

INEQUALITY, SOCIAL INSURANCE
AND REDISTRIBUTION

Michael Wallerstein and Karl Ove Moene

Estudio/Working Paper 1999/144
November 1999

Michael Wallerstein is Professor of Political Science and Chairman of the Department of Political Science, Northwestern University. This paper is based on a seminar that he presented at the Center for Advanced Study in the Social Sciences, Juan March Institute, Madrid, on 16 March 1999, entitled "Inequality, Redistribution and Social Insurance".

Inequality, Social Insurance and Redistribution*

Abstract

Is the political support for welfare policy higher or lower in less egalitarian societies? We answer the question using a framework in which welfare policies are modelled as publicly financed insurance policies that pay benefits in a redistributive manner. With the majority of voters having both redistributive and insurance motives for supporting welfare spending, the impact of inequality on voters' support for welfare depends on how benefits are targeted. If benefits are mostly paid to persons who are employed, redistribution dominates and greater inequality increases support for welfare spending. If benefits are mostly targeted to those without earnings, however, the insurance aspect dominates and greater inequality reduces support for welfare spending. When voters choose both the level and the targeting of benefits, we find that the insurance aspect dominates in the sense that greater inequality is associated with lower benefits targeted to those without earnings. This prediction is contrary to the dominant view in the literature but, as we show, consistent with the evidence among advanced industrial societies.

* We thank Steve Davis, David Austen-Smith and Daniel Diermeier for helpful comments. We thank the MacArthur Foundation and the Norwegian Research Council for financial support.

I. Introduction

What is the effect of changes in the inequality in pre-tax and transfer income on political support for welfare policy? The conventional intuition that is expressed in almost all economic models of welfare policy starting with Romer (1975), Roberts (1977) and Meltzer and Richard (1981), is that increased inequality leads to greater political demand for redistributive policies. The logic is simple and compelling. Welfare policy, it is assumed, redistributes income from persons with above-average incomes to persons with below-average incomes. Given that the income distribution is skewed such that a majority of the population receives an income below the mean, it is natural to think that an increase in inequality that causes above-average incomes to rise and below-average incomes to fall would provoke to increasing demands for public policies to reduce the gap between the rich and the poor. In the Romer-Roberts-Meltzer and Richard formalization, political competition drives the level of welfare spending towards the ideal point of the voter with median income. The greater the distance between the median and the mean income, the greater the equilibrium level of welfare spending.

This simple model of voting over redistributive policies has several comforting implications. If greater equality reduces the political demand for redistributive policies and redistributive policies inhibit growth, reducing income inequality promotes economic development as argued by Alesina and Rodrik (1994) and Persson and Tabellini (1994). In addition, if an exogenous change in the economic environment causes the distribution of income to grow more unequal, the Romer-Roberts-Meltzer and Richard model predicts that political support for redistributive policies will increase. Thus, political support for redistributive policies is predicted to increase when redistribution is most needed. Although each voter is assumed to care only about his or her welfare, the result is a welfare policy that varies appropriately with the needs of the poor.

In this paper, we demonstrate that a more complete theory of voting over welfare spending can lead to the opposite conclusion: The relationship between pre-tax and transfer inequality and support for welfare policies may be perverse in the sense the increased inequality is associated with lower, not higher, support for policies that benefit the bottom half of the income distribution. In the second part of the paper, we show that the perverse relationship between inequality and redistributive policies is what the data show among OECD countries. Together, the theory and the evidence implies that the political consequences of inequality magnify, rather than diminish, differences between countries and change within countries over time in the inequality of wages and salaries, at least among advanced industrial societies.

The framework we develop combines two different approaches to understanding the sources of political support for welfare policy. In the first view, as expressed in the Romer-Roberts-Meltzer and Richard model, as well as in the large literature in political science and sociology that emphasizes the political strength of the working class in cross-national studies of welfare spending, welfare policy is fundamentally about redistribution from rich to poor.¹ Self-interested voters support welfare policy up to the point where their gain from the redistribution of income is matched by their share of the cost that redistributive policies impose. In the second view, the essence of welfare policy is the public provision of insurance. In the insurance framework, self-interested voters support welfare policy to obtain insurance against risks for which private insurance markets fail to cover.²

Whether welfare policy is viewed as redistributing income or providing insurance has significant rhetorical implications. Viewed as redistributive policies, welfare programs benefit the poor at the expense of the rich, and generally reduce the efficiency of the economy. Viewed as social insurance policies, welfare policies provide insurance benefits to all and enhance efficiency to the extent that the public sector provides insurance against risks that are difficult or impossible to cover through private insurance markets. In fact, the policy implications of the two different understandings welfare policies are less clear

than the rhetoric suggests. On the one hand, one might think that welfare policies are fundamentally redistributive yet support increases in spending on the grounds that the benefit of increased equality outweighs the efficiency loss. On the other hand, one might consider welfare policies as social insurance policies yet think that the demand for insurance could be better satisfied by private firms.

In this paper, our purpose is to investigate the contrast between the redistributive and the insurance model of welfare policies in terms of the impact of inequality on the political support for welfare spending. Consider an increase in the inequality of income that lowers the income of the median voter, while leaving average income constant. In the redistributive model, the lower the income of median voter relative to the average income, the more the median voters gains from welfare expenditures. In the insurance model, in contrast, the demand for insurance declines with income, holding risk constant, assuming that insurance is a normal good. If the median voters' income declines, and the risks covered by social insurance do not change, support for spending on social insurance declines.

In our judgment, both approaches provide essential ingredients for an adequate understanding of the politics of welfare policy. Social insurance policies comprise a large part of the welfare budget, at least in advanced industrial societies, and even policies of social assistance can be viewed as providing insurance against the residual risk of income loss that social insurance policies do not cover. At the same time, public insurance is commonly provided and financed in a manner that is redistributive *ex ante* in that voters with lower expected incomes receive insurance on more favorable terms than voters with higher expected incomes. Thus, one cannot tell which aspect of welfare policy, redistribution or insurance, dominates in determining the impact of inequality on support for welfare spending without specifying how the policy is designed.

In the model we develop, government spending is characterized by two parameters that determine the level of government spending and the group to which the benefits of

government spending are targeted. This framework encompasses both the model of welfare policy as fundamentally redistributive and the model of welfare policy as fundamentally about insurance depending on the type of targeting. When the benefits of government spending are predominantly received by persons who are employed, we are able to reproduce the Romer-Roberts-Meltzer and Richard model. When benefits of government spending are targeted to those without earnings, the response to greater inequality is as predicted by the model of welfare policy as insurance. Finally, when the type of targeting is determined endogenously through a democratic process, we show that the equilibrium consists of an outcome in which the insurance aspect of welfare policy dominates if the pre-tax and transfer inequality is not too great. If the initial inequality is sufficiently large, the political equilibrium is one in which both models retain some predictive power.

The next section of the paper presents the basic assumptions regarding the economy and the policy choice. In Section III, we consider the impact of inequality on voters' choice of the level of benefits for a given type of targeting. In this section, we show that the conclusions of both the redistributive model and the insurance model of welfare spending can be generated within a single framework by assuming different values for the parameter that determines the extent to which benefits are targeted to those without earnings. In section IV, we investigate the simultaneous choice of the level and targeting of benefits and demonstrate that, if the initial distribution of income is not too unequal, the comparative statistics of the political equilibrium correspond to the insurance model of government spending. The fifth section discusses the empirical association of welfare spending and wage inequality in OECD countries. Section VI concludes.

II. Basic Assumptions

In order to make our analysis as transparent as possible, we present our theory in the context of the simplest model of the economy that we can think of that contains the essential ingredients of (a) uncertainty with regard to future income on the part of a significant fraction of the population and (b) heterogeneity among voters in terms of

both their income and the risks they face. We will assume that the population is divided into three groups. The share σ_0 of the population are permanently outside the labor market and have no income other than what they receive as transfer payments. The share σ_L consist of wage-earners who receive a wage of w_L when they are employed. The share σ_H are higher income workers, or salary-earners for short, who receive earnings of w_H , with $w_H > w_L$. We assume that the three groups exhaust the population, so that $\sigma_0 + \sigma_L + \sigma_H = 1$.

Wage-earners may be gainfully employed or not. We assume that the probability of employed wage-earners losing their source of income, whether because of being laid off, going bankrupt, or falling sick, within the time period dt is αdt . The probability that workers who are temporarily without earnings will recover their earnings within dt is βdt . For simplicity, both α and β are assumed to be constant.³ The Markov process described by the parameters α and β converges to a steady state distribution of wage-earners in which the fraction θ are working where

$$\theta = \frac{\beta}{\alpha + \beta}$$

Alternatively, θ denotes the fraction of time that each wage-earner expects to be employed in the long-run. We assume that salary-earners face no risk of losing their earnings.

The population not working, therefore, consists of the share who are permanently outside the labor market, σ_0 , plus the share of the population who are temporarily without employment, $(1 - \theta)\sigma_L$, while the work force consists of salary-earners, σ_H , plus the share of wage-earners who are employed, $\theta\sigma_L$. We will assume that a majority of the population is employed, or that $\sigma_0 + (1 - \theta)\sigma_L < 1/2$. In addition, we assume that salary-earners who, by assumption, face no risk of losing their income, constitute a minority of the population, or that $\sigma_H < 1/2$. It follows that the group of employed wage-earners are the median income earners.

