

Mulligan, Casey B.

Working Paper

Economic Limits on "Rational" Democratic Redistribution

Working Paper, No. 171

Provided in Cooperation with:

George J. Stigler Center for the Study of the Economy and the State, The University of Chicago
Booth School of Business

Suggested Citation: Mulligan, Casey B. (2002) : Economic Limits on "Rational" Democratic Redistribution, Working Paper, No. 171, The University of Chicago, George J. Stigler Center for the Study of the Economy and the State, Chicago, IL

This Version is available at:

<https://hdl.handle.net/10419/262573>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Economic Limits on “Rational” Democratic Redistribution*

by

Casey B. Mulligan

The University of Chicago

March 2002

Abstract

I begin with an economic environment familiar from welfare- and political-economic literatures and show how, with quantitatively reasonable distributions of labor productivity and tax-price-elasticities of taxable income, middle class consumers are (personally) worse off with any negative income tax scheme than they would be with no redistribution at all. This finding has important implications for political-economic theories of redistribution, because it implies that the fully informed median voter cannot be expected to support programs of cash redistribution from rich to poor – such as the negative income tax – merely on the basis of his personal benefits from the program. It also implies that the “rational” median voter model of redistribution is, in the empirically relevant range, consistent with a zero correlation between income distribution skewness, or enfranchisement of the poor, and the amount of rich-poor redistribution. The paper also presents some comparable cross-country measures of one of the determinants of the tax-price-elasticity of taxable income, tax base breadth.

*I appreciate financial support of the Smith-Richardson Foundation and the Stigler Center for the Study of the Economy and the State, and the comments of Gary Becker, Sven Feldman, Jeff Milyo, Tomas Philipson, Raj Sah, Guido Tabellini, and GSB Macro Lunch eaters. Numerical calculations were made on a Quattro-Pro spreadsheet, and are available at <http://www.src.uchicago.edu/users/cbm4/>

Table of Contents

I. The Economic Environment	3
The Effect of Taxes on Taxable Income	3
The Laffer Curve: General Case	6
The Laffer Curve: Log Utility	8
II. Quantifying the Elasticity of Taxable Income and Productivity Distribution Skewness	10
Parametric Versions of Utility and Distribution Functions	11
Breadth of the Tax Base	13
Elasticities of Labor Supply and Taxable Income with Respect to the After-tax Share	15
.....	15
Income Variance and Skewness	17
“Nonredistributive” Government Spending	20
III. A Negative Income Tax Benefits Few and Costs Many	23
Calculations Derived Directly from the Empirical Estimates	23
Model Simulations	25
IV. Cash Flow Indicators of the Size of the Coalition Gaining From Redistribution	29
V. Conclusions	33
VI. Appendix I: Some Properties of the Distribution $F(x;\sigma,\rho)$	37
VII. Appendix II: Nonunitary Elasticities of Substitution	40
VIII. References	40

Almost two centuries ago, de Tocqueville (1835) made a prediction, or should I say expressed the fear, that democracy would lead to, among other things, excessive redistribution from rich to poor. Economists have relatively recently qualified his prediction (although perhaps only redirecting the fear) by pointing out that redistribution has aggregate costs and that these costs might be incorporated into the actions of political actors, thereby limiting the amount of redistribution. With this economic contribution,¹ both costs and benefits of redistribution had been specified, thereby permitting comparative static analyses of the amount of redistribution. For example, Brennan and Buchanan (1980), Becker (1983), and Stigler (1986) suggest that redistribution should grow in response to the reduction in the costs of redistribution, such as the reduction associated with the development of new and better means of taxation. Meltzer and Richard (1981), Tabellini (1992), Alesina and Rodrik (1994), Persson and Tabellini (1994), Benabou (1996a, p. 23), Krusell et al (1996, p. 489), and many others use median voter models to suggest that more skewness of the income distribution should lead to more redistribution, and quite a number of researchers have followed their suggestion by searching (without much success) for a positive correlation between redistribution or government spending and income skewness or inequality (e.g., Meltzer and Richard 1983, Benabou 1996a, Lindert 1996, Perotti 1996).

The purpose of this paper is to show that, contrary to the conclusions reached by the median voter literature, the voter with median income might not benefit from redistribution because the government also has some nonredistributive spending, so that even a small amount of redistribution is a first order distortion. Indeed, when nonredistributive spending is added to Meltzer and Richard's model, and the model has realistically parameterized utility functions and income distributions, the voter with median income prefers no redistribution to any negative

¹De Tocqueville seems aware of the costs of redistribution (eg., "Freedom...produces far more goods than it destroys" 1835, Book 1, Chapter 13), so perhaps the difference between his and many economic approaches is whether costs of redistribution are expected to affect the policies of democratic governments.

income tax policy, and his preferred policy is insensitive to marginal changes in the skewness of the income distribution.

My section I sketches the economic environment featured in Meltzer and Richard's (1981) model of rational democratic redistribution, which shares several key features with other models in the literature.² I show how the shape of the Laffer curve is an important determinant of the size of the coalition that can benefit from redistribution. Section II shows how, in turn, the Laffer curve derives from hypothesized effects of tax policy on taxable income (which depend in part on the amount of nonredistributive government spending), and the shape of the distribution of labor productivity. Section III shows that, with realistically parameterized utility and distribution functions, the median voter supplies some taxable income and is better off with no redistribution than he would be with almost any other government policy. In summary, when expressed in the median voter model, the economist's qualification of de Tocqueville's (1835) prediction is actually a reversal of that prediction – we should expect essentially no income redistribution in democracies, or at least in democracies situated in economies like those of the modern developed nations.

I, and many authors preceding me in the literature, model income redistribution with a linear tax and transfer schedule, when in fact tax rules are much more complicated. But this simplification has little consequence for my basic conclusion – that deadweight costs reduce the size of the poor coalition than can benefit from distribution, and that the coalition might be in the minority. To show this, Section IV proposes another test of the proposition that redistribution hurts the majority: whether or not the person with median taxable income is a net taxpayer. The test is applicable to nonlinear tax and transfer systems, and, although excessively powerful, can be conducted using results from the literature on the income incidence of taxes and transfers.

²e.g., Tabellini (1992), Alesina and Rodrik (1994), Persson and Tabellini (1994), Benabou (1996a), and Krusell and Rios-Rull (1999). Four key features shared by these models are: (1) redistribution is from rich to poor, (2) redistribution has aggregate costs in terms of distorting behavior away from taxable activities, (3) voters are aware of the costs of redistribution and take them into account in calculating their personal gains or losses from one policy or another, and (4) government policy relates only to the amount of redistribution, and the parameters of that policy are chosen by voters in a majoritarian election.

I. The Economic Environment

I.A. *The Effect of Taxes on Taxable Income*

With one modification, the economic environment is familiar from welfare- and political-economic studies of income redistribution (eg., Mirrlees 1971, Meltzer and Richard 1981)³; I will use Meltzer and Richard's (1981) notation. There are two sectors, which I call "work" and "leisure," and one period. Everyone is endowed with a unit of time and effort (hereafter "time"). Time can be devoted to leisure or work, or some combination of the two; l denotes leisure time and n denotes work time. Work time produces consumption goods at rate wx , where the productivity rate x varies continuously across persons according to the distribution function F . I normalize the median of F to be one, and let w denote the rate at which labor productivity in the taxed sector is translated into goods for consumption.⁴ Denoting "consumption" as c , we suppose that each person has the same utility function $u(c,l)$. c and l are supposed to be normal goods, u 's indifference curves are supposed to be convex, and Inada conditions apply as c or l approach 0.

When it comes to the economic environment, my departure from Mirrlees, Meltzer and Richard, and others is that not all labor product can be taxed for the purpose of redistribution. Mueller (1989, p. 330), Krusell et al (1996, pp. 497-98) and others have suggested that a "realistic" model would allow for various uses of taxation,⁵ and my purpose is to show that this particular dimension of realism has dramatic implications for the political-economics of redistribution. From the consumer's point of view, the relevant modification is that there are two consumption goods, one of which generates labor income tax deductions. Formally, the "consumption" c

³Tabellini (1992) shares several features with the Melzer-Richard model, except a labor supply distortion, and he adds an age dimension. Krusell and Rios-Rull (1999) have a dynamic version.

⁴This normalization is my one notational departure from Meltzer and Richard's (1981), and reduces the notation necessary to discuss numerical calibration of the model.

⁵A number of theories of government also suggest that costly redistribution could not occur in the absence of nonredistributive policy. For example, we might expect the rich to join a government with the power to tax only if some of that power were used for nonredistributive policies. Or the costs of redistribution could themselves enhance the political success of nonredistributive policy (as in Becker 1983).

referenced above is a composite good $c(c_1, c_2)$, where c is a homogeneous function of the amounts consumed of the two goods. Each consumer allocates his resources between taxed and untaxed goods in order to maximize his utility u subject to his budget constraint, which is determined by his “productivity” x and the parameters (t, r) of the negative income tax policy chosen by the government:

$$c_1 + c_2 = wxn - t(wxn - c_2) + r \quad (1)$$

where r denotes a “guaranteed minimum income” or “universal benefit,” and t is a scalar tax rate. r is provided by the government, and is financed from labor income tax revenue, which is collected in the amount $t(wxn - c_2)$ from a person with productivity x , taxable time n , and deductible consumption c_2 . All people consume at least r , and consume $c > r$ to the extent that their income after taxes exceeds zero. Since those with little taxable income are net beneficiaries from the government, and those with lots of taxable income are net taxpayers, the tax system embedded in equation (1) is sometimes called a “negative income tax” (hereafter, NIT).

Given the constraint (1), let $n[r/p(1-t), (1-t)wx/p]$ denote the utility maximizing time devoted to the taxable sector:

$$\begin{aligned} n[r/p(1-t), (1-t)wx/p(1-t)] &\equiv \underset{n \in [0,1]}{\operatorname{argmax}} u([(1-t)wxn + r]/p(1-t), 1-n) \\ 1/p(1-t) &\equiv \max_{c_2} c(1 - (1-t)c_2, c_2) \end{aligned} \quad (2)$$

where we again see how the departure from Meltzer-Richard has an effect on the relative price of consumption goods $p(1-t)$ in addition to the usual effect of taxes on the after tax wage rate.

