a_i) \\ &= U_1^i(0, 0) Y F(\tilde{a}) + \int_{\tilde{a}}^{\hat{a}} U_1^i(c(a_i), h(a_i)) [Y - c(a_i)] dF(a_i) \\ &\quad + \int_{\hat{a}}^{\infty} U_1^i(c(a_i), h(a_i)) [Y - c(a_i)] dF(a_i) \\ &> U_1^i(0, 0) Y F(\tilde{a}) + U_1^i(c(\hat{a}), h(\hat{a})) \int_{\tilde{a}}^{\hat{a}} [Y - c(a_i)] dF(a_i) \\ &\quad + U_1^i(c(\hat{a}), h(\hat{a})) \int_{\hat{a}}^{\infty} [Y - c(a_i)] dF(a_i) \\ &= U_1^i(0, 0) Y F(\tilde{a}) + U_1^i(c(\hat{a}), h(\hat{a})) \int_{\tilde{a}}^{\infty} [Y - c(a_i)] dF(a_i) > 0. \end{aligned}$$

■

References

- Atkinson, Anthony B. 1973 Minimax and Optimal Income Taxation. Discussion Paper no. 47, University of Essex Department of Economics.
- Diamond, Peter 1980 Income Taxation with Fixed Hours of Work, *Journal of Public Economics* 13, 101-110
- Dixit, Anivash and Sandmo 1977 Some Simplified Formulae for Optimal Income Taxation, *The Scandinavian Journal of Economics* 79(4), 417-423.
- Harsanyi, John 1955 Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility, *Journal of Political Economy* 63(4), 309-321.
- Jackson, Shirley 1948. The Lottery. in *the New Yorker* (June 26).
- Manski, Charles 2014 Choosing Size of Government under Ambiguity: Infrastructure Spending and Income Taxation, *The Economic Journal* 124(576), 359-376.
- Mirlees, James A., 1971, An exploration in the theory of optimal income taxation, *Review of Economic Studies* 38, 135-208.
- Nagel, Thomas 1973 Rawls on Justice, *The Philosophical Review* 82(2), 220-234.
- Piketty, Thomas and Saez, Emmanuel 2013 "Optimal Labor Income Taxation" with Thomas Piketty, in the *n*, Volume 5, 391-474
- Rawls, John 1958 Justice as Fairness. *The Philosophical Review*, Vol. 67(2), 164-194.
- Rawls, John 1971 *A Theory of Justice*, Harvard University Press, Cambridge, MA.
- Scanlon, Thomas, 1973, Rawls' Theory of Justice, in *Reading Rawls*, edited by Normal Daniels, 1989, Basic Books, New York, NY.
- Stiglitz, Joseph 1982 "Self-selection and Pareto Efficient Taxation." *Journal of Public Economics* 17, 213-240.
- Stiglitz, Joseph 1987. "Pareto efficient and optimal taxation and the new welfare economics," in: A. J. Auerbach and M. Feldstein (ed.), *Handbook of Public Economics*, volume 2, chapter 15, 991-1042. Amsterdam: North-Holland.
- Stiglitz, Joseph 2009. "Simple Formulae for Optimal Income Taxation and the Measurement of Inequality: An Essay in Honor of Amartya Sen." Oxford online scholarship, DOI: 10.1093/acprof:oso/9780199239115.001.0001.
- Werning, Ivan 2007 Pareto Efficient Income Taxation, working paper.