We represent fiscal policy with two parameters. The first is the flat tax rate on earnings, t , that determines total government spending per capita, $T(t)$. We write the

requirement that tax receipts equal expenditures as

$$T(t) = \tau(t)[\theta\sigma_L w_L + \sigma_H w_H] = \tau(t)[\theta\sigma_L + \sigma_H]\bar{w} \quad (1)$$

where \bar{w} is the average wage ($\bar{w} \equiv [\theta\sigma_L/(\theta\sigma_L + \sigma_H)]w_L + [\sigma_H/(\theta\sigma_L + \sigma_H)]w_H$) and $\tau(t)$ represents tax revenues as a share of earnings as a function of the tax rate. The function $\tau(t)$ implicitly incorporates the deadweight cost of taxation. We assume, therefore, that $\tau(t)$ is a strictly concave function with $\tau(0) = \tau(1) = 0$ and $\tau'(0) = 1$.⁴

The second policy parameter, γ , represents the share of welfare spending received by persons who are employed. The remaining share, $(1 - \gamma)$, is assumed to go to programs targeted to those without earnings. Thus, the post-tax and transfer consumption of a person with a pre-tax income of w is

$$c_E(w) = (1 - t)w + \frac{\gamma T(t)}{\theta\sigma_L + \sigma_H} \quad (2)$$

where $\gamma T(t)/[\theta\sigma_L + \sigma_H]$ is the welfare benefit received by each employed person. The consumption of those without earnings is

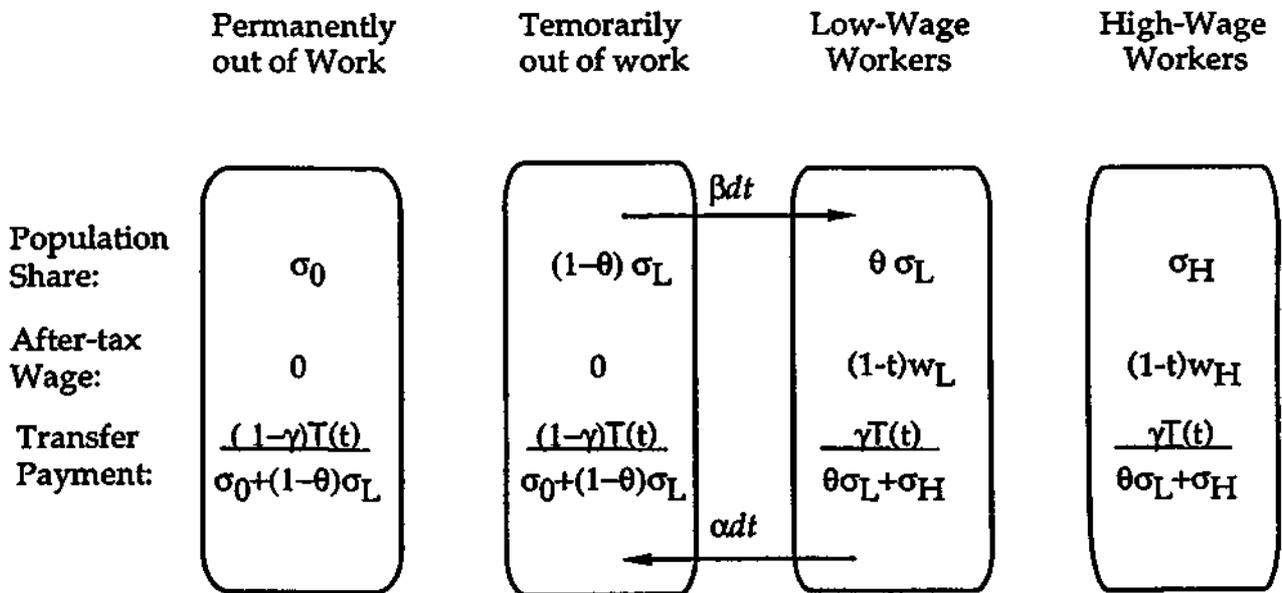
$$c_N = \frac{(1 - \gamma)T(t)}{\sigma_0 + (1 - \theta)\sigma_L} \quad (3)$$

Implicit in (6) is an assumption that all persons without earnings receive the same benefit, regardless of their history of past employment or earnings.⁵ If $\gamma = 0$, we have a welfare policy that is targeted at those without work. A universalistic policy that paid the same benefit to all individuals regardless of employment status is implied by $\gamma = \theta\sigma_L + \sigma_H$. If $\gamma = 1$, the benefits go exclusively to those with earnings. (We assume throughout that $0 \leq \gamma \leq 1$.) Our assumptions regarding the division of the population with respect to the labor market and the distribution of post-tax and transfer income are summarized in Figure 1.

Figure 1 about here

Finally, we assume that all individuals have identical preferences over consumption, described by a standard utility function $u(c)$ with the following characteristics (i) $u''(c) < 0$,

Figure 1
Distribution of Private Income



(ii) $u'(c) \rightarrow \infty$ as $c \rightarrow 0$, and (iii) $\mu \equiv -cu''(c)/u'(c) \geq 1$. Assumption (i) states that individuals are risk averse. Assumption (ii) insures that individuals always want some insurance to cover a non-negligible risk of having nothing. Assumption (iii) implies that the demand for insurance remains constant or increases as income rises. Empirical estimates of μ , usually called the coefficient of relative risk aversion, consistently conclude that $\mu \geq 1$ (Friend and Blume 1975).

Assuming that individuals live forever, the expected lifetime utility of a wage-earner, can be derived from the asset equations:

$$rV^E = u(c_E(w)) - \alpha(V^E - V^N) \quad (4)$$

$$rV^N = u(c_N) + \beta(V^E - V^N) \quad (5)$$

where V^E is the expected lifetime utility of a person currently employed, V^N is the expected lifetime utility of a person temporarily not employed, $u(c_i)$ is the instantaneous utility of consumption when employed ($i = E$), or when not employed ($i = N$) and r is the discount rate. Equations (4) and (5) can be solved for the expected lifetime utilities of starting out in the two different states. We will concentrate on the expected lifetime utility of employed wage-earners, which is conveniently written in the form:

$$v \equiv rV^E = \left(\frac{\beta + r}{\alpha + \beta + r} \right) u(c_E(w_L)) + \left(\frac{\alpha}{\alpha + \beta + r} \right) u(c_N) \quad (6)$$

As equation (6) shows, the expected lifetime utility of an employed wage-earner consists of a weighted average of expected utility in the two states, with the current state of being employed weighted more heavily the greater the discount rate. If employed wage-earners cast their votes to maximize (6), and if r equals the rate at which wage-earners discount future income, we have a model of self-interested voting. If r is less than the interest rate, we have a model that includes some degree of altruism with additional weight on the utility of those without earnings.

III. Choosing the Level of Benefits with Exogenous Targeting

We first investigate the political choice of the level of benefits when the targeting of benefits is fixed. The investigation of the choice of t for a fixed γ provides a general framework in which the contrasting predictions of the two models of welfare spending—welfare as redistribution and welfare as publicly provided insurance—can be compared and shown to depend on the extent to which benefits are targeted to those without work. In addition, in circumstances where changing the level at which existing programs are funded is politically easier than altering the programs' designs, the model of choosing t for a fixed γ may be applicable.

With γ fixed, the level of taxation and benefits preferred by wage-earners is given by the first-order condition

$$(\alpha + \beta + r) \frac{dv}{dt} = (\beta + r) u'(c_E) [\gamma \tau'(t) \bar{w} - w_L] + \alpha u'(c_N) \left[\frac{(1 - \gamma) \tau'(t) (\theta \sigma_L + \sigma_H) \bar{w}}{\sigma_0 + (1 - \theta) \sigma_L} \right] = 0 \quad (7)$$

or

$$\left(\frac{\beta + r}{\alpha} \right) \frac{u'(c_E)}{u'(c_N)} - \left[\frac{\theta \sigma_L + \sigma_H}{\sigma_0 + (1 - \theta) \sigma_L} \right] \left[\frac{(1 - \gamma) \tau'(t)}{(w_L/\bar{w}) - \gamma \tau'(t)} \right] = 0 \quad (8)$$

The first term in (8) is the slope of the wage-earners' indifference curve in the $c_N - c_E$ space while the second term is the slope of the transformation frontier between c_N and c_E . The first-order condition for a wage-earner temporarily without earnings is identical to (7) or (8) except that the coefficients are β in front of $u'(c_E)$ and $(\alpha + r)$ in front of $u'(c_N)$. The first-order condition for the σ_0 voters who are permanently outside the labor force is $\tau'(t) = 0$. Finally, for salary-earners, $du(c)/dt = u'(c_E) [\gamma \tau' \bar{w} - w_H] < 0$ for all t in the absence of altruistic voting. The strict concavity of $u(c)$ and $\tau(t)$ guarantees that the second order condition for a maximum is satisfied. Therefore, all voters have single-peaked preferences and we can identify the political equilibrium with the preferred policy of the median voter. Assuming that $\sigma_0 + (1 - \theta) \sigma_L < 1/2 < \sigma_0 + \sigma_L$, the median voter is an employed wage-earner.

From (8), it is apparent that an decrease in the discount rate r or an increase in voters' altruism induces voters to raise $u'(c_E)/u'(c_N)$ or to increase the redistribution of

income from c_E to c_N . Conversely, an increase in the cost of taxation, as represented by a decrease in the marginal tax yield $\tau'(t)$, induces voters to lower $u'(c_E)/u'(c_N)$ or to reduce the redistribution of income from c_E to c_N . It is sometimes argued that the more policy benefits are targeted to the majority with earnings, the higher the level of political support.⁶ Within the framework of this paper, this argument is partially correct. Differentiating (8) with respect to γ yields

$$\frac{dt^*}{d\gamma} = \left\{ \left[\frac{(1-\gamma)\tau'(t)}{(w_L/\bar{w}) - \gamma\tau'(t)} \right] + \left[\frac{(1-\gamma)\tau\bar{w}\mu}{(1-t)w_L + \gamma\tau\bar{w}} \right] + (\mu - 1) \right\} \frac{1}{Q} > 0$$

where

$$Q \equiv \left\{ \frac{[(1-t)\tau'(t) + \tau]\mu}{[(1-t)w_L + \gamma\tau\bar{w}]\tau} - \frac{\tau''(t)}{\tau'(t)[w_L - \gamma\tau'(t)\bar{w}]} \right\} (1-\gamma)w > 0$$

since $\mu \geq 1$, $\tau''(t) < 0$ and $w_L - \gamma\tau'(t)\bar{w} > 0$ from (8). Thus, the more benefits are targeted to the employed, the higher the tax rate the employed majority will support.