A second reason that all labor income is not taxable for the purpose of redistribution is that governments spend resources on public goods, administration, etc. Let g denote this spending, expressed as an amount per consumer. It follows that the government faces the constraint (3) in choosing the two parameters (t, r) of its negative income tax system – namely,

aggregate spending on the universal benefit r and the nonredistributive spending g may not exceed the revenue from the proportional part of the tax:⁶

$$r + g \leq t \int_0^{\infty} \{wxn[r/p(1-t), (1-t)wx/p(1-t)] - \lambda(t)((1-t)wxn[r/p(1-t), (1-t)wx/p(1-t)] + r)\} dF(x) \quad (3)$$

$$\lambda(t) \equiv \underset{c_2}{\operatorname{argmax}} c(1 - (1-t)c_2, c_2)$$

where $\lambda(t)$ is the fraction of consumption expenditure $[(1-t)wxn+r]$ that is spent on the tax-deductible good. The government budget constraint (3) has expenditure per consumer on the left-hand-side, and tax revenue per consumer on the right-hand-side. The latter is a proportion t of labor product (the integral's first term) minus deductible consumption (the integral's second term). Equation (3)' algebraically simplifies (3), showing how allowing for tax-deductible consumption effectively increases the government's price of the universal benefit (ie, the term $1+t\lambda$ multiplying r), and decreases the fraction of labor income that can be collected even for the nonredistributional good g (ie, the term in square brackets multiplying the integral).

$$[1 + t\lambda(t)]r + g \leq [1 - (1-t)\lambda(t)]tw \int_0^{\infty} xn[r/p(1-t), (1-t)wx/p(1-t)] dF(x) \quad (3)'$$

It is important for the analyses of de Tocqueville (1835), Meltzer and Richard (1981), and many other studies of income redistribution that the relevant dimension for redistribution is rich vs poor, because this leads to a perfect correlation between the factors determining coalition formation and the factors determining income tax liability. Hence "nonredistributive" spending

⁶In addition to putting nonredistributive spending g in the budget constraint, g could also affect utility or wages without changing my results, as long as g did not affect the marginal rate of substitution between consumption and leisure.

g should be interpreted to include not only purchases of goods and services that might be in the general interest, but probably also transfers (such as old age pensions) and other policies that are financed with income taxes and may only benefit a fraction of the population. “Nonredistributive” spending may well be determined by political factors (eg., as modeled by Browning 1975), but the point of my analysis and those cited above is to study the political economics of rich-poor redistribution holding those other policies constant.

I.B. The Laffer Curve: General Case

I assume that some redistribution is feasible, which in technical terms means that g is small enough that there exists a t solving equation (3)' for some $r > 0$. As Meltzer and Richard (1981, footnote 8) explain,⁷ the normality of leisure implies that the function $n[\cdot]$ is nonincreasing (and strictly decreasing for $n > 0$) in its first argument, so that the government budget constraint (3) implicitly defines the universal benefit as a function $r(t)$ of the marginal tax rate. This function $r(\cdot)$, commonly known as the Laffer curve, is an important determinant of the politics of redistribution, so my next sections numerically study that curve. In turn, $r(\cdot)$ is determined by the functions $n[\cdot]$, $c(\cdot)$ and $F(\cdot)$ – the responses of taxable activity to tax policy, and the productivity distribution, respectively.

An important question raised by any political-economic analysis of redistribution via a NIT is: “Who, if anyone, benefits from a NIT?” In order to see the determinants of the answer, consider the indirect utility function $v(t, r(t); x)$ implicitly defined by the labor supply program (2). Consider the total derivative $dv(t, r(t); x)/dt$ in the neighborhood of the no-universal-benefit policy $(t_0, 0)$, where t_0 is the tax rate needed to finance nonredistributive spending alone (ie, the minimum t solving equation (3) with $r = 0$). This derivative is of some interest because we expect the gains to a consumer from redistribution, if any, to be the greatest with the first dollar of redistribution.⁸ In other words, any consumer with $dv(t, r(t); x)/dt \leq 0$ in the neighborhood of $(t_0, 0)$ is better off with the no-universal-benefit policy $(t_0, 0)$ rather than with any positive

⁷Their explanation still applies despite my addition of the terms g , $p(t)$, and $\lambda(t)$ because, given t and g , government expenditure is monotonic increasing and government revenue nonincreasing in r .

⁸Meltzer and Richard (1981, p. 920) assume this.

redistribution. The Inada conditions on u imply that we can use the envelope theorem to calculate the total derivative in this neighborhood, and show that it is positive if and only if:

$$r'(t_0) > \beta(t_0) w x n [0, (1-t_0) w x / p(1-t_0)] \quad (4)$$

$$\beta(t) \equiv 1 - \frac{d \ln p(1-t)}{d \ln (1-t)}$$

The left hand side is the slope of the Laffer curve for the first dollar of redistribution. β is the elasticity of the after-tax real wage $(1-t)w/p(1-t)$ with respect to the after-tax share $(1-t)$, so the right hand side of (4) can be interpreted as the amount of income of a type x consumer that is taxable at the margin.

The slope of the Laffer curve has a lot of determinants, except in some special cases. One of those includes those studied by Mirrlees, Meltzer and Richard, and others, who consider the case where all government spending is redistributive ($t_0 = 0$). Here $r'(t_0) = r'(0)$ is merely the average taxable income in the economy, and $\beta(t_0) = 1 - \lambda(0)$. In this case, equation (4) says that anyone with taxable income less than the average benefits from at least some redistribution. The point of my analysis is that, when some government spending is not redistributive, $t_0 > 0$ and some consumers with less than average taxable income prefer no redistribution. In other words, nonredistributive spending reduces $r'(t_0)$ below average taxable income. Moreover, existence of tax deductions magnifies this effect, because the Laffer curve slope declines more quickly with g and with r . So, for a given g , an economy with a narrower tax base will have a flatter Laffer curve for the first dollar of redistribution, and hence fewer consumers who can benefit from the first dollar of redistribution. Indeed, since we know that the Laffer curve slope is eventually zero, 20 or 30% reductions in its slope due to nonredistributive spending and tax base narrowness would not be extraordinary.

Let $T(t,r;x)$ be taxable income of a consumer with productivity x facing the policy (t,r) , and $\bar{T}(t,r)$ the cross-section average taxable income under that policy. Equation (4)' divides both sides of equation (4) by average taxable income in the absence of any government spending

$(\bar{T}(o,o) = r'(o))$:

$$\frac{r'(t_o)}{r'(o)} > \frac{\beta(t_o)}{1 - \lambda(t_o)} \frac{\bar{T}(t_o, o)}{\bar{T}(o, o)} \frac{T(t_o, o; x)}{\bar{T}(t_o, o)} \quad (4)'$$

the left-hand side is the amount by which nonredistributive spending and tax base narrowness reduce the slope of the Laffer curve. The right-hand side has three ratios. The first is the ratio of the marginal to average fraction of income that is taxable. The second is the effect of nonredistributive spending and tax base narrowness on aggregate taxable income. The third is a type x consumer's taxable income in the absence of redistribution as a fraction of average. If the consumer in question is the median ($x = 1$), the third term is the inverse of the skewness of the taxable income distribution in the absence of redistribution. Hence, whether the median consumer prefers some redistribution to none depends on whether the skewness of the taxable income distribution, combined with the first two terms on the RHS, is enough to overwhelm the reduced slope of the Laffer curve due to nonredistributive spending and tax base narrowness. The next subsection considers the case of log utility, for which the first two terms on the RHS are one, and the slope of the Laffer curve is simply calculated.

I.C. The Laffer Curve: Log Utility

An important determinant of the slope of the Laffer curve is the compensated microeconomic taxable income elasticity with respect to the after-tax share $(1-t)$, $\eta(t,r;x)$, which is calculated by differentiating $\ln[T(t,r;x)]$ with respect to $\ln(1-t)$, taking into account the amount by which r must be decreased in order to keep utility constant. This elasticity, and hence the effect of nonredistributive spending on the size of the coalition that might benefit from redistribution (hereafter, "the NIT coalition"), can be simply calculated for a productivity x worker when utility is logarithmic: $u(c_1, c_2, l) = \beta \ln c_1 + (1-\beta) \ln c_2 + (1/n^* - 1) \ln l$, where β and

n^* are preference parameters:⁹

$$\eta(t, r; x) = \frac{(1-t)\beta x + r/w}{(1-t)\beta x/(1-\beta n^*) + r/(wn^*)} \quad (5)$$

Notice that the taxable income elasticity is constant across persons in the neighborhood of the universal-benefit policy $(t_0, 0)$: $\eta(t_0, 0; x) = (1-\beta n^*)$. Let the scalar $\eta_0 \in (0, 1)$ denote the taxable income elasticity at this point. Log utility implies that a worker with productivity x has taxable income $T(t, r; x) = (1-\eta_0)wx - \eta_0 r/(1-t)$. Since the Laffer curve $r(t)$ is defined by $r(t) + g = t \int_0^\infty T(t, r(t); x) dF(x)$, we have a closed form solution for the Laffer curve's slope at t_0 :¹⁰

$$r'(t_0) = \frac{\beta n^* w E(x)}{1 + \frac{t_0}{1-t_0} \eta_0} \quad (6)$$

where $E(x)$ is average productivity, as well as the ratio of mean to median productivity (remember that F is normalized so that its median is one). The numerator is average taxable income in the neighborhood of the universal-benefit policy $(t_0, 0)$. The denominator reflects the inframarginal loss in tax revenue due to a reduced tax base. Hence, in the case where no government spending is redistributive ($t_0 = 0$), the denominator is one and the slope of the Laffer curve equals average taxable income.

Combining (4) and (6), we have the condition for a type x log-utility consumer to gain from redistribution:

$$x < \frac{E(x)}{1 + \frac{t_0}{1-t_0} \eta_0} \quad (7)$$

⁹Log utility implies that the elasticity $\beta(t)$ discussed above not longer varies with t . The parameter n^* turns out to be the amount worked by someone with log utility and receiving transfer $r = 0$.

¹⁰The fact that the government can afford g implies that $\eta_0 t_0 / (1-t_0) < 1$.

The size of the NIT coalition is the fraction of consumers for whom (7) is satisfied, and grows with the skewness of the productivity distribution. More elastic taxable income, or a higher tax rate needed to finance nonredistributive programs, decrease the size of the NIT coalition.

Might the NIT coalition be a minority? Equation (7) shows that the answer requires numerical values for the skewness of the productivity distribution $E(x)$, the tax rate needed to finance nonredistributive government spending t_o , and the after-tax-price elasticity of taxable income η_o . Section II discusses reasonable values for these parameters.

II. Quantifying the Elasticity of Taxable Income and Productivity Distribution Skewness

I suggest that numerical versions of the model should be compared with six observations:

1. the ratio of observed average labor income tax rates to statutory rates (ie, measures of tax base breadth)
2. observed elasticities of taxable income with respect to the after tax share $(1-t)$
3. the relative constancy of labor hours per capita over long periods of time or across broad country cross-sections
4. the skewness of the income distribution
5. the variance of the income distribution
6. the quantity of government spending that does not redistribute from rich to poor

The first three observations relate to the utility parameters n^* , β , and θ . Items 5 & 6 relate to the distributional parameters σ and ρ (defined below), and observation 4 to the model's parameter g . Equation (7) shows that items 2 and 4 are of direct relevance for determining the size of the NIT coalition, while the other items matter as they determine the tax rate t_o needed to finance nonredistributive government spending.

One difficulty in comparing the model to observations is that the former is presumably about long run, lifetime incomes and effort (since the model has only one period) while most observations of tax elasticities and income distributions are for one calendar year. My discussion below tries to address these issues and relate them to the interpretation of my results.

II.A. Parametric Versions of Utility and Distribution Functions

In order to numerically characterize the Laffer curve, the functions $n[\cdot]$, $c(\cdot)$ and $F(\cdot)$ (or, equivalently, the functions $u(\cdot)$, $c(\cdot)$ and $F(\cdot)$) must be represented numerically. The purpose of this section is to do so in a way consistent with the theory and consistent with observed income distributions and observed responses to public policy. My first step is to parameterize u , c and F as follows:

$$u(c, l; (1-n^*)w^{\theta-1}/n^*, \theta) \equiv \begin{cases} \frac{\theta}{\theta-1} \ln \left[c^{(\theta-1)/\theta} + \left(\frac{1-n^*}{n^*} w^{\theta-1} \right)^{1/\theta} l^{(\theta-1)/\theta} \right] & \text{if } \theta \neq 1 \\ \ln c + \frac{1-n^*}{n^*} \ln l & \text{if } \theta = 1 \end{cases}$$

$$c(c_1, c_2; \beta) \equiv c_1^\beta c_2^{1-\beta} \quad (8)$$

$$F(x; \sigma, \rho) \equiv \begin{cases} \Phi\left(\frac{\ln x}{\sigma}\right) & \text{if } \Phi\left(\frac{\ln x}{\sigma}\right) \leq \rho \\ 1 - (1-\rho) \left(e^{\sigma \Phi^{-1}(\rho)/x} \right)^{\Phi(\Phi^{-1}(\rho))/[(1-\rho)\sigma]} & \text{if } \Phi\left(\frac{\ln x}{\sigma}\right) > \rho \end{cases}$$

where ϕ , Φ , Φ^{-1} are the standard normal density, distribution, and inverse distribution functions, respectively. Notice that u is a two parameter constant elasticity of substitution function. It is important that one of these parameters, θ , dictates the elasticity of substitution, since that elasticity has deservedly received the most attention in empirical studies of the Laffer curve and other aspects of the supply of taxable income. $(1-n^*)w^{\theta-1}/n^*$ is the second parameter, which I have written this way so that n^* is the labor supplied by the median consumer in the absence of taxation. β also affects the supply of taxable income, because it dictates preferences for goods whose expenditures are deductible from taxation. F is a two parameter function, with median $x = 1$ when $\rho \geq 1/2$, and is a composite of the lognormal and Pareto distributions.¹¹ lognormal distributions have been used in other numerical studies of this economic environment (eg., Mirrlees 1971, p. 193f), but my analysis stresses the importance of the skewness of F , so

¹¹ F becomes lognormal as $\rho \rightarrow 1$ and Pareto as $\rho \rightarrow 0$. See Appendix I for more properties of this distribution.

composing lognormal and Pareto permits numerical exploration of distributions more skewed but otherwise similar to the lognormal.

Let Γ be the parameter space to be explored with my numerical calculations. In other words, each numerical calculation draws $\{w, \beta, \theta, n^*, \sigma, \rho, g\} \in \Gamma$, where Γ is:

$$\Gamma \equiv \{(w, \beta, \theta, n^*, \sigma, \rho, g) \mid \\ w = 1, \beta \in [0.6, 0.9], \theta = 1, n^* \in \{[0.6/\beta, 0.1/\beta] \cap [0.08, 0.8]\}, \\ \rho \in [0.7, 0.1], \sigma \in [(\rho - 0.7)/3 + 0.4, (\rho - 0.7)/3 + 0.5], g \in [0.1wn^*, 0.4wn^*]\}$$

The “benchmark parameters” are $\gamma_b \equiv \{(w, \beta, \theta, n^*, \sigma, \rho, g) = (1, 0.75, 1, 0.5, 0.5, 0.72, 0.175)\}$, and are intended to error on the size of overestimating the size of the NIT coalition. The “alternate parameters” $\gamma_a \equiv \{(w, \beta, \theta, n^*, \sigma, \rho, g) = (1, 0.65, 1, 0.5, 0.5, 0.8, 0.175)\}$ and are intended to produce my best estimate of the size of the NIT coalition for postwar OECD countries.

The analysis below explains why I believe Γ to either: (a) include the parameter space empirically relevant to a study of redistribution by the governments of OECD member countries, or (b) exclude those parameters only at the cost of *overstating* the benefits of redistribution to the middle class. Since some of these parameters are not directly estimated here or in the public finance literature, Table 1 summarizes implications of Γ for estimates that can be found there.

Table 1: Public Finance Calculations Implied by the Parameter Space Γ			
observable functions of model parameters	evaluated at:		range implied by parameter space
	benchmark (γ_b)	alternate (γ_a)	
	parameters		
Micro effects of taxes on behavior			
compensated microeconomic taxable income elasticity at $x=1$ & $r=0$, η_o	.63	.68	[.4,.9]
average tax rate/marginal tax rate at $t = 0$, β	.75	.65	[.5,.9]
Frisch labor supply elasticity, θ	1	1	1
Characteristics of the productivity distribution			
std. deviation of log productivity	.55	.52	[0.5,0.6]
ratio of mean/median productivity	1.25	1.19	[1.13,1.26]
skewness relative to lognormal	1.06	1.03	[1,1.08]
fraction with less than avg product.	.67	.64	[.60,.68]
Government budget			
nonredistributive public expenditure/GDP for $r = 0$.28	.29	[.07,.35]
total public expend./GDP for $t = 0.5$.36	.32	[.17,.45]
maximum feasible public expenditure share of GDP	.45	.40	[.19,.70]
Addendum: largest coalition that can benefit from a NIT policy:	.42	.30	[.02,.66]
<p>Notes: (1) The benchmark parameters are $\gamma_b \equiv \{(w,\beta,\theta,n^*,\sigma,\rho,g) = (1,0.75,1,0.5,0.5,0.72,.175)\}$. (2) The alternate parameters are $\gamma_a = (1,0.65,1,0.5,0.5,0.8,0.175)$ (3) The “maximum feasible expenditure share of GDP” maximizes with respect to the tax rate, holding fixed all parameters (including β).</p>			

II.B. Breadth of the Tax Base

Tax bases are rarely as broad as they are in the economic environments modeled by Mirrlees (1971), Meltzer and Richard (1981), Tabellini (1992), Gallasso and Ruiz (1999) and many

others. Quantifying this fact is difficult in general, but there are some interesting cases where an accurate calculation is possible. Consider, for example, Social Security payroll taxes for which a single statutory tax rate $t > 0$ typically applies to some labor income while a zero rate applies to the rest of labor income (eg., labor income above the Social Security cap, labor income paid as fringe benefits, or labor income earned in uncovered sectors). Table 2 reports, for 12 OECD countries with available 1995 data, the ratio of Social Security payroll tax revenue collected to $(t \cdot \text{labor income})$.¹² Since $t \cdot \text{labor income}$ is the revenue that would be collected if the rate t applied to ALL labor income, the ratio calculated in the table can be interpreted as the fraction of labor income that is taxable by the payroll tax in these countries.

country	fraction	country	fraction
Belgium	0.905	Netherlands	0.616
Czech Republic	0.854	Norway	0.886
France	0.950	Portugal	0.473
Italy	0.576	Spain	0.772
Japan	0.698	UK	0.580
Korea	0.369	US	0.750

Notes: (1) Countries reported are OECD countries with 1995 Social Security payroll taxes, minimal nonpayroll Social Security revenue sources, and 1995 compensation of employees reported in OECD *National Accounts 1984-96*. 1994 data is used for the Czech Republic and 1993 data for Portugal
(2) Fraction of labor income taxed is calculated as $(t \cdot \text{labor income}) / (\text{SS payroll tax revenue})$
(3) The Statutory Social Security payroll tax rate is calculated as $(t_e + t_f) / (1 + t_f)$, where t_e and t_f are Social Security payroll tax rates (for old age, survivors, disability, sickness and maternity programs) for employee and employer, respectively, as reported by SSA (1995).
(4) The cross-country average fraction is 0.70.

¹² t is calculated from SSA's *Social Security Programs Throughout the World 1995* as $(t_e + t_f) / (1 + t_f)$, where t_e and t_f are Social Security payroll tax rates (for old age, survivors, disability, sickness and maternity programs) for employee and employer, respectively.

The ratios in the table are significantly less than one, so that a significant fraction of labor income is untaxed. I model this fact above with the second consumption good (c_2) that is deductible from labor income in the calculation of taxable income. Notice that β is exactly the fraction of labor income tax that is taxable as taxes and transfers become small, and that β exceeds this fraction for tax rates significantly greater than zero. On the other hand, labor income tax bases are more narrow than suggested by Table 2 because the calculations there exclude an important fraction illegal labor income that is not included in the national accounts. Hence, values of β that are consistent with Table 2 are undoubtedly less than one, and probably in the range covered by the Table.

Table 2 probably overstates β , because it measures only payroll taxes, and personal income taxes are less broad-based than are payroll taxes. In the U.S., for example, federal personal income tax revenue has recently been about 10% of aggregate personal income,¹³ but hardly anyone faces a federal marginal tax rate less than 10%. The calculations of Barro and Sahasakul (1983) and Stephenson (1995) suggest that this pattern is typical of the entire history of the federal personal income tax: cross-section average statutory tax rates have exceeded the tax's share of personal income by a factor of more than two. Even the payroll tax has been less broad than suggested by Table 2: less than half of aggregate wages and salaries were subject to payroll taxes prior to the mid 1970's (Barro and Sahasakul 1983). With these caveats of Table 2 in mind, I use a range of $\beta \in [.5,.9]$ for my numerical calculations. My benchmark value of β is 0.75, and "alternate value" is 0.65.

II.C. Elasticities of Labor Supply and Taxable Income with Respect to the After-tax Share

Since the model is often applied to economies with quite different aggregate labor productivity w but not very different labor supply per capita (e.g., the 20th century American changes studied by Meltzer and Richard 1983 – a century during which real output per manhour grew by almost a factor of 7 or 8 but manhours per capita by a factor of only 1.2; or across countries as studied by Benabou 1996a), it seems that θ must be pretty close to one so that labor productivity's wealth and substitution effects on aggregate labor supply cancel, or that the model

¹³Tax revenues from OMB (2001, Table 2-1) and personal income from BEA (2001, Table 2.1).

be amended to allow for other determinants of labor supply that would cancel the substitution (wealth) effect that would otherwise dominate behavior with $\theta > 1$ ($\theta < 1$). I use $\theta = 1$ for the algebraic derivations and numerical calculations remaining in the main text, and explore other values of θ in Appendix II.