However, the argument is not really about whether or not less targeting to the poor leads to higher taxes, but whether less targeting to the poor increases welfare spending to such an extent that the poor benefit. The question of whether or not less targeting to the poor benefits the poor concerns the sign of

$$\frac{dc_N}{d\gamma} = \left[\frac{\theta\sigma_L + \sigma_H}{\sigma_0 + (1-\theta)\sigma_L} \right] \bar{w} \left[(1-\gamma)\tau'(t)\frac{dt^*}{d\gamma} - \tau \right]$$

which can be either positive or negative, depending on the concavity of the function $\tau(t)$. It is straightforward to show that $dc_N/d\gamma > 0$ when $\tau(t) = t$ while $dc_N/d\gamma < 0$ when $|\tau''(t)|$ is sufficiently large.

Our primary concern, however, is with the effect of differences in inequality on the support for welfare. Consider the impact of a mean-preserving spread in the wage distribution, that is an increase in w_H and a reduction in w_L such that \bar{w} remains constant. To investigate the impact of w_L on wage-earners' support for welfare, define

$$\Psi(w_L, \gamma) \equiv u'(c_E)[\gamma\tau'(t)\bar{w} - w_L] \tag{9}$$

as the part of equation (7) that depends on w_L . Let t^* be the median voters' preferred tax rate. From the second-order condition, we have $\partial^2 v / \partial t^2 < 0$. Therefore, the sign of dt^*/dw_L is the same as the sign of

$$\frac{\partial \Psi(w_L, \gamma)}{\partial w_L} = u'(c_E)(\mu\xi - 1) \quad (10)$$

where

$$\xi \equiv \frac{dc_E}{d(1-t)} \frac{(1-t)}{c_E} = \frac{(1-t)[w_L - \gamma\tau'(t)\bar{w}]}{(1-t)w_L + \gamma\tau\bar{w}}$$

is the elasticity of consumption when working with respect to $1-t$. Since $\mu \geq 1$ and $\xi \leq 1$, the sign of $\partial \Psi / \partial w$ in (10) is not clear.

There are two special cases in which the sign of $\partial \Psi / \partial w_L$ is immediate, however. The first is the case of $\gamma = 1$ where benefits are exclusively targeted at those with employment. If $\gamma = 1$, we have $w_L - \gamma\tau'(t)\bar{w} = 0$ from (7) which implies that $\xi = 0$. In this case, equation (10) reduces to

$$\frac{\partial \Psi(w_L, 1)}{\partial w_L} = -u'(c_E) < 0 \quad (11)$$

Thus, $dt^*/dw_L < 0$ when benefits are targeted at the employed. Workers with lower wages prefer higher benefits.

The second special case is the case of $\gamma = 0$ where benefits are exclusively targeted at those without employment. Since $\gamma = 0$ implies that $\xi = 1$, equation (10) reduces to

$$\frac{\partial \Psi(w_L, 0)}{\partial w_L} = u'(c_E)(\mu - 1) \geq 0 \quad (12)$$

with strict inequality if $\mu > 1$. Thus, $dt^*/dw_L > 0$ when benefits are targeted at those without employment. In this case, workers with higher wages prefer higher benefits, provided that higher wage workers face the same probability of losing their source of earnings.

The impact of increased inequality on the political support for welfare spending is summarized in the following proposition:

Proposition 1. *A mean-preserving spread in the income distribution (i) reduces the median voters' preferred level of benefits when benefits are targeted on those without*

employment (unless $\mu = 1$, in which case an change in inequality has no effect on benefits), but (ii) increases the median voter's preferred level of benefits when benefits are targeted on the employed.

Proof: The proposition follows immediately from equations (11) and (12) for any mean-preserving change in the wage distribution that reduces the wage received by the median voter. According to (11), a decline in the median voter's wage will reduce her demand for welfare benefits if $\gamma = 0$. According to (12), a fall in the median voter's wage will increase her demand for welfare benefits if $\gamma = 1$. Since $\Psi(w_L, \gamma)$ is continuous in γ , the conclusions hold for γ near zero and γ near one as well.⁷

A mean-preserving spread of the pre-tax income distribution has two effects on the choice of benefits. On the one hand, an increase in inequality represents a decline in income for voters with income below the mean. The effect is to induce voters to reduce c_N in order to reduce the loss of welfare when employed. On the other hand, greater inequality lowers the ratio of the median voter's income to mean income, thereby lowering the tax that must be paid by low-wage workers to finance a given level of benefits. A reduction in the price of providing benefits increases the willingness of low-wage voters to support higher benefit levels. Thus, in addition to an income effect that leads the median voter to reduce his or her preferred level of expenditures, there is a substitution effect that works in the opposite direction.

Alternatively, the two effects can be described as a redistribution effect and an insurance effect. For any value of $\gamma < 1$, welfare policy both redistributes income and provides insurance. A rise in inequality increases the redistributive impact of the welfare system, to the benefit of low-wage workers. At the same time, an increase in inequality, holding average income constant, implies a reduction in the income of workers whose wages are below average. Voters whose wages have declined prefer to reduce the amount of insurance they buy. Which effect dominates depends on the coefficient of relative risk aversion, μ , and on the targeting of welfare benefits as determined by the parameter γ .

Consider first the effect of voters' risk aversion, for a fixed value of $0 \leq \gamma \leq 1$. If $\mu = 1$, i.e. if voters' risk aversion is as low as we allow, the insurance benefit provided by the welfare system is less important and an increase in inequality increases the median voters' preferred tax rate (unless $\gamma = 0$, in which case the median voters' preferred tax rate is unchanged). If μ is sufficiently large, i.e. if voters are sufficiently risk averse, the insurance aspect of welfare dominates, in which case an increase in inequality lowers the median voters' preferred tax rate.

Alternatively, for a fixed $\mu \geq 1$, whether the redistributive or the insurance effect dominates can be understood as a function of the targeting of the benefits. When benefits are mostly paid to those who are employed (when γ is close to one), the redistributive aspect dominates and the median income earners' preferred tax rate increases as inequality grows. This is the case described by the standard Romer-Richards-Meltzer and Richard model. When benefits are mostly paid to those without earnings (when γ is close to zero), however, the insurance aspect dominates and the median income earners' preferred level of spending declines as inequality increase.

The model in this section can be generalized in a variety of ways without altering the conclusions. The simplifying assumption that there are only two types of workers, lower paid wage-earners and higher paid salary-earners, can be replaced by the assumption of any finite number of types, or even of a continuum of types. The assumption that only wage-earners are subject to the risk of income loss can be replaced by assuming a general distribution of risk of income loss. As long as the risk of job loss does not increase as one moves up the wage scale, all of the results of this section go through with little change. Finally, one might consider a different environment in which workers' wages each period are random draws from a known wage distribution. In this case, one can investigate the impact of an increase in workers' uncertainty regarding their future wages, holding their expected wage and their risk of job loss (drawing a zero wage) constant. The result is very similar to the case explored above. If γ is close to one, an increase in workers' uncertainty

regarding future wages raises the preferred tax rate of the median voter. If γ is close to zero, however, an increase in workers' uncertainty regarding future wages reduces the median voters' preferred level of taxes and benefits.⁸

The assumption that the transition probability into employment, βdt , is not affected by changes in taxes and benefits could be relaxed to allow β to be a function of the difference in welfare between those with earnings and those without earnings. The result would be to introduce an additional mechanism whereby increased inequality causes reduced welfare spending, the mechanism described by Moffitt, Ribar and Wilhelm (1998) in which voters cut benefits to the poor as wages fall to restore work incentives. Another extension would be to allow for self-insurance (or savings) to compete with public insurance against income loss. Introducing the possibility of saving would enable the number of wealthy, defined as that part of the population with no need for social insurance, to be made endogenous. Such an extension might introduce a third reason for the association of greater inequality with lower welfare spending insofar as greater inequality increases the relative share of the electorate who prefer self-insurance to public insurance.

IV. Choosing both Benefit Levels and Targeting

The targeting of benefits is as much a political decision as the level of welfare spending. Thus, a general model of the politics of welfare must include the political choice of targeting. Consideration of a second dimension of political choice is made difficult, however, by the general absence of a majority rule equilibrium in two dimensions without restrictive assumptions about the political process. Our strategy is to start with the simple case where a majority of voters have identical preferences and hence can dictate the outcome. We then relax the assumption that there is a majority of a single type and present a particular model of political competition that generates the same equilibrium as the case with a majority of like-minded voters.