Even with $\theta = 1$, the model is still consistent with a variety of long run responses of lifetime taxable income to changes in the after-tax share $(1-t)$, although it is necessarily inconsistent with some of the higher estimates. There are a number of estimates of taxable income elasticities in the public finance literature, although these are typically estimates of short run responses. For example, Slemrod (1998) surveys studies of short run, compensated, microeconomic responses of taxable income to $(1-t)$,¹⁴ which I denote η , and suggests that estimates range from 0.4 to 1.8. It is not clear whether long run compensated lifetime responses would be smaller or larger. Since my results are stronger when η is larger, I consider a range $[0.4, 0.9]$ for η which overlaps only the lower end of Slemrod's report of the empirical range.¹⁵

In order to see how $\eta \in [0.4, 0.9]$ defines boundaries of the parameter space, recall that, in the model with $\theta = 1$, a typical worker with productivity x has taxable income $T(t, r; x) = \beta w x n^* - (1 - \beta n^*)r / (1 - t)$. The compensated microeconomic taxable income elasticity with respect to the after-tax share $(1-t)$, $\eta(t, r; x)$, is calculated above (equation 5), and its formula is reproduced here for the reader's convenience:

$$\eta(t, r; x) = \frac{(1-t)\beta x + r/w}{(1-t)\beta x / (1 - \beta n^*) + r / (w n^*)} < 1 \quad (5)$$

Notice that log utility ($\theta = 1$) rules out taxable income elasticities greater than or equal to one – and hence the higher estimates found in the literature. If the higher estimates are right, my

¹⁴As I note below, these estimates are typically for groups of people with positive taxable income.

¹⁵Of course, there is debate about the magnitude of the elasticity. Among the studies surveyed by Slemrod (1998), Feldstein (1995) has the higher estimates. Recent work by Goolsbee points to the lower end of Slemrod's surveyed range. Hence, I keep my calculations conservative by using numbers from the lower end of the range.

calculations will overstate the size of the NIT coalition because they overstate the slope of the Laffer curve for the first dollar of redistribution.

Log utility also implies $\eta(t_o, o; x) = (1 - \beta n^*)$, so we need n^* less than $0.6/\beta$ and larger than $0.1/\beta$ in order for η_o to be in the range $[0.4, 0.9]$.

II.D. *Income Variance and Skewness*

There are two distributions which are relevant for my calculations (which are equivalent in the log utility model): the distribution of lifetime-average productivity and the distribution of lifetime-average taxable income in the absence of redistribution. Two statistics of these distributions are especially relevant, the coefficient of variation and the skewness. I am not aware of precise calculations of lifetime income or productivity variance and skewness, but I point to five calculations that can be used to make an educated guess:

- skewness of family incomes in one year cross-sections (from CPS)
- skewness of family incomes in one year cross-sections partitioned by age (from CPS)
- variance and skewness of multiyear-averaged male wage rates (from PSID)
- variance and skewness of multiyear-averaged family incomes (from PSID)
- variance and skewness of multiyear-averaged family incomes (from SCF)

The Census Bureau (2000) reports mean and median family incomes for each of the years 1967-99. Over the period 1967-92, the ratio of mean to median varied from 1.11 to 1.21. The ratio has varied from 1.26 to 1.28 since 1992, when the CPS increased top-codes for measuring earnings and other incomes.¹⁶ A year is not a lifetime, so we might expect income inequality and skewness to be different when annual income is measured rather than lifetime income. For example, if all person's had the same life-cycle income profile, we would have inequality in a year's income due to the inequality across age groups, but no inequality in lifetime incomes. The Census Bureau calculated mean and median family incomes for samples partition by age of

¹⁶The top-code increased from \$299,999 to \$999,999 in 1993. The top-code was also increased in 1985, but this change is not substantially correlated with increases in family income skewness.

householder. There is less skewness in the each of the three prime age groups 25-34, 35-44, and 45-54 than there is in the distribution pooling all families. For example, the mean-median ratio for the 35-44 group varies from 1.08 to 1.15 prior to 1993, and 1.20 to 1.23 after 1993.

There are many microeconomic studies looking at labor and total incomes time-averaged over several years. Mulligan (1999) reports a standard deviation of log two year average wages of 0.51 for men aged 26-34 (years 1990 and 1991). I use his data to calculate the skewness of the same distribution, and it is 1.14 – a distribution about equally skewed as the CPS distribution of annual family incomes among those headed by a person aged 25-34 (mean/median=1.12) and about equally skewed as a lognormal distribution (mean/median=1.14 when $sd(\ln x) = 0.51$). Mulligan (1997, Appendix F) reports statistics which can be used to calculate a standard deviation of log 5 year average (1967-71) hourly wages of 0.58 in a sample of PSID-SRC household heads with children and a standard deviation of log 5 year average (1967-71) of household income of 0.54 in a sample of PSID-SRC households with children. His reported statistics can be used to calculate a standard deviation of log 5 year average (1984-88) hourly wages of 0.55 in a sample of PSID-SRC men aged 27-36 and a standard deviation of log 5 year average (1984-88) of household income of 0.67 in a sample of PSID-SRC households with heads aged 27-36. I use his data to calculate mean-median ratios for the four aforementioned sample distributions of 1.11, 1.11, 1.12, and 0.99, respectively.¹⁷ Given the standard deviation of logs reported above, these distributions are slightly less skewed than the lognormal.

Income skewness – the gap between mean and median in particular – is crucial for my analysis, and may not be well measured in a samples such as the PSID or CPS that top-code incomes and/or may undersample the very rich. The 1983-89 household panel from the Survey of Consumer Finances (SCF) may be better suited for this purpose, because it over-samples rich households, does not top-code, and inquires about incomes in four different calendar years (1983, 1986-88). The SCF provides weights for estimating population statistics, but even the weighted

¹⁷The PSID can also be used to directly assess the hypothesis that multi-year wage and income distributions are less skewed than single-year distributions. There is a slight tendency in this direction, with single year measures skewed 1.12 or 1.13 while multi year measures are skewed 1.11.

average real household income is too high in the sample: 53,357 1996 dollars.¹⁸ But the SCF panel may still tell us something about income skewness relative to the lognormal. Consider the following ratio:

$$\frac{\left(\frac{\text{mean income}}{\text{median income}} \right)}{e^{[\text{sd}(\ln \text{income})]^2/2}}$$

where income is average for each household over the four years reported. This ratio would be one if incomes were distributed lognormal, so it is an index of income distribution skewness relative to the lognormal distribution. In the SCF panel – limited to households with heads aged 33-61 in order to mitigate life cycle biases – the (weighted) standard deviation of log income is 0.789, implying that the denominator above is 1.37. The ratio of (weighted) mean to median income is 1.40, so that the observed income distribution is 2% more skewed than lognormal (1.02 = 1.40/1.37).

The SCF panel measures household incomes in the 1980's and in the United States, so it probably overstates that amount of income inequality and income skewness¹⁹ that might be found in a typical postwar OECD economy. Nor are distributions of income or wages necessarily the same as distributions of productivity, although these are the same in the model with log utility and no redistribution. However, equation (4)' suggests that what matters for the size of the NIT coalition is the skewness of taxable income in the absence of redistribution. This distribution may be much less skewed – or even skewed to the left (ie, with median greater than mean) – because payroll tax caps might be interpreted as the nontaxation of higher incomes.

Considering these various caveats, the calculations above suggest productivity distribution parameters in the ranges $\sigma \in [0.4, 0.6]$ and $\rho \in [0.7, 1]$. The benchmark parameters

¹⁸The Census Bureau (2000) reports mean family incomes for the years 1982, 1986-8 of \$50, 425. Averaging real household income across SCF households and over the four years without weights yields an average of 237,851 1996 dollars.

¹⁹That skewness has trended significantly over time is consistent with the 1968-99 CPS data on annual family incomes.

are $\sigma = 0.5$ and $\rho = 0.72$, which imply $\text{sd}(\ln x) = 0.55$ and 6% more skewness than the lognormal. The “alternate” parameters are $\sigma = 0.5$ and $\rho = 0.8$, which imply $\text{sd}(\ln x) = 0.52$ and 3% more skewness than the lognormal. Appendix I reports more details on how these two parameters determine $\text{sd}(\ln x)$ and the skewness of F relative to the lognormal.

II.E. “Nonredistributive” Government Spending

Perhaps the most important parameter in my calculations is g , the amount of “nonredistributive” government spending. Since governments are involved in so many activities, and the income incidence of those activities can be complicated, any measurement of g is subject to some doubt. Table 3 reports four measures for 26 OECD countries with available 1992 data. The first measure is government consumption, and is of interest because it excludes the transfer payments which are presumably of primary relevance to the theory.²⁰ It varies from 7 to 23% of GDP.

²⁰Even if government consumption expenditures were primarily designed to tax rich and subsidize poor, they would have deadweight costs above and beyond those modeled here (because a cash transfer is preferred, by the beneficiary, to government purchases), and hence my analysis would overstate the attractiveness of such policies to the middle class.

country	government consumption	government expenditure minus:		
		Social Spending	SSp-health	SSp-health-OA
Australia	0.14	0.23	0.29	0.32
Austria	0.13	0.26	0.31	0.41
Belgium	0.10	0.29	0.36	0.43
Canada	0.13	0.31	0.38	0.42
Czech Republic	na	0.33	0.38	0.44
Denmark	0.19	0.31	0.37	0.44
Finland	0.18	0.26	0.33	0.42
France	0.15	0.24	0.31	0.41
Germany	0.13	0.22	0.30	0.40
Greece	na	0.16	0.19	0.27
Iceland	0.19	na	na	na
Ireland	0.12	0.22	0.27	0.31
Italy	0.11	0.28	0.35	0.46
Japan	0.07	0.21	0.26	0.31
Korea	na	0.18	0.20	0.20
Luxembourg	0.10	na	na	na
Mexico	0.09	na	na	na
Netherlands	0.11	0.28	0.35	0.42
Norway	0.19	0.25	0.32	0.38
Portugal	na	0.29	0.34	0.39
Spain	0.13	0.24	0.30	0.37
Sweden	0.23	0.32	0.39	0.47
Switzerland	0.09	0.30	0.36	0.42
Turkey	0.11	na	na	na
United Kingdom	0.18	0.22	0.28	0.35
United States	0.13	0.21	0.27	0.32

The three other measures make various subtractions from government *expenditure*: “social expenditure,” “social expenditure” except health, and “social expenditure” except health and old age. Social expenditure by the government is measured by the OECD (1997) and includes: old age cash benefits, disability cash benefits, occupational injury and disease program spending, sickness benefits, services for the elderly and disabled, survivors benefits, family cash benefits,

training and employment programs, unemployment benefits, health programs, housing benefits, low income programs, and immigrant & refugee programs. Since some of these programs could arguably be designed to subsidize the poor, my second measure subtracts all of them from government expenditure. Perhaps health spending is more about redistribution to healthcare workers and/or old rather than the poor, or perhaps the poor are not particularly interested in consuming health care,²¹ so my third measure leaves them included in “nonredistributive” government spending. A similar argument can be made about old age cash benefits, so my final measure subtracts leaves these and health program spending in “nonredistributive” government spending.²² These estimates range from 16 to 47% of GDP.

The point of these estimates is not to make an accurate estimate of the income incidence of all government programs taken together, just to reiterate Mueller’s (1989, p. 330) suggestion that there is a significant amount of government spending that neither has the primary intention nor primary effect of taxing the rich in order to help the poor. My estimates suggest that reasonable numerical calculations of a model like Meltzer and Richard’s should look at $g/(wn^*)$ in the range of [0.1,0.4]. This range roughly corresponds to a range of 10-40% of GDP – remember that GDP varies with the amount of redistribution and that wn^* is median labor income per capita in the absence of any taxation.

My analysis hinges on the concavity of the Laffer curve, so it would be nice to have direct estimates of that concavity. I am not aware of such estimates in the literature, although a model like mine allows us to infer Laffer curve concavity from estimates of taxable income elasticities and some of the other observations above. Another empirical check on the model is whether the laffer curve is not so concave to render infeasible the fairly high levels of public expenditure observed for some countries. To make this check I calculate two statistics at various points in the parameter space:

²¹See McClellan and Skinner (1997) for some estimates that suggest this effect is of primary relevance for evaluating the income-incidence of Medicare.

²²Note that my fourth measure still nets out spending on some programs that may be designed to help the elderly: disability cash benefits, occupational injury and disease program spending, sickness benefits, services for the elderly and disabled, survivors benefits, and housing benefits.

- the model's implied amount of public expenditure at $t = 0.50$, as a fraction of aggregate income
- the model's implied maximum amount of public expenditure, as a fraction of aggregate income

With the benchmark parameters, a 50% rate of taxation of taxable income leads to a total government budget (ie, the sum of nonredistributive spending g and redistributive spending r) of 36% of GDP. The share, which depends on the responses of both GDP and taxable income to tax rates, ranges from 17% to 45% at various points in the parameter space Γ . The maximum feasible government spending share of GDP is 45% with the benchmark parameters and 70% at other points in the parameters space Γ .

III. A Negative Income Tax Benefits Few and Costs Many

Numerical versions of the model can be used to calculate the size of the NIT coalition and the magnitude of the median consumer's preferred tax rate. Subsection A uses some formulas from the log case and some of the numerical parameters estimates references above to make a direct, although approximate, calculation of the size of the NIT coalition. Section B uses full parameter vectors to simulate the full model, exactly calculate the size of the NIT coalition, and of the median consumer's preferred tax rate.

III.A. Calculations Derived Directly from the Empirical Estimates

Remember that many of the model parameters are relevant for the size of the NIT coalition only through their effects on t_o , the rate at which taxable income is taxed by the government needing only to finance only nonredistributive spending g . With log utility, t_o is:

$$t_o = \frac{\frac{g}{wn^+ E(x)}}{\beta}$$

t_o might be computed directly from the numbers in Table 3, since the numerator is the ratio of nonredistributive government spending to GDP for an economy with no redistribution, except

that economies with redistribution are expected to have reduced their GDP as a result of their redistribution. Ignoring this distinction for a moment, we can estimate the size of the NIT coalition with the benchmark values of β , η_o , $E(x)$, and the nonredistributive spending share of GDP:

$$\text{NIT coalition size} = F\left(\frac{1.25}{1 + \frac{0.37}{0.63} \cdot 0.63}\right) = F(0.91) < 0.5$$

where 1.25 is the skewness $E(x)$ implied by the benchmark parameters, 0.37 is a guess of t_o (nonredistributive spending share of 0.28 divided by $\beta = .75$), and 0.63 is the benchmark taxable income elasticity. Regardless of the parameters of the productivity distribution F , this coalition is a minority because F has been normalized so that its median is one. Even well beyond the border of the parameters space with $E(x) = 1.4$ – significantly more skewness than lognormal, and more than the evidence suggests – the coalition is barely a majority ($F(1.02)$) if we stick with the benchmark values of η_o and t_o . If we take the smallest value of the taxable income elasticity Slemrod found in his survey of the literature, 0.4, the coalition is barely a majority with the other benchmark values ($F(1.01)$).

I chose the benchmark parameter vector γ_b so as to be within the range of empirical estimates, but also to error in terms of overestimating the size of NIT coalition. The alternate parameter vector γ_a was chosen to generate my best guess of the size of NIT coalition, given the other model assumptions, by using values from the middle of the range of empirical estimates. A direct estimate of the size of the NIT coalition with the alternate parameter values of β , η_o , $E(x)$, and the nonredistributive spending share of GDP:

$$\text{NIT coalition size} = F\left(\frac{1.19}{1 + \frac{0.43}{0.57} \cdot 0.68}\right) = F(0.83) < 0.5$$

where 1.19 is the skewness $E(x)$ implied by the alternate parameters, 0.43 is a guess of t_o

(nonredistributive spending share of 0.28 divided by $\beta = .65$), and 0.68 is the alternate taxable income elasticity. Beginning from the alternate parameters, it is hard to create a majority NIT coalition by changing the parameter vector in only one dimension. For example, the smallest value of the taxable income elasticity Slemrod found in his survey of the literature, 0.4, implies a coalition size of $F(0.91)$ with the other alternate values. Even at a border of the parameters space with $E(x) = 1.4$, the NIT coalition is a minority ($F(0.98)$) if we stick with the alternate values of η_0 and t_0 . With $E(x) = 1.19$, $\eta_0 = 0.68$, and $\beta = 0.65$, t_0 must be less than 0.22 – and nonredistributive spending's share of GDP less than 0.14 – for the NIT coalition to be a majority.

III.B. Model Simulations

Figure 1 plots the redistribution Laffer curve $r(t)$ for three economies. The solid line corresponds to the benchmark economy (ie, with parameter vector γ_b). The dash-dot line is for the economy with the same parameters as the benchmark economy, except with $g = 0$ (and hence with parameter vector $\gamma_0 \notin \Gamma$). We see how eliminating nonredistributive government spending shifts up the redistribution Laffer curve, so that its slope is steeper at $r = 0$. As I argue qualitatively above, this Laffer curve slope is why eliminating nonredistributive government spending enlarges the coalition of consumers that might benefit from a NIT.

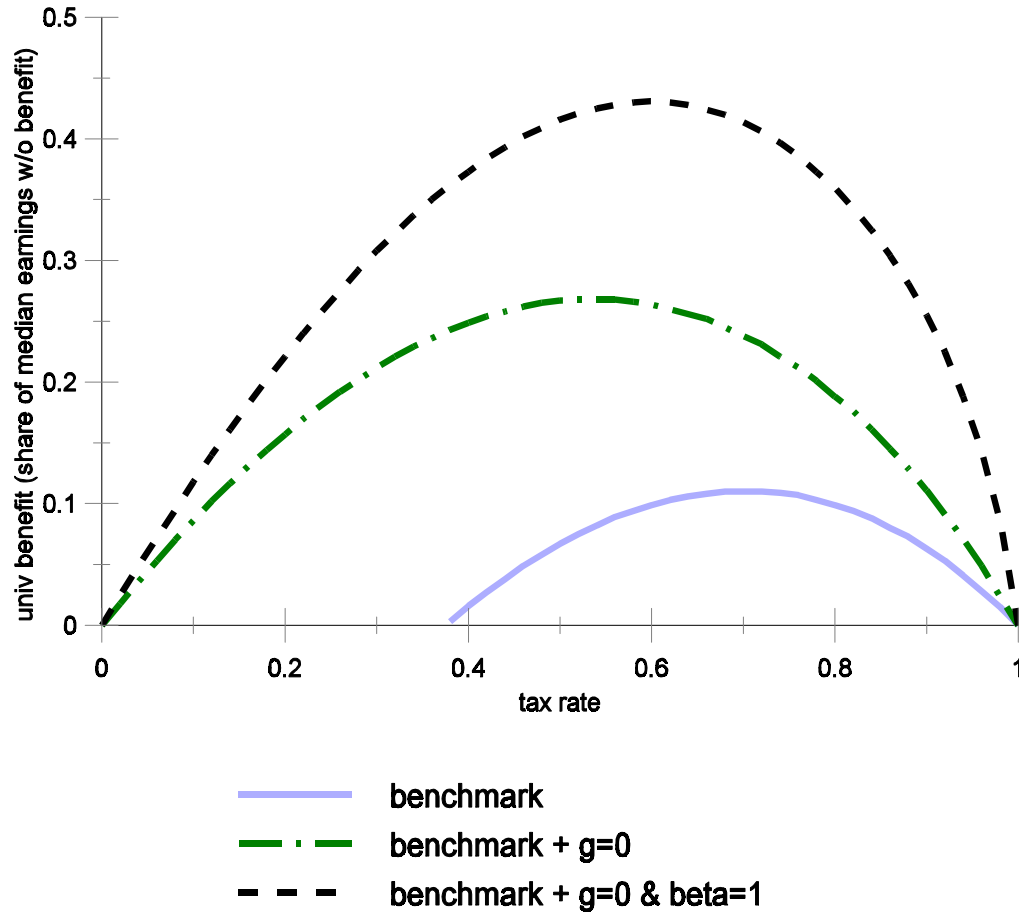


Figure 1 Nonredistributive spending, tax base breadth, and the redistribution Laffer curve

The dashed line is for the economy with the same parameters as the benchmark economy, except with $\beta = 1$ and $g = 0$ (and hence with parameter vector $\gamma_1 \notin \Gamma$), and thereby for an economy like those considered by Mirrlees (1971) and Meltzer and Richard (1981). We see that the Laffer curve is more concave with a narrow tax base (dash-dot line) than with a broad tax base (dashed line). Less concavity of the Laffer curve enlarges the coalition of consumers who can benefit from significantly more redistribution than zero.

To see how unrealistic are Laffer curves derived from models with $\beta = 1$ or $g = 0$, relative to the benchmark Laffer curve, consider their implications for the year 2000 U.S. economy. Let the median earnings in the absence of redistribution be \$39,000 per year. With a tax rate of 50%

– 12 percentage points more than needed to finance nonredistributive spending – the universal benefit in the benchmark economy would be 7% of median earnings in the absence of redistribution (the 7% is read off the solid curve in the Figure). 7% of median earnings is redistribution of \$2730 per adult per year. There were 197,522,000 American adults in the year 2000, so this is \$539 billion of redistribution per year!

Hence, while my benchmark Laffer curve looks quite low in comparison to those implied by Mirrlees and Mettler-Richard's model, \$539 billion is not a trivial amount of redistribution for the U.S. economy. To put it another way, consider the amount of redistribution that a 50% rate could bring without $g > 0$ (the dashdot line; parameters γ_0) or without $g > 0$ and $\beta < 1$ (the dashed line; parameters γ_1). The parameters γ_0 and γ_1 imply 4.5 and 7 times the redistribution as in the benchmark economy or \$2.5 trillion and \$3.8 trillion, respectively.