We assume, to begin with, that $\theta\sigma_L > 1/2$ or that the group of employed wage-earners who receive w_L and face the risk αdt of losing their source of earnings within the period

dt constitute a majority of the population. It is convenient to change variables and use t and c_N instead of t and γ as the choice variables. Using (2) and (3) to eliminate γ , we can write $c_E(w_L)$ as

$$c_E(w_L) = (1-t)w_L + \frac{T(t) - [\sigma_0 + (1-\theta)\sigma_L]c_N}{\theta\sigma_L + \sigma_H} \quad (14)$$

As before, $T(t) = \tau(t)[\theta\sigma_L + \sigma_H]\bar{w}$. The welfare benefits received by the employed are equal to tax revenues minus the benefits received by those without earnings, or $T(t) - [\sigma_0 + (1-\theta)\sigma_L]c_N$, divided by the fraction of the electorate with earnings $[\theta\sigma_L + \sigma_H]$. The majority's ideal point is given by the combination of t and c_N that solves the problem

$$\max(\alpha + \beta + r)v = (\beta + r)u(c_E) + \alpha u(c_N) \quad (15)$$

$$\text{such that } T(t) - [\sigma_0 + (1-\theta)\sigma_L]c_N \geq 0$$

The constraint $T(t) - [\sigma_0 + (1-\theta)\sigma_L]c_N \geq 0$, or $\gamma \geq 0$, rules out lump-sum taxation. The constraint that $\gamma \leq 1$, or $c_N \geq 0$, is never binding given our assumption that $u'(c) \rightarrow \infty$ as $c \rightarrow 0$. The first-order conditions for the solution to (15) can be written as

$$\begin{aligned} (\beta + r)u'(c_E)\frac{\partial c_E}{\partial t} + \lambda T'(t) &= 0 \\ (\beta + r)u'(c_E)\frac{\partial c_E}{\partial c_N} + \alpha u'(c_N) - \lambda[\sigma_0 + (1-\theta)\sigma_L] &= 0 \\ \lambda\{T(t) - [\sigma_0 + (1-\theta)\sigma_L]c_N\} &= 0 \end{aligned}$$

where λ is the Lagrangian multiplier.

There are two cases to consider. Consider first the case in which the constraint is not binding, or in which $T'(t) - [\sigma_0 + (1-\theta)\sigma_L]c_N > 0$ ($\gamma > 0$). In this case $(\partial c_E/\partial t) = [\tau'(t)\bar{w} - w_L]$, $(\partial c_E/\partial c_N) = -[\sigma_0 + (1-\theta)\sigma_L]/[\theta\sigma_L + \sigma_H]$ and $\lambda = 0$. The first-order conditions can be written as

$$\tau'(t) = w_L/\bar{w} \quad (16)$$

$$\frac{u'(c_E)}{u'(c_N)} = \left(\frac{\alpha}{\beta + r}\right) \left[\frac{\theta\sigma_L + \sigma_H}{\sigma_0 + (1-\theta)\sigma_L}\right] \quad (17)$$

Equation (16) determines the optimal tax rate as a function of the ratio of the median to the mean wage in line with the standard model. The optimal tax is zero when $w_L = \bar{w}$ (since

$\tau'(0) = 1$, and rises as w_L/\bar{w} declines (since $\tau''(t) < 0$). According to (17), the majority's preferred level of consumption for those without earnings, c_N , is a rising function of the wage, since

$$\frac{dc_N}{dw_L} = \frac{(1-t)c_N}{(1-t)w_L + \tau\bar{w}} > 0$$

The second case to consider is when the constraint $T(t) - [\sigma_0 + (1-\theta)\sigma_L]c_N \geq 0$ is binding. In this case, $(\partial c_E/\partial t) = -w_L$ and $(\partial c_E/\partial c_N) = 0$. The first-order conditions can be written as

$$\begin{aligned} T(t) &= [\sigma_0 + (1-\theta)\sigma_L]c_N \\ \frac{u'(c_E)}{u'(c_N)} &= \left(\frac{\alpha}{\beta+r} \right) \left[\frac{\theta\sigma_L + \sigma_H}{\sigma_0 + (1-\theta)\sigma_L} \right] \frac{\tau'(t)}{(w_L/\bar{w})} \end{aligned} \quad (18)$$

Since $\tau'(t) < w_L/\bar{w}$ when the constraint is binding, the RHS of (18) is less than the RHS of (17). The majority would like to lower t and raise money with a lump-sum tax (i.e., set γ below zero), but lump-sum taxes are ruled out by the constraint. Therefore, the majority transfers less money from c_E to c_N than they would if lump-sum taxes were allowed. Since (18) is identical to (8) with $\gamma = 0$, this case was already investigated in the previous section. From Proposition 1, we know that $dt^*/dw_L > 0$. With $\gamma = 0$, we have

$$\frac{dc_N}{dw_L} = \left[\frac{T'(t)}{\sigma_0 + (1-\theta)\sigma_L} \right] \frac{dt^*}{dw_L} > 0$$

as well.

Combining the two cases, the majority's preferred policy is illustrated in Figure 2. The horizontal axis represents the wage of the majority, or w_L . The curve labelled $T(t^*)$ represents the majority's unconstrained optimal level of tax revenues while the curve $[\sigma_0 + (1-\theta)\sigma_L]c_N^*$ represents the unconstrained optimum with respect to the benefits targeted to those without earnings. Since $T(t^*) = 0$ when $w_L = \bar{w}$ while c_N^* is always positive and rising with w_L , the two curves must cross at a wage level below \bar{w} , denoted w_0 in the figure. If $w_L < w_0$, the majority's choice of benefits targeted to themselves is given by the difference between $T(t^*)$ and $[\sigma_0 + (1-\theta)\sigma_L]c_N^*$. For $w_L \geq w_0$, the constraint that

$T(t) \geq [\sigma_0 + (1 - \theta)\sigma_L]c_N^*$ binds. The constrained optimum with $T(t) = [\sigma_0 + (1 - \theta)\sigma_L]c_N^*$ is represented by the curve labelled $T(t^{**})$.

Figure 2 about here

The comparative static results implicit in Figure 2 are summarized as follows:

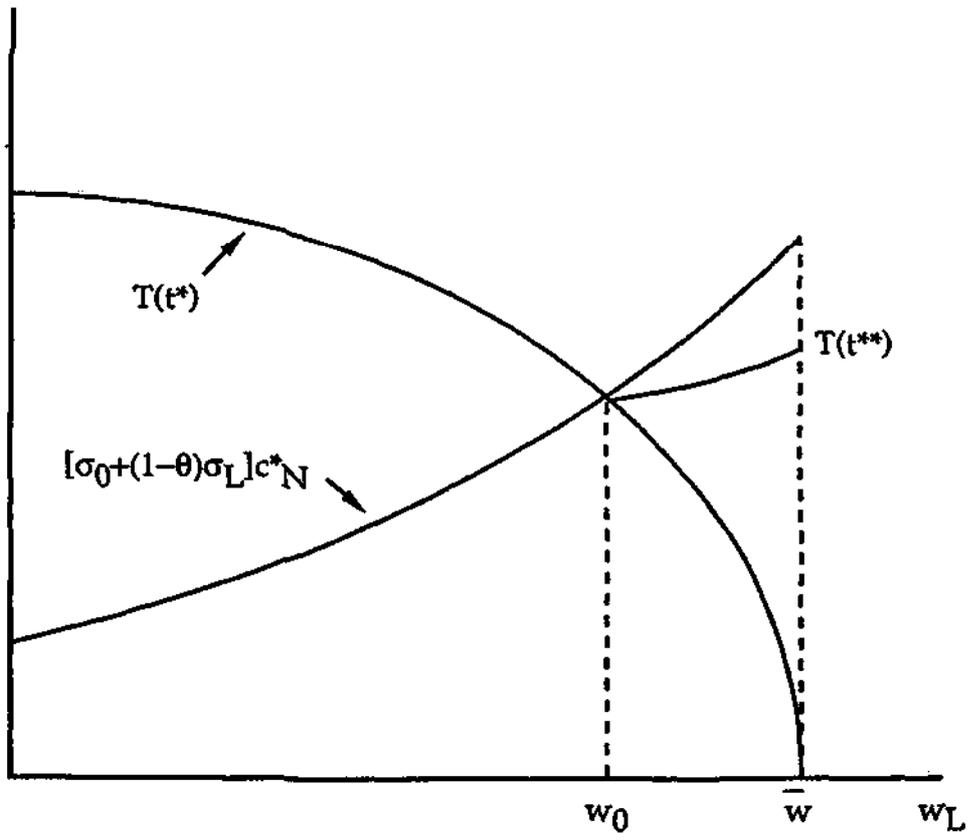
Proposition 2. *When a majority of voters receive the wage w_L with $w_L < \bar{w}$, a mean-preserving spread in the income distribution (i) reduces the majority's preferred level of redistribution to those with no income, (ii) reduces the majority's preferred level of redistribution from high-wage workers when initial inequality is sufficiently small, but (iii) increases the majority's preferred level of redistribution from high-wage workers when initial inequality is sufficiently large.*

Proof: Since \bar{w} represents a weighted average of w_L and w_H , a mean-preserving spread in the income distribution implies a decline in w_L . Thus, part (i) of the proposition states that the optimal c_N increases with w_L , or that the curve $[\sigma_0 + (1 - \theta)\sigma_L]c_N^*$ for $w_L \leq w_0$ and $T(t^{**}) = [\sigma_0 + (1 - \theta)\sigma_L]c_N^*$ for $w_L \geq w_0$ slopes upward. Parts (ii) and (iii) state that the optimal t is a V-shaped function of w_L , or that $T(t^*)$ slopes downward for $w_L \leq w_0$ while $T(t^{**})$ slopes upward for $w_L \geq w_0$.

When employed wage-earners choose both the type of targeting and the level of benefits, a mean-preserving increase in inequality always reduces redistribution towards those without income but has a non-monotonic effect on redistribution away from high-wage workers. In sufficiently unequal societies, further increases in inequality increase redistribution from the rich to the middle, while reducing redistribution from the middle to the bottom. In less unequal societies, increases in inequality reduce redistribution, both from the rich and towards the poor.