Not surprisingly, the NIT coalition is pretty large when there exist tax rates that can raise \$3 or 4 trillion for redistribution. The parameters γ_0 and γ_1 both imply coalitions of 67%, as compared to 42% for the benchmark economy. With the parameters γ_0 and γ_1 , the median consumer prefers tax rates of 20% and 28%, respectively. Compare these rates with the 0% rate preferred by the majority of consumers in the benchmark economy.

Consider a border of the parameters space creating the largest NIT coalitions: $w = 1$, $\beta = 0.9$, $\theta = 1$, $n^* = 0.67$, $\sigma = 0.5$, $\rho = .7$. Along this border, the compensated micro elasticity of taxable income with respect to the after-tax price, η_o , is 0.4 – at the bottom of the range surveyed by Slemrod. Mean/median productivity is 1.26, the productivity distribution is 8% more skewed than lognormal, and 68% of consumers are less productive than average. Any economy with parameters on this border, and a value of g in the parameter space, has a NIT coalition that is a majority. In other words, the median consumer's preferred redistributive spending over GDP is positive. Figure 2 graphs this preferred transfer spending versus nonredistributive spending's share of GDP.²³

²³In other words, a point in the graph is generated by: (1) picking the parameters above, (2) picking some g in the parameters space, (3) calculating the tax rate t preferred by the median consumer, and (4) graphing the quantity of redistributive spending r and nonredistributive spending g as fractions of aggregate income produced under the policy t, r, g .

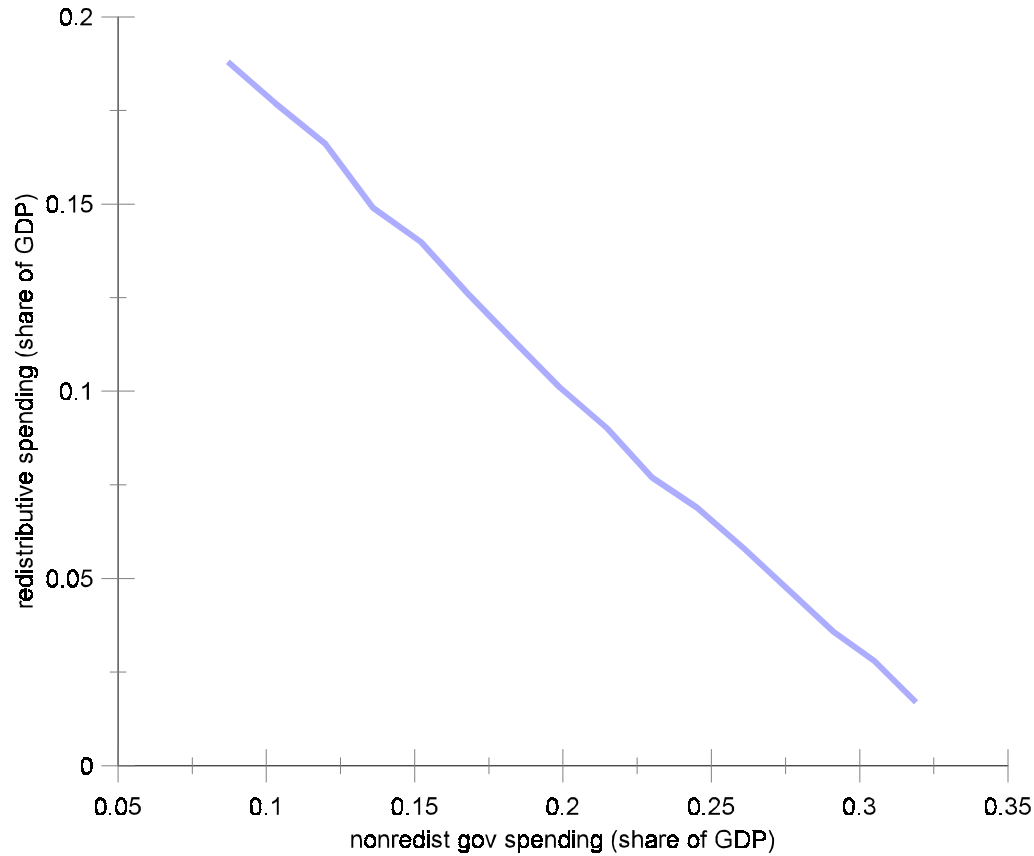


Figure 2 Median-Preferred Redistribution with Extreme Parameter Values

Not surprisingly, median-preferred transfer spending declines with nonredistributive spending. But the interesting result reported in the Figure is that – even at the border of the parameter space where, given the various empirical estimates, the NIT benefits so many people as to strain credibility – preferred redistributive spending is a much smaller share of GDP than we observe in OECD countries.²⁴ Remember that total government spending is 40% of GDP or more in those countries. As discussed above, it is hard to say exactly how much of this spending is “nonredistributive,” but suppose it were 15% of GDP. Then the Figure suggests that, with

²⁴The result can also be seen in Galasso and Ruiz (1999), where the model government spends resources on pensions as well as rich-poor redistribution, and thereby relatively little rich-poor redistribution is found in the political equilibrium. Roughly speaking, their model can be considered an “extreme” parameterization of mine, because their tax bases are broad, and because the deadweight costs of nonpension-nonredistributive government policies are ignored.

these extreme parameter values, the median consumer would prefer 14% of GDP spent on redistributive programs rather than the 25% of GDP or more observed. With 25% of GDP nonredistributive, the median consumer would prefer 7% of GDP spent on redistribution, rather than the 15% of GDP or more observed.

The NIT coalition can be very small at the opposite border of the parameter space, where $w = 1$, $\beta = 0.5$, $\theta = 1$, $n^* = 0.2$, $\sigma = 0.5$, $\rho = 1$, and $g = 0.08$. Along this border, the compensated micro elasticity of taxable income with respect to the after-tax price, η_o , is 0.9 – in the middle of the range surveyed by Slemrod. The productivity distribution is lognormal, with mean/median productivity equal to 1.13, and 60% of consumers are less productive than average. Only 2% of consumers can benefit from a NIT in this economy!

IV. Cash Flow Indicators of the Size of the Coalition Gaining From Redistribution

My results suggest that negative income tax policies, conducted in economies where the government has nonredistributive obligations and tax bases as narrow as in OECD countries, will hurt middle class consumers. The purpose of this section is to relate my result to taxpayer cash flows, with an analysis that applies to any utility function with the assumed regularity properties.

Consider the effect of NIT program (t,r) on the net real NIT subsidy $S(t,r;x)$ of a consumer with productivity x . $S(t,r;x)$ is computed graphically in the $[l,c]$ plane as the vertical distance between the line $c = wxn(1-t_o)/p(t_o)$ (which is the consumer's budget constraint with the no-universal-benefit policy $(t_o,0)$) and the solution to the consumer's program given (t,r) , as in Figure 3.

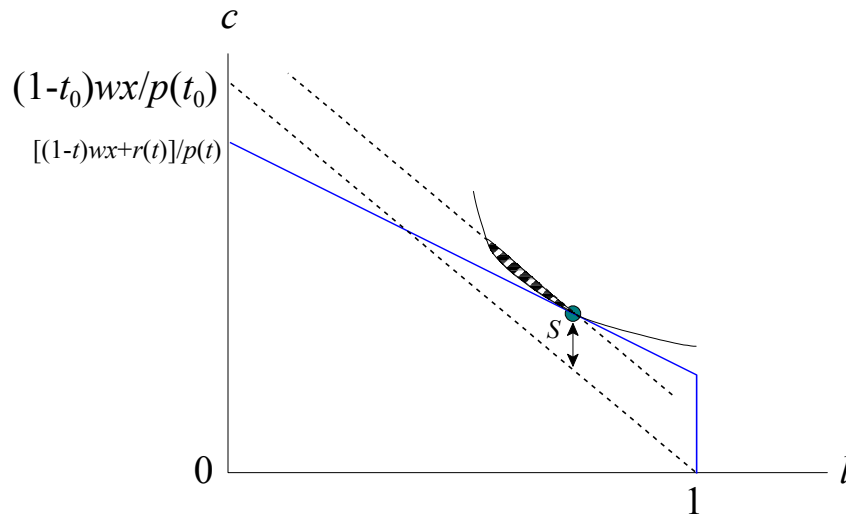


Figure 3 Welfare Implications of the Net Subsidy

It can be shown that, when $dS(t,r(t);x)/dt < 0$, then any consumer with productivity x is worse off with a slightly higher tax rate. To see this, consider a position like the solid dot that is optimal for some initial policy (t,r) , and a marginal increase in the tax rate from that initial policy. $dS(t,r(t);x)/dt < 0$ means that the new optimal choice is below the dashed line drawn through the solid dot with slope $-wx(1-t_0)/p(t_0)$. If we suppose that the consumer is better off, then the new allocation must be in the shaded area and involve less leisure which, because of the convexity of preferences and the normality of consumption, means that the marginal rate of substitution has increased, which in turn contradicts the supposed increase in the tax rate. Hence, we have proved by contradiction that $dS(t,r(t);x)/dt < 0$ implies reduced utility from $dt > 0$. A similar argument can be constructed to show that a consumer with productivity x and $S(t,r(t);x) < 0$ is better off with the no-universal-benefit policy $(t_0,0)$ than with policy $(t,r(t))$. A similar argument also applies to tax systems that are not linear.

The algebraic expression for $S(t,r(t);x)$ is complicated, but it is positive if and only if the

consumer's net cash flow vis-a-vis the government is greater than it would be without redistribution. Hence, an upper bound on the fraction of the population benefitting from a redistributive policy is the fraction of consumers with more positive cash flows as a result of that policy. Since this is an upper bound (some people with positive cash flows may be harmed by the policy because they changed their behavior in order to gain cash flow), we have a powerful test of my result and a weak test of the median voter models of redistribution in the literature: does a majority of consumers (or a majority of voters) have more positive lifetime cash flow as a result of redistributive policy?

Carefully conducting this test is beyond the scope of this paper, but the calculations like those of Musgrave et al (1974), Pechman and Okner (1974), Reynolds and Smolensky (1977), and Peckman (1985) might be used in conducting such a test for the U.S., or those of Ruggles and O'Higgins (1981) for the U.K. Some of Pechman's (1985) calculations are reproduced in Figure 4 below. Pechman aligns households by their market incomes for the year, and graphs their 1980 taxes, transfers, and net taxes as a fraction of market incomes. We see that 30% of U.S. households received more in transfers than they paid in taxes in 1980. Since 20 million of 60 million families²⁵ received Social Security Old Age and Survivors benefits in 1980 – which were paid to the old regardless of their lifetime incomes – and thereby especially likely to receive more transfers than they paid in taxes in 1980, a conservative guess is that 10 percentage points of those 30% do not enjoy a positive net lifetime cash flow from means-tested programs.

²⁵Total OAS beneficiary families from SSA (2000, Table 5.H.1). Total families from Council of Economic Advisers (1995, Table B-31).