So far, we have assumed away the problems of modeling political competition in two-dimensional policy space by assuming the existence of a majority with identical policy preferences. To introduce political competition in a non-trivial way, it is sufficient to

Figure 2
The Preferred Policy of the Majority



relax the assumption that employed wage-earners constitute a majority of the electorate. Suppose, instead, that σ_0 , the share of the electorate that consists of persons outside the labor market, $\theta\sigma_L$, the share of the electorate that consists of employed wage-earners, and σ_H , the share of the electorate that consists of salary-earners who face no risk of job loss, are each between one quarter and one half so that any two of these three groups constitute a majority.

In the case in which w_L is sufficiently close to \bar{w} ($w_0 \leq w_L \leq \bar{w}$), there is a majority in favor of eliminating government spending for benefits received by employed persons. Both wage-earners and those outside the labor market prefer to have all welfare spending targeted on those without earnings such that $T(t) = [\sigma_0 + (1 - \theta)\sigma_L]c_N^*$ or $\gamma = 0$. With this constraint, the policy space is one-dimensional and the median voter theorem applies. The ideal policy of the employed wage-earners is preferred by a majority to any other feasible policy alternative. In the case in which w_L is sufficiently low such that employed wage-earners prefer a policy combination with some government spending targeted to employed workers ($w_L < w_0$), however, the policy space is two-dimensional and no policy has the property of being majority-preferred to all feasible alternatives.

The preferences of the different groups of voters in the $c_N - t$ policy space with $w_L < w_0$ are illustrated in Figure 3. The curve $T(t)/[\sigma_0 + (1 - \theta)\sigma_L]$ represents the maximum attainable c_N corresponding to each value of t . The set of feasible policies is given by $c_N \leq T(t)/[\sigma_0 + (1 - \theta)\sigma_L]$. Those outside the labor force always prefer higher to lower values of c_N . Thus, the indifference curves of the group outside the labor market are horizontal lines, such as the line $A - A'$. Salary-earners have above average earnings and, by assumption, face no risk of income loss. Thus, salary-earners prefer higher to lower values of $c_E(w_H)$. Their ideal point is at the origin with $c_N = t = 0$. The line $B - B'$ represents a typical indifference curve of salary-earners along which c_E is constant. I is the ideal point of employed wage-earners for the case in which $w_L < w_0$. The tax rate t^* is the tax rate that satisfies the first-order condition $\tau'(t) = w_L/\bar{w}$. The indifference curves

of employed wage-earners are ellipses around I , as shown in the figure. The ideal point of unemployed wage-earners, not shown in Figure 3, is above I on the vertical line segment IC at $t = t^*$ or on the $T(t)/[\sigma_0 + (1 - \theta)\sigma_L]$ curve above C .

Figure 3 about here

Note that if $w_L > w_0$, then I would be on the feasible policy frontier, $c_N = T(t)/[\sigma_0 + (1 - \theta)\sigma_L]$, and any deviation from I would be opposed by a majority of voters. If $w_L < w_0$, however, both salary-earners and voters who are outside the labor force would prefer a policy between A and B on the $c_N = T(t)/[\sigma_0 + (1 - \theta)\sigma_L]$ curve or in the interior of the triangle ABI over I . In this case, the ideal point of employed wage-earners can be overturned by an alliance of the rich and the poor in which lower taxes and higher transfers to those with no earnings are simultaneously obtained by reducing the spending on benefits targeted to persons who are employed. Of course, any policy between A and B on the $c_N = T(t)/[\sigma_0 + (1 - \theta)\sigma_L]$ curve could itself be overturned by an alliance of the middle with either the rich or the poor, and so on.

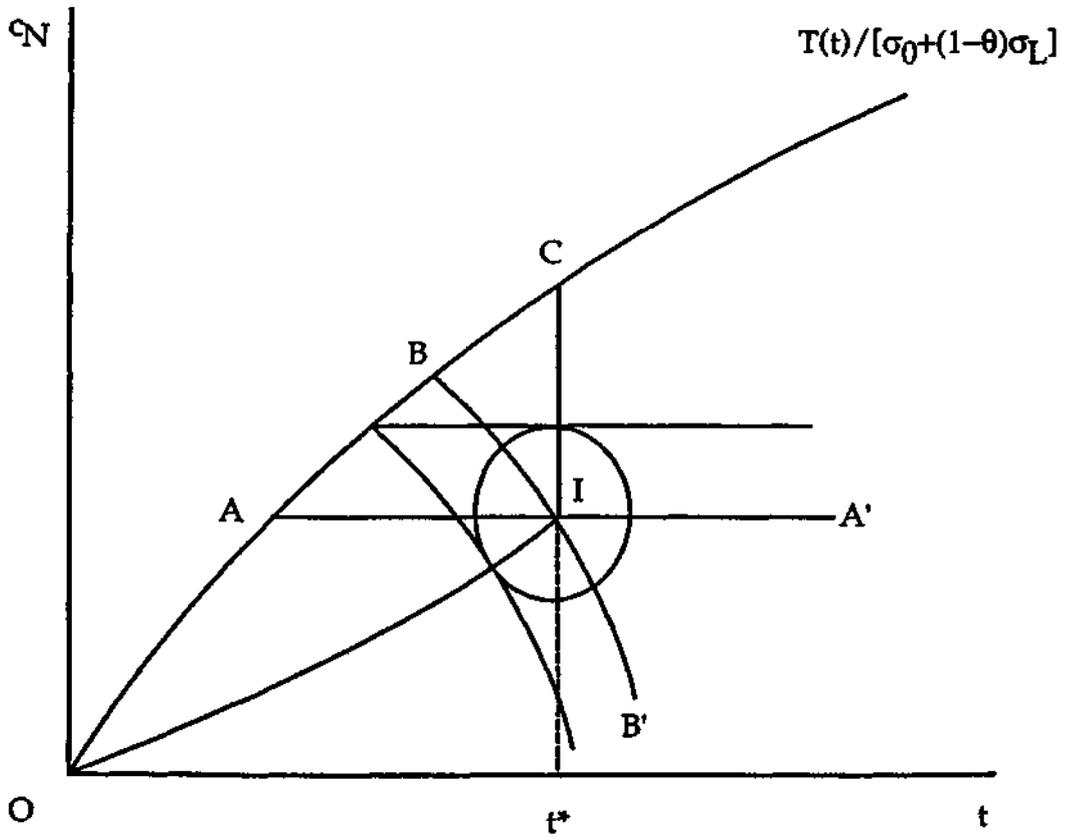
To obtain a political equilibrium when $w_L < w_0$, we need to restrict the nature of political competition in some way. Here we adopt the approach developed by Roemer (1998, 1999) in which political competition is assumed to take place between two main parties, or two families of parties such as the Social Democratic block and the Bourgeois block in Scandinavia, that have distinctive constituencies.⁹ In particular, we assume that the preferences of the left party are given by a convex combination of the welfare of employed wage-earners and those outside the labor force, or

$$v^L = \omega_L u(c_N) + (1 - \omega_L) \left[\left(\frac{\beta + r}{\alpha + \beta + r} \right) u[(c_E(w_L))] + \left(\frac{\alpha}{\alpha + \beta + r} \right) u(c_N) \right] \quad (19)$$

where $0 \leq \omega_L \leq 1$ is the weight given to those with no earnings by the left party. Similarly, we assume that the preferences of the right party are given by a convex combination of the welfare of employed wage-earners and salary-earners, or

$$v^R = \omega_R u[c_E(w_H)] + (1 - \omega_R) \left[\left(\frac{\beta + r}{\alpha + \beta + r} \right) u[(c_E(w_L))] + \left(\frac{\alpha}{\alpha + \beta + r} \right) u(c_N) \right] \quad (20)$$

Figure 3
Voters' Preferences over cN and the Tax Rate



where $0 \leq \omega_R \leq 1$ is the weight given to high-wage workers by the Right party. The parties' preferences could be the result of a compromise between party militants, committed to each party's core constituency, and opportunists who are committed to the median group because they want to win, as in Roemer (1998, 1999). Alternately, the parties' preferences could be the result of a compromise between representatives of the different income groups who belong to the same party. Voters vote for the party that offers the most attractive platform.

In Figure 3, the line segment IC represents the contract curve for the Left party, or the set of efficient compromises between employed wage earners and those who are outside the labor market. Similarly, the line segment OI represents the contract curve of the Right party. Equations (19) and (20) imply that the parties' ideal points are located on their respective contract curves. Moreover, the electoral platforms of both parties must lie on their respective contract curves as well. Since the contract curves can be defined as the set of points that maximize the parties' support among the median group of voters subject to the constraint of promising a given level of welfare for the parties' core constituents, any party platform that is off the contract curve is dominated by another that is on the contract curve. If voters in the median group vote for the party that promises the best policy for the middle group, the unique equilibrium is for both parties to promise to implement I , the median income group's ideal point. Any other choice by the left party would allow the right party to win the election with a platform that is worse than I for the left's constituency, and vice versa for the right party. The critical assumption in the model of two party competition is that the party system blocks the formation of an alliance of the rich and the poor against the middle. Political competition is essentially reduced to a single dimension and the logic of the median voter theorem can be applied.

Complete convergence on the median groups' ideal point implies that the outcome of elections is irrelevant, since both parties implement the same policy. A minor extension of the model, however, allows for only partial convergence in equilibrium. Suppose that

the vote of the middle group of voters is influenced by perceptions of the relative competence of the two parties (or the charisma of the party leaders), as well as their platforms. If the parties are uncertain about the distribution of voters' perceptions of the relative competence of the two parties, the relationship between party platforms and vote shares is stochastic from the parties' point of view.