Federal, State, and Local Transfers and Taxes as a Percent of Market Income, by Population Percentile, 1980^a

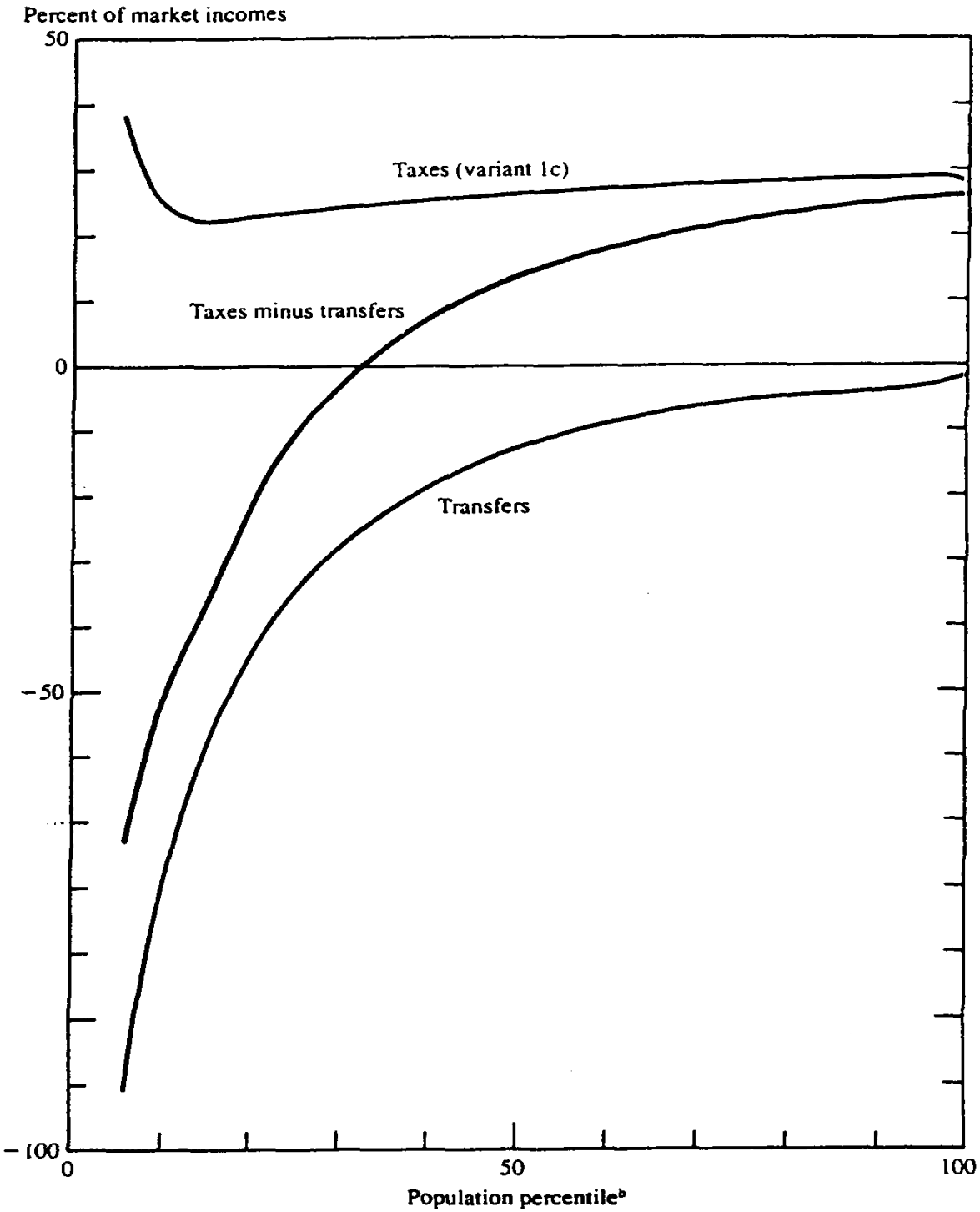


Figure 4 Overall US Taxes and Transfers, 1980. [Source: Pechman (1985), Fig 4-2, p 54.]

On the other hand, not all taxes counted by Pechman were for redistributive programs. But if we look at the 50th or 60th percentile (remember the 10 percentage points or more in the left half of the Figure who are not lifetime poor) in Pechman's Figure, we see that transfers are less than one third of taxes, so that total transfers are less than taxes paid to redistributive programs unless those taxes are less than one third of total taxes. Hence, it seems that lifetime incomes well below the median are required if one is to receive more in one's lifetime from redistributive programs than he pays in taxes toward those programs.

V. Conclusions

I begin with an economic environment familiar from welfare- and political-economic literatures, and show how the shape of the redistribution Laffer curve has an important effect on the size of the coalition that can benefit from redistribution. Inequality (4)' is my main analytical result in this regard, so I repeat it below as inequality (4)", in the simplified form used in much of my numerical analysis. The inequality holds for a type x consumer if and only if that consumer is a member of the NIT coalition:

$$\frac{r'(t_0)}{r'(0)} > \frac{T(t_0, 0; x)}{\bar{T}(t_0, 0)} \quad (4)''$$

where $r(t)$ is the redistribution Laffer curve: the relation between guaranteed minimum income r and marginal tax rate t . $T(t, r; x)$ is the taxable income of a consumer with productivity x facing the policy (t, r) , and $\bar{T}(t, r)$ cross-section average taxable income under that policy. The ratio on the left is the effect of tax base narrowness and nonredistributive spending on the slope of the Laffer curve for the first dollar of redistribution. The ratio on the right is type x consumer's taxable income as a fraction of cross-section average taxable income, in the absence of redistribution. The NIT coalition is larger when the Laffer curve is steeper for the first dollar of redistribution, or the taxable income distribution is more skewed.

I then show how the shape of the redistribution Laffer curve is determined by the microeconomic compensated elasticity of taxable income with respect to the after-tax price, the amount of nonredistributive government spending, the breadth of the tax base, and the skewness of the lifetime income distribution. The first item has been estimated in the literature, and I present some estimates of the other determinants. The estimates suggest that the coalition of consumers that might benefit from redistribution is small, so that the median consumer – and most of the middle class – is better off with no redistribution. In terms of the inequality (4)”, these numbers suggest that the Laffer curve slope is reduced at least 25% by tax base narrowness and nonredistributive spending.

A corollary to this result is that costs of redistribution not only qualify de Tocqueville's (1835) prediction of excessive democratic redistribution, they reverse it – at least when those costs are included in the median voter model. We should expect essentially no income redistribution in democracies, or at least in democracies situated in economies like those of the modern developed nations. Moreover, the median voter model does *not* predict that the amount of redistribution is related to income distribution skewness, or to the fraction of the poor that is enfranchised.

V.A. Possible Implications for Voting-Based Political Theories of Redistribution

In making my argument, I have taken the median voter model, and its applications to income redistribution found in the economic literature, at face value. Others have questioned the other implications of that model, such as whether the median voter has less than mean income (eg., Nelson 1999), or whether it is rational for voters to vote in their self-interest (eg., Brennan and Lomansky 1983), or whether voters have much influence on policy however they vote, or whether much government spending can be understood as redistribution from rich to poor. Hence, while “rational” democratic models of income redistribution like Meltzer and Richard’s imply that democracies will not redistribute from rich to poor, these questions seem important enough to conjecture that other models of the public sector would derive different implications. Perhaps the necessary extension is merely to suppose that the middle class have

significant altruism for the poor.²⁶ Or, if I am right that the majority are hurt by a negative income tax, perhaps political economic theory should unify its explanation of rich-poor redistribution with its explanations of tariffs, price supports, and other policies that also benefit a few at the expense of many.

Much, but not all, of the economic literature stresses the efficiency costs of redistribution, and my analysis maintains that emphasis. But it has been argued that redistributive policy can enhance efficiency. One reason for this is that redistributive policies may provide significant insurance. Or policy might intermediate savers and investors by taxing the cash-rich and paying the cash poor who have access to high return investment projects, such as investments in children. Or the rich may be sufficiently altruistic (Hochman and Rogers 1969) or paternalistic (Olsen 1969), that the rich have higher utility when they are taxed to fund poverty programs. These arguments can be used to derive redistributive policies in voting models (eg., Benabou 1996b describes a case where all people agree to income-tax-financed schooling programs), but they can be used to derive redistributive policies in any public choice model in which efficient choices have advantages over inefficient ones (eg., the interest group model of Becker 1983, or the leviathan model of Buchanan and Congleton 1979). The challenge for the median voter model is to explain redistribution from rich to poor even when that redistribution has aggregate efficiency costs.

Perhaps the median of an uninformed voting populace would approve of redistribution in the economic environment studied in this paper, because he does not fully understand the costs of taxation? While it seems plausible that voters might not be fully informed, and that policy incidence can be complicated, a voter ought to know something about his personal cash flows, and my Section IV shows how (in the economic environment modeled) this information can be enough to determine whether a consumer is worse off as a result of redistributive

²⁶However, we also expect deadweight costs to discourage altruistic redistribution (eg., Becker and Mulligan 1998). For example, in the model above with the benchmark parameters, the median voter's utility loss from a small amount of redistribution is 70% of the average utility gain, which means that an altruistic median voter favoring redistribution would be weighting others' utility at least 70% as much as his own. It seems unlikely that the median voter would have so much altruism, given that most estimates of the degree of altruism of parents for their *own children* (eg., Becker 1993) are 60% or less.

government programs. Namely, any consumer who pays more into transfer programs than he receives in transfer payments is (personally) better off without government redistribution. This calculation seems fairly easy, at least if redistributive taxes were earmarked and the timing of taxes and transfers were not too complicated.

V.B. Other Implications

De Tocqueville's (1835) fear and median voter formalizations of it have motivated a number of empirical studies of the correlation between redistribution or government spending and income skewness or inequality (e.g., Meltzer and Richard 1983, Benabou 1996, Lindert 1996, Perotti 1996). A positive correlation is hard to find, and some (eg., Peltzman 1980) have even suggested that the correlation is negative.²⁷ My results suggest that, if the middle class has an important weight in policy-making as they do in the median voter model, one might expect a zero correlation because, in the empirically relevant range, the middle class prefers exactly zero redistribution regardless of the gap between mean and median income.

Determinants of redistribution, and public policy more generally, might be partitioned into two categories: political and economic. The former include the institutions of public decision-making, such as voting, while the latter include deadweight costs and other variables determining feasibility constraints, and tax incidence. While economists boldly study both, they must admit that De Tocqueville and many other noneconomists have some skill at analyzing the effects of political variables. My study is an illustration of the economist's comparative advantage in the study of deadweight costs, and that the results of such a study can be dramatically different than results derived from a study that ignores them. Since I stress the incidence of deadweight costs, and how incidence relates to political support, perhaps my results also illustrate that there may be important interactions between the political and economic variables.