Let the probability that a voter in the median group will vote for the left party is written as

$$\Pr[x < V^E(t^L, c_N^L) - V^E(t^R, c_N^R)] = F[V^E(t^L, c_N^L) - V^E(t^R, c_N^R)]$$

where x represents the perceived difference in competence of the right versus the left party among the middle group of voters and $V^E(t^i, c_N^i)$, $i = L, R$, represents the expected welfare of a member of the middle group, as defined in equation (4), should the platform (t^i, c_N^i) be implemented. If $F[\cdot]$ is a continuous cumulative density function, both parties face a smooth tradeoff between improving their probability of winning by adopted policies close to those preferred by the median group, and lowering the value of winning in terms of the welfare of the parties' core constituents. If both parties choose their platforms to maximize the welfare of their constituents, as defined in equations (19) and (20), each party, in equilibrium, will adopt a platform located on its contract curve in the interval between its own ideal point and the ideal point of the median voters.¹⁰ The equilibrium, then, consists of a lottery between the two parties.

With this extension, election results can alter the policy that is implemented. Nevertheless, the pressure of electoral competition continues to propel the party platforms of the left and right in the direction the ideal point of the median voter. The comparative static results of Proposition 2 remain valid in the sense that the median voters' ideal point continues to define the political center, with left parties implementing policies somewhat to the left of center and right parties implementing policies somewhat to the right of center. While there is no end to the number of different models of politics that could be

investigated, it is clear that the results summarized in Proposition 2 do not depend on the assumption that there is a majority of voters with identical preferences.¹¹

V. Inequality and Welfare Spending Among OECD Countries, 1980-1995

The two central results of the model with endogenous targeting are that (1) total spending is a non-monotonic function of inequality, first falling and then rising as inequality increases and (2) welfare spending targeted to those without employment is a monotonic, declining function of inequality. It is difficult to reject the prediction of a non-monotonic relationship between inequality and total spending. Without some way of estimating w_0 in Figure 2, the model is compatible with finding either a positive or negative relationship between inequality and total spending in the data. Moreover, it is the second claim that most sharply distinguishes our approach from the Romer-Roberts-Meltzer and Richards model in which more inequality is predicted to lead to more redistribution towards the poor. In this section, therefore, we conduct a preliminary empirical test of the sign of the relationship between the inequality of wages and salaries and the level of welfare spending targeted to those without earnings.

Most social insurance and social assistance policies provide benefits to persons who are not currently employed. The most important exception is health care. In the US, government spending on health care is targeted to the elderly (Medicare) or the poor (Medicaid). In all other countries in our data set, however, government financed health care is a universalistic benefit provided to all regardless of income or employment status. To including government spending on health care in our dependent variable would significantly overstate expenditures targeted to those without employment in advanced industrial societies other than the US. To exclude government spending on health, however, would significantly understate expenditures targeted to those without employment in the US.

Let SI_{it} (for Social Insurance) denote government spending on old-age cash benefits, disability cash benefits, occupational accidents and disease benefits, sickness benefits, services for elderly and disabled, survivors' benefits, family cash benefits, family services,

active labor market programs, unemployment benefits, housing benefits and cash benefits for other contingencies, all as a share of GDP in country i in year t . The variable SI_{it} includes both government expenditures and mandated private expenditures. Let H_{it} denote government spending on health as a share of GDP in country i and year t . To include the fraction of government spending on health that goes to those without earnings, we constructed a variable

$$\lambda = \begin{cases} 1 & \text{if the country is the US} \\ .42 & \text{otherwise} \end{cases}$$

where .42 was the share of total health care spending in the US that was spent by the government in 1990 (OECD 1994). We use $SI_{it} + \lambda H_{it}$ as our primary dependent variable.

Our measure of the inequality of wages and salaries is derived from the ratio of the earnings of a worker at the 90th decile to the earnings of a worker at the 10th decile. This data is available for 18 OECD countries from 1979/80 through 1995/96 (OECD 1993, 1996). The data is skewed to the right, with the US and Canada having much higher levels of wage inequality than the rest. Taking the natural log of the wage differential received by the worker at the 90th percentile relative to a worker at the 10th percentile reduces the skewness of the data and improves the fit slightly. Therefore, we use

$$\text{Inequality} = \ln \left(\frac{w_{90} - w_{10}}{w_{10}} \right) \quad (21)$$

as our measure of inequality. The results we report do not depend in any important way, however, on whether one uses the log of the wage differential, as defined in equation (21) or the 90/10 ratio, w_{90}/w_{10} , as the measure of inequality. Because wage inequality data is not available on an annual basis for many of the countries in the data set, we took the average of all data points for each country in the five year periods 1980 – 84, 1985 – 89 and 1990 – 94, yielding a data set with $3(18) = 54$ possible data points. After removing cases in which there is no inequality data within one or more of the specified five year time periods for a particular country, we are left with 49 observations in all.¹²

Welfare spending today is highly correlated with welfare spending in the recent past. If one regresses $SI_{it} + \lambda H_{it}$ in the years 1985, 1990 and 1995 for the 18 countries on

$SI_{it-5} + \lambda H_{it-5}$, one obtains

$$SI_{it} + \lambda H_{it} = \underset{(2.48)}{2.51} + \underset{(17.7)}{0.95} [SI_{it-5} + \lambda H_{it-5}], \quad R^2 = 86.9$$

The t -statistics are listed in the parentheses below the coefficients. Although it is not intellectually exciting to explain current spending on the basis of past spending, it is clear that the impact of past spending on current spending cannot be ignored in empirical work.

In addition to a lagged dependent variable, we control for the share of population over 65, the rate of unemployment, government by conservative parties, per capita GDP, and the turnout for elections to the lower house of parliament or Congress in the case of the US. The share of the elderly in the population and the rate of unemployment are potentially important determinants of both the need for social insurance expenditures and the share of the electorate that rely on transfer payments as their main source of income.¹³ Whether one views the fraction of the population who are over 65 or who are unemployed as measuring need or measuring political influence has a subtle implication for how the data is measured. As an indicator of need, the relevant control is the share of each group in the year at which expenditures are measured. As an indicator of political influence, the relevant control would be an average of the size of each group in the preceding five-year period, since policy changes in response to shifts in the size of different groups in the electorate with a lag. We let the data decide this issue. The unemployment rate fits the data much better if measured in the same year that we measure welfare spending. With regard to the share of elderly in the population, using the average over the preceding five years fits slightly better, although the difference between using the current value versus the average over the preceding five years is small.

In the literature on the impact of the partisan composition of government on welfare spending, the early emphasis was on the division between the left-socialist or social democratic parties—and the center/right (Stephens 1979, Korpi 1983). Like many before us, we find that the most important partisan division with regard to the level of welfare spending is not between the left and the center/right, but between the left/center and the right

(Castles 1982, Esping-Andersen 1990). Since the left versus center/right division was never significant in any of our regressions, we only report our results based on the classification of parties into conservative versus center/left. Our measure of conservative government is the share of cabinet seats held by conservative parties following the classification based on expert judgements collected by Castles and Maier (1984) and, more recently, Huber and Inglehart (1995). We do not include any controls for union strength such as density or centralized wage-setting, since cross-national differences in union density and the centralization of wage-setting are the dominant causes of cross-national differences in the equality of wages and salaries (Freeman 1988, Wallerstein 1999).

Among the earliest findings of the empirical literature on welfare spending as share of GDP is that spending is higher in countries with higher levels of GDP per capita (Wilensky 1975). Although it is common to find that GDP per capita has limited explanatory power in data sets that consist only of countries with relatively high levels of per capita GDP, we include GDP per capita as an independent variable. Finally, it is a common belief that the low level of turnout among lower income voters in the US is an important part of the explanation for the lower level of welfare spending in the US (Lijphart 1997, Franzese 1998). Therefore, we include the average turnout in elections for the lower house of parliament (or Congress in the US) in the preceding five years as a control. Summary statistics and data sources for all of the variables used in the analysis are presented in Table 1.

Table 1 about Here

Regressions results with $SI_{it} + \lambda H_{it}$ as the dependent variable are presented in the first two columns of Table 2. The first column of Table 2 presents the results with the full set of controls. The estimated coefficient on turnout is statistically significant with the “wrong” sign, indicating that countries with higher turnout spend less on social insurance than countries with lower turnout.¹⁴ Since there is no reason to believe that higher turnout reduces redistributive spending, turnout was removed from the data set with the result

**Table 1:
Summary Statistics**

Variable	Mean	Standard Dev.	Minimum	Maximum
$SI + \lambda H$	19.62	5.68	8.64	29.96
Inequality	.609	.381	-.020	1.500
Elderly Share	13.6	2.1	9.5	17.7
Unemployment	7.2	3.1	1.7	17.2
Right Govt.	41.7	37.1	0	100
GDP per capita	12.97	2.25	5.72	18.10
Turnout	78.3	13.3	40.0	95.6

Notes: $SI + \lambda H$ is social insurance plus an estimate of the share of health expenditures paid to those without earnings as described in the text, all as a share of GDP. Data is for 1985, 1990 and 1995 in the case of $SI + \lambda H$ and unemployment. All of the other variables represent the average value for the periods 1980 – 84, 1985 – 89 and 1990 – 94. See footnote (13) for the list of countries included. The source for $SI + \lambda H$ is OECD (1999). The share for the share of elderly in the population, and unemployment is OECD (1997). The inequality data is from OECD (1996) and, in the case of the US, OECD (1993). GDP per capita (in thousands of 1985 US dollars calculated using the chain index) is from the Summers and Heston dataset, described in Summers and Heston (1991). Conservative government is from the Swank data set (Swank 1992), updated using recent issues of Keesings Contemporary Archive. The classification of parties in terms of right versus center and left is based on Castles and Mair (1984) updated with Huber and Inglehart (1995). Turnout refers to turnout in elections in the lower house of parliament, or for president in the United States. The source for turnout is Blais and Dobryzyska (1998). The data set is available upon request.

Table 2:
The Impact of Wage Inequality on Social Insurance Spending as a Share of GDP
in 18 OECD Countries, 1980-95

(Upper parentheses contain robust standard errors.)
(Lower parentheses contain associated t-statistics.)

	1	2	3	4	5
Dependent Variable:	$SI + \lambda H$	$SI + \lambda H$	SI	SI	$SI + \lambda H$ (exc. Finland 1995)
<i>Lagged Dep. Var.</i>	.745 (.074) (10.1)	.719 (.076) (9.41)	.754 (.075) (10.1)	.726 (.075) (9.63)	.724 (.080) (9.07)
<i>Inequality</i>	-2.71 (.711) (3.81)	-1.29 (.611) (2.11)	-3.01 (.714) (4.21)	-1.87 (.526) (3.56)	-0.96 (.684) (1.41)
<i>Elderly Share</i>	.351 (.247) (1.42)	.473 (.264) (1.79)	.337 (.242) (1.39)	.444 (.255) (1.74)	.497 (.301) (1.65)
<i>Unemployment</i>	.253 (.101) (2.50)	.240 (.099) (2.43)	.257 (.107) (2.41)	.249 (.103) (2.41)	.127 (.109) (1.17)
<i>Right Govt.</i>	-.016 (.008) (1.98)	-.013 (.007) (1.86)	-.013 (.007) (1.79)	-.011 (.007) (1.66)	-.014 (.007) (1.95)
<i>GDP per capita</i>	-.021 (.100) (0.21)	.100 (.088) (1.13)	-.087 (.102) (0.85)	.009 (.087) (0.10)	.111 (.093) (1.19)
<i>Turnout</i>	-.072 (.022) (3.24)		-.059 (.022) (2.76)		
adjusted R^2	91.6	90.1	91.7	90.7	91.1
Number of obs.	49	49	49	49	48

Notes: OLS estimation. All regressions include a constant. The standard errors are estimated assuming that $E(\epsilon\epsilon') = \Sigma \otimes I_3$ where Σ is a general 18x18 variance-covariance matrix for the error terms associated with the 18 countries and I_3 is the 3x3 identity matrix (for 3 time periods). See Beck and Katz (1995) or Greene (1997: 651-669) for discussions of the estimation of robust standard errors.

shown in the second column.

Table 2 about Here

In specification (2), all of the estimated coefficients have the expected sign and all but GDP per capita are statistically significant at the conventional .95 level. The estimated impact of an increase of wage inequality by one standard deviation (roughly the difference between Sweden and Germany in 1995 or between Germany and the UK in 1995) is to reduce social insurance spending as a share of GDP by $(1.29)(.38) \approx 0.5$ percentage points in five years. In the long run, the impact of an increase of one standard deviation of wage inequality is to reduce social insurance spending as a share of GDP by $(1.29)(.38)[1/(1 - .719)] \approx 1.75$ percentage points. A similar reduction of social insurance spending is associated with a decrease in the share of the population over 65 of one percentage point or a reduction in the unemployment rate of two percentage points. The impact of five consecutive years of majority government by a conservative party, as opposed to five years of governments composed entirely of center and left parties, is estimated to reduce spending on social insurance as a share of GDP by $(.013)(100) = 1.3$ percentage points.

In columns (3) and (4), we present the parallel results when health expenditures are excluded from the dependent variable. The effect is to reduce the estimated impact of GDP per capita to zero and to increase the estimated impact of wage inequality. Finally, inspection of the residuals revealed that the case of Finland in 1995, an outlier with an unemployment rate of 17.4 per cent, has a significant impact on the estimates. Reestimating specification (2) with Finland 1995 removed from the data set results in the estimates of column (5). The exclusion of Finland 1995 reduces the estimated impact of unemployment on welfare spending by one-half and reduces the estimated impact of inequality by one-quarter. Whether or not Finland in 1995 should be excluded is not clear. On the one hand, the high level of unemployment in Finland in 1995 significantly increases the range over which unemployment varies in the data, thereby allowing a more precise estimate of the impact of unemployment of social insurance spending. On the other hand, estimates

that change significantly when a single case is removed must be treated with caution.

The point estimates of the coefficient on inequality imply that inequality is substantively important, whether or not we use the estimates of specification (2) or specification (5). The difference between the average log of the 90/10 wage premium in the US, the country with the most unequal wage distribution in the sample, and Norway, the country with the most equal wage distribution in 1995, is 1.38. Using the point estimates of columns (2) and (5) as lower and upper bounds, one obtains an estimate of the long-run impact of the difference of wage inequality in the US and Norway of between $(1.38)(1.29)[1/(1-.719)] \approx 6$ and $(1.38)(.96)[1/(1-.743)] \approx 5$ percentage points. Given that the actual difference between spending on social insurance and health care for those without earnings as a share of GDP is 7 percentage points (21.6 per cent is the average for Norway while 14.6 per cent is the average for the US), the difference in the inequality of wages and salaries in the US and Norway explains between 70 and 85 per cent of the difference in welfare spending.

Given the inclusion of a lagged dependent variable that alone explains 87 per cent of the variance, the results concerning the impact of inequality on social insurance spending are reasonably strong. The weight of the evidence indicates the OECD countries with more unequal distributions of wages and salaries spend less, not more, on social insurance and social assistance policies.

VI. Conclusion

In this paper, we develop the implications of the view of welfare policies as publicly financed insurance policies that pay out benefits relative to contributions in a redistributive manner. At the extreme ends of the income scale, the insurance aspect is dwarfed by the redistributive aspect. The poor in our model receive benefits and don't contribute at all while the rich have no need for publicly financed insurance. But the middle group of voters in our model benefit both from both the insurance and the redistribution that welfare policies provide.

The impact of increasing inequality on the political support for redistributed policies

depends critically on the way in which benefits are targeted when targeting is exogenous. In particular, greater inequality or greater uncertainty of income increases support for redistributive policies when benefits are mostly targeted to those with earnings. The intuition is that the wider the gap between the income of the decisive voter and the income of the average voter, the greater the incentive for redistribution. When benefits are mostly targeted to those without earnings, however, greater inequality of income reduces support for redistributive policies. Here the intuition comes from the insurance framework. When the income of the median voter falls, the median voter wishes to reduce his or her purchase of insurance against the possibility of income loss. Finally, we demonstrated, in the context of a model of electoral competition with two parties or two competing blocks of parties, that when voters choose both the type of targeting and the level of benefits, it is the intuition from the insurance model that prevails with regard to benefits targeted to those without earnings. Public support for benefits targeted to those without earnings goes down as inequality increases. Regression results indicating that higher inequality is associated with less spending on benefits targeted to those without earnings among OECD countries provide empirical support for the model.

The model also implies that the impact of redistribution from the rich is non-monotonic with respect to income inequality. A small increase in income inequality from an initial condition of relative equality reduces majority support for redistribution from the rich since, if incomes are relatively equal, voters prefer to target all benefits to those without earnings and support for benefits to those without earnings declines as the median voters' wage declines. An increase in inequality from an initial condition of sufficient inequality, however, increases majority support for redistribution from the rich to the middle. In highly unequal societies, the rich may have a good reason to fear democracy even while the poor may not find that democracy brings many benefits.

Endnotes

¹ The early literature on that emphasized the strength of social democratic parties and organized labor in explaining the expansion of the welfare policies is surveyed by Shalev (1983). See Esping-Andersen (1990), Hicks and Swank (1992), Hicks (forthcoming), Huber, Ragin and Stephens (1993) and Huber and Stephens (forthcoming) for more recent studies.

² For studies of social welfare as publicly provided insurance, see Barr (1992), Wright (1996), Sinn (1995), and Casamatta, Cremer and Pestieau (1998).

³ Our assumptions regarding the nature of the risk covered by social insurance are similar to those made by Wright (1996). In a more general model, the probability of obtaining employment, βdt , would be partly a matter of agents' efforts rather than a parameter. Making β endogenous is discussed briefly at the end of Section III.

⁴ For example, we could have $\tau(t) = t\bar{h}(t)$ where $\bar{h}(t)$ is the appropriately defined average number of hours worked (i.e. $\bar{h}(t) \equiv E[wh(t)]/\bar{w}$, with $\bar{h}(0) = 1$, $\bar{h}(1) = 0$, $\bar{h}'(t) < 0$ and $\bar{h}''(t) < 0$). Pre-tax earnings would then be $(1 - t)wh(t)$. Let utility be separable in consumption and leisure with the part that depends on consumption given by $u(c_E)$ as presented in equation (2). Then h can be held constant around its equilibrium value by the envelope theorem and all of the arguments go through with w replaced by wh wherever w appears. Alternately, we could assume that hours worked are fixed and that $t - \tau(t)$ represents the cost of collecting taxes, with the cost assumed to be an increasing, convex function of t .

⁵ Such an assumption is stronger than necessary. All of the results go through in a more general model in which the benefits targeted to those without earnings depend on past wages or contributions according to the formula

$$c_N = \left[\delta + (1 - \delta) \frac{w}{\bar{w}} \right] \frac{(1 - \gamma)T(t)}{\sigma_0 + (1 - \theta)\sigma_L}$$

for $0 < \delta \leq 1$. See Casamatta, Cremer and Pestieau (1998) for a paper that explores the impact of δ on voters' choice of t in a model with $\sigma_L = 1$ and $\gamma = 0$.

⁶ See, for example, the exchange between Skocpol (1991) and Greenstein (1991). For formal models that investigate the impact of targeting on the level of support for welfare spending, see De Donder and Hindriks (1998), Gelbach and Pritchett (1997) and Moene and Wallerstein (1999).

⁷ We conjecture that ξ is generally a monotonic declining function of γ , which would imply the existence of a unique $\gamma^* \in (0, 1)$ such that $dt/dw > 0$ for $\gamma < \gamma^*$ and $dt/dw < 0$ for $\gamma > \gamma^*$. However, we cannot prove that

$$\frac{d\xi}{d\gamma} = \frac{\partial\xi}{\partial\gamma} + \frac{\partial\xi}{\partial t} \frac{dt}{d\gamma} < 0$$

for all $0 < \gamma < 1$. While $\partial\xi/\partial\gamma < 0$, the sign of $\partial\xi/\partial t$ depends on the specification of $\tau(t)$.

⁸ The details of the mathematical analysis of the impact of an increase in workers' uncertainty on support for welfare spending can be found in Moene and Wallerstein (1998).

⁹ See Austen-Smith (1998) for a model with three constituency-based parties that compete along a single dimension.

¹⁰ The extent to which the parties converge to the ideal point of the median voters in this framework depends on the weight given the parties' core constituents in determining the parties' ideal points, ω_L and ω_R , the risk aversion of voters and the distribution of views regarding the parties' competence, $F(\epsilon)$.

¹¹ In Moene and Wallerstein (1998) we show that the same policy equilibrium can be generated by a model of voting in parliament in which each dimension of choice is under the jurisdiction of a different committee, as in Shepsle (1979).

¹² The countries included in the data set are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Sweden, Switzerland, UK and the US. The cases that are excluded because of missing data are Belgium 1980 – 84, Netherlands 1980 – 84, Portugal 1980 – 84 and Switzerland 1980 – 89. The method of calculating wage inequality for the US was changed in the early 1990s. To construct a continuous series starting in 1980 for the US, we used

the older series reported in OECD (1993) together with the extension of the old series reported in OECD (1996).

¹³ Pampel and Williamson (1989) emphasize the importance of the size of the groups receiving welfare benefits as determinants of the level of expenditures.

¹⁴ If one adds an interaction term between turnout and inequality, the coefficient on the interaction term is positive as found by Franzese (1998). However the impact of turnout remains negative at every level of inequality. The positive interaction indicates that turnout is less negatively associated with welfare spending, on average, at higher levels of inequality than at lower levels.

References

- Alesina, Alberto and Dani Rodrik. 1994. Distributive Politics and Economic Growth. *Quarterly Journal of Economics*, 109: 465-490.
- Austen-Smith, David. 1998. Redistributing Income under Proportional Representation. Northwestern University, Evanston IL: unpublished paper.
- Barr, Nicholas. 1992. Economic Theory and the Welfare State: A Survey and Interpretation. *Journal of Economic Literature*, 30: 741-803.
- Beck, Neal and Jonathan Katz. 1995. What To Do (and Not to Do) with Time-Series Cross-Sectional Data in Comparative Politics. *American Political Science Review*, 89: 634-647.
- Biais, Andre and Agnieszka Dobrzynska. 1998. Turnout in Electoral Democracies. *European Journal of Political Research*, 33: 239-261.
- Casamatta, Georges, Helmuth Cremer and Pierre Pestieau. 1998. On the Political Sustain-ability of Redistributive Social Insurance Systems. GREMAQ, University of Toulouse, France: Unpublished paper.
- Castles, Francis G. 1982. The Impact of Parties on Public Expenditure. In Francis Castles, ed., *The Impact of Parties: Politics and Policies in Democratic Capitalist States* (London: Sage Publications).
- Castles, Francis and Peter Mair. 1984. Left-Right Political Scales: Some 'Expert' Judgements. *European Journal of Political Research*, 12: 73-88.
- De Donder, Philippe and Jean Hindriks. 1998. The Political Economy of Targeting. *Public Choice*, 95: 177-200.
- Esping-Andersen, Gøsta. 1990. *The Three Worlds of Welfare Capitalism*. Princeton: Princeton University Press.
- Franzese, Robert J. 1998. Political Participation, Income Distribution and Public Transfers in Developed Democracies. The University of Michigan, Ann Arbor: unpublished paper.
- Freeman, Richard B. 1988. Labour Market Institutions and Economic Performance. *Economic Policy*, 3: 64-80.
- Friend, Irwin and Marshall E. Blume. 1975. The Demand for Risky Assets. *American Economic Review*, 65: 900-922.

- Gelbach, Jonah B. and Lant H. Pritchett. 1997. Indicator Targeting in a Political Equilibrium: Leakier can be Better. Washington DC: World Bank Policy Research Working Paper 1523.
- Greene, William H. 1997. *Econometric Analysis*, (Third Edition). Upper Saddle River, NJ: Prentice Hall.
- Greenstein, Robert. 1991. Universal and Targeted Approaches to Relieving Poverty: An Alternative View. In Christopher Jencks and Paul Peterson, eds., *The Urban Underclass*. Washington DC: Brookings Institution.
- Hicks, Alexander. Forthcoming. *Social Democracy and Welfare Capitalism: A Century of Income Security Politics*. Ithaca NY: Cornell University Press.
- Hicks, Alexander and Swank, Duane. 1992. Politics, Institutions and Social Welfare Spending in the Industrialized Democracies, 1960-1982. *American Political Science Review*, 86: 658-674.
- Huber, Evelyne and John D. Stephens. Forthcoming. *Political Choice in Global Markets: Development and Crisis in Advanced Welfare States*. Chicago: University of Chicago Press.
- Huber, Evelyne, Charles Ragin and John D. Stephens. 1993. Social Democracy, Christian Democracy, Constitutional Structure and the Welfare State. *American Journal of Sociology*, 99: 711-749.
- Huber, John and Ronald Inglehart. 1995. Expert Interpretations of Party Space and Party Locations in 42 Societies. *Party Politics*, 1: 73-111.
- Korpi, Walter. 1983. *The Democratic Class Struggle*. London: Routledge and Kegan Paul.
- Lijphart, Arend. 1997. Unequal Participation: Democracy's Unresolved Dilemma. *American Political Science Review*, 91: 1-14.
- Meltzer, Allan H. and Scott F. Richard. 1981. A Rational Theory of the Size of Government. *Journal of Political Economy*, 89: 914-927.
- Moene, Karl Ove and Michael Wallerstein. 1998. Rising Inequality and Declining Support for Redistribution. University of Oslo and Northwestern University: unpublished manuscript.
- Moene, Karl Ove and Michael Wallerstein. 1999. Targeting and Political Support for Welfare Spending. University of Oslo and Northwestern University: unpublished manuscript.

- Moffitt, Robert, David Ribar and Mark Wilhelm. 1998. The Decline of Welfare Benefits in the US: The Role of Wage Inequality. *Journal of Public Economics*, 68: 421-452.
- Organization for Economic Cooperation and Development. 1993. *Employment Outlook*, July 1993. Paris: OECD.
- Organization for Economic Cooperation and Development. 1994. *New Orientations for Social Policy*. Paris: OECD.
- Organization for Economic Cooperation and Development. 1996. *Employment Outlook*, July 1996. Paris: OECD.
- Organization for Economic Cooperation and Development. 1997. *Statistical Compendium 1997/2*, [CD-ROM]. Paris: OECD.
- Organization for Economic Cooperation and Development. 1999. *Social Expenditure Database, 1980-1996*, [CD-ROM]. Paris: OECD.
- Pampel, Fred C. and John B. Williamson. 1989. Welfare Spending in Advanced Industrial Democracies 1950-1980. *American Journal of Sociology*, 93: 1424-1456.
- Persson, Torsten and Guido Tabellini. 1994. Is Inequality Harmful for Growth? *American Economic Review*, 84: 600-621.
- Roberts, Kevin W. S. 1977. Voting over Income Tax Schedules. *Journal of Public Economics*, 8: 329-340.
- Roemer, John E. 1998. Why the Poor Do Not Expropriate the Rich in Democracies: A New Argument. *Journal of Public Economics*, 70: 399-424.
- Roemer, John E. 1999. The Democratic Political Economy of Progressive Income Taxation. *Econometrica*, 67: 1-20.
- Romer, T. 1975. Individual Welfare, Majority Voting, and the Properties of a Linear Income Tax. *Journal of Public Economics*, 14: 163-185.
- Shalev, Michael. 1983. The Social Democratic Model and Beyond: Two Generations of Comparative Research on the Welfare State. *Comparative Social Research*, 6.
- Shepsle, Kenneth A. 1979. Institutional Arrangements and Equilibrium in Multidimensional Voting Models. *American Journal of Political Science*, 23: 27-59.
- Skocpol, Theda. 1991. Targeting within Universalism: Politically Viable Policies to Combat Poverty in the United States. In Christopher Jencks and Paul Peterson, eds., *The Urban Underclass*, Washington DC: Brookings Institution.

- Stephens, John D. 1979. *The Transition from Capitalism to Socialism* (Atlantic Highlands, NJ: Humanities Press).
- Sinn, Hans-Werner. 1995. A Theory of the Welfare State. *Scandinavian Journal of Economics*, 97: 495-526.
- Summers, Robert and Alan Heston, 1991, The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988, *Quarterly Journal of Economics*, 106: 327-368.
- Swank, Duane. 1992. Politics and the Structural Dependence of the State in Democratic Capitalist Nations. *American Political Science Review*, 86: 38-54.
- Wallerstein, Michael. 1999. Wage-Setting Institutions and Pay Inequality in Advanced Industrial Societies. *American Journal of Political Science*, 43: 649-680.
- Wilensky, Harold L. 1975. *The Welfare State and Equality*. Berkeley: University of California Press.
- Wright, Randall. 1996. Taxes, Redistribution, and Growth. *Journal of Public Economics*, 62: 327-338.