Redistributive taxation problems are analytically quite similar to optimal risk sharing

²⁷Persson and Tabellini (1994) suggest that much rich-poor redistribution may occur via off-budget regulatory policies, and lead to a weak correlation between on-budget redistribution and inequality even when total redistribution is related to inequality. If this is true, my results are still relevant, showing how a majority of citizens cannot be expected to support on-budget redistribution on the basis of their personal benefits.

programs (eg., Fudenberg and Tirole 1991 interpreted Mirrlees' 1971 problem in this way). Deriving the political implications of risk sharing when tax bases are narrow and there are nonredistributive/non-risks-sharing government expenditures is beyond the scope of this paper, but my analysis does suggest that introducing these factors would significantly increase the cost of risk sharing without enhancing the benefits.

My analysis – and those of Meltzer and Richard (1981), Tabellini (1992), Alesina and Rodrik (1994), Persson and Tabellini (1994), Benabou (1996), and others – take the tax system as given, but the β preferred by the median consumer in my model is one. In other words, while the median consumer prefers no redistribution in my benchmark economy, he would prefer to have $\beta = 1$ and to spend 4% of GDP on redistribution.²⁸ How can tax deductions survive in the political environment described in the literature? I expect that answering this question is difficult in a median voter model with income as the relevant dimension, especially when we recognize that an important reason for tax base narrowness is that the payroll taxes used to fund transfer programs are *capped* in most countries (according to SSA 1995, they are capped in 7 of the 12 countries studied in my Table 2).

I have followed Mirrlees (1971) and Meltzer and Richard (1981) by studying a model with two composite goods interpreted as “consumption” and “leisure” and a single tax interpreted as a labor income tax, but the main issues and many of the calculations also apply to models of current consumption, future consumption, and capital taxes, such as the models of Bertola (1993), Persson and Tabellini (1994), and Benabou (1996, p. 23). All of these studies assume that taxes are broad-based, and that the first dollar of redistribution has negligible marginal deadweight costs. When these assumptions are relaxed, it is a numerical question whether the median voter, or the middle class more generally, can benefit from linear redistribution programs, and the public finance literature provides many of the numbers required to formulate an answer.

VI. Appendix I: Some Properties of the Distribution $F(x;\sigma,\rho)$

²⁸Krusell et al (1996) do not model tax base narrowness as I do, but their results may suggest that a median voter in a dynamic model might prefer $\beta < 1$. Nevertheless, I think the question remains as to why, if the median voter wants redistribution, why he is satisfied with tax bases as narrow as observed.

As discussed in the text, I parameterize the distribution of productivity as:

$$F(x; \sigma, \rho) \equiv \begin{cases} \Phi\left(\frac{\ln x}{\sigma}\right) & \text{if } \Phi\left(\frac{\ln x}{\sigma}\right) \leq \rho \\ 1 - (1 - \rho) \left(e^{\sigma \Phi^{-1}(\rho)/x} \right)^{\phi(\Phi^{-1}(\rho))/[(1-\rho)\sigma]} & \text{if } \Phi\left(\frac{\ln x}{\sigma}\right) > \rho \end{cases}$$

where ϕ , Φ , Φ^{-1} are the standard normal density, distribution, and inverse distribution functions, respectively. The median of F is 1 as long as $\rho \geq 0.5$. Clearly, when $\rho = 1$, F is lognormal with the σ^2 equal to the variance of $\ln(x)$. As ρ becomes small, F approaches the Pareto distribution with shape parameter $\phi(\Phi^{-1}(\rho))/[(1-\rho)\sigma]$ and scale parameter $(1 - \rho)^{-\phi(\Phi^{-1}(\rho))/[(1-\rho)\sigma]} e^{\sigma \Phi^{-1}(\rho)}$. I leave it to the reader to show that the density of F is continuous and integrates to one.

The Figure below graphs the density of F corresponding to $\sigma = .5$, and $\rho = 0.6, 0.75$, and 1.

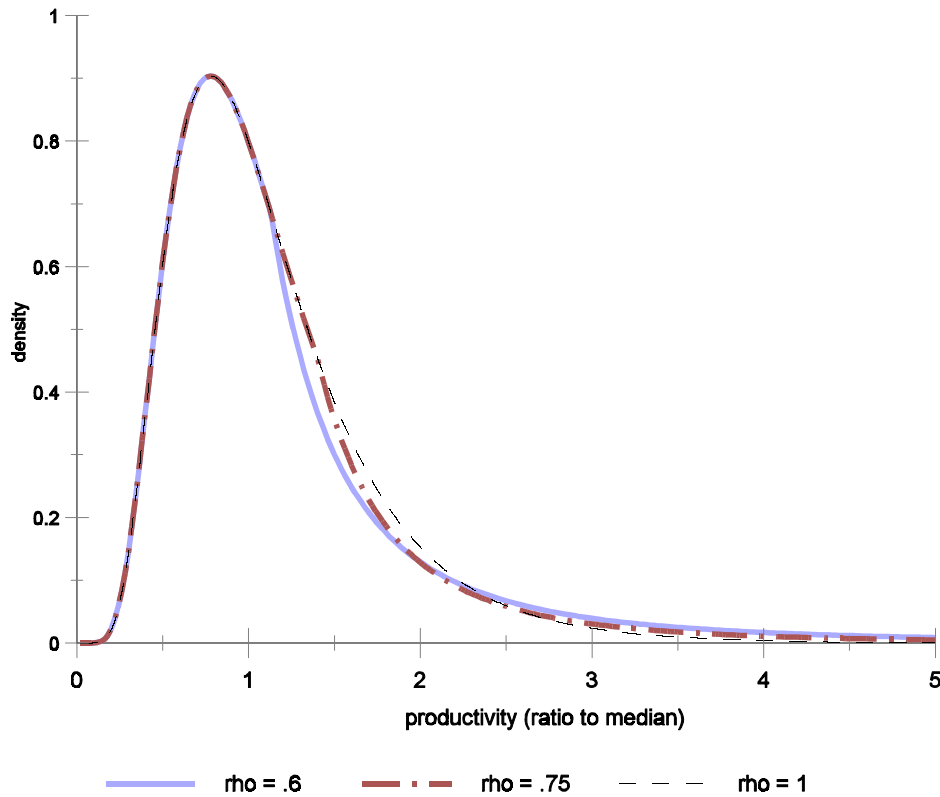


Figure 5 Three Density Functions ($\sigma = 0.5$), including lognormal special case

Notice that F 's skewness decreases with ρ , because the Pareto distribution is more skewed than the lognormal. The mean (and therefore the ratio of mean to median) of F is:

$$E(x; \sigma, \rho) = \int_0^{\infty} x dF(x; \sigma, \rho) = e^{\sigma^2/2} \Phi(\Phi^{-1}(\rho) - \sigma) + \frac{1 - \rho}{1 - [(1-\rho)\sigma/\phi(\Phi^{-1}(\rho))]} e^{\sigma\Phi^{-1}(\rho)}$$

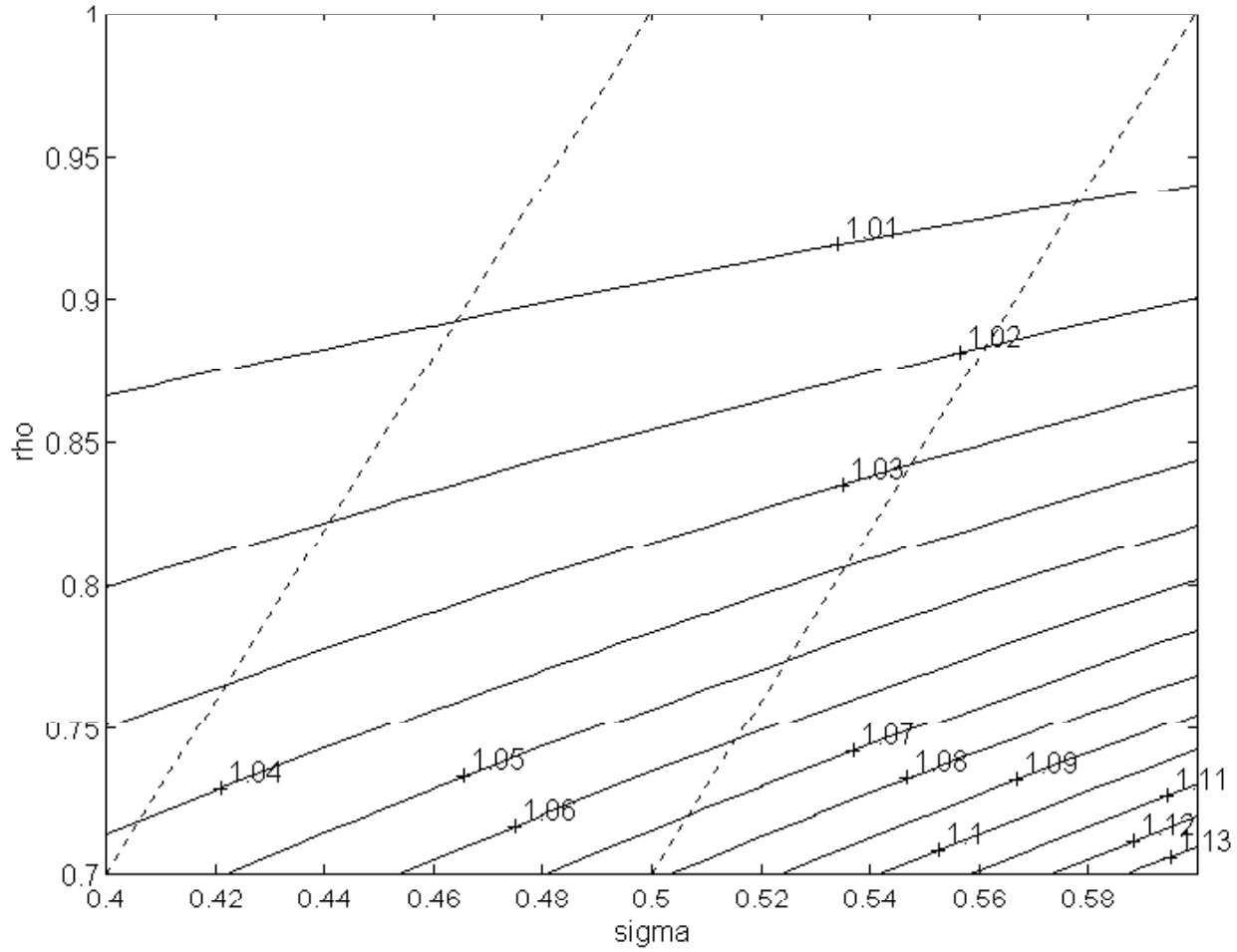
which, not surprisingly, is the weighted average of the lognormal and Pareto means with weights ρ and $(1-\rho)$, respectively. Since the median of F is 1 for all (σ, ρ) , $E(x)$ is also a indicator of F 's skewness.

Let $\text{sd}(\ln x)$ denote the standard deviation of $\ln x$ implied by F and consider the following calculation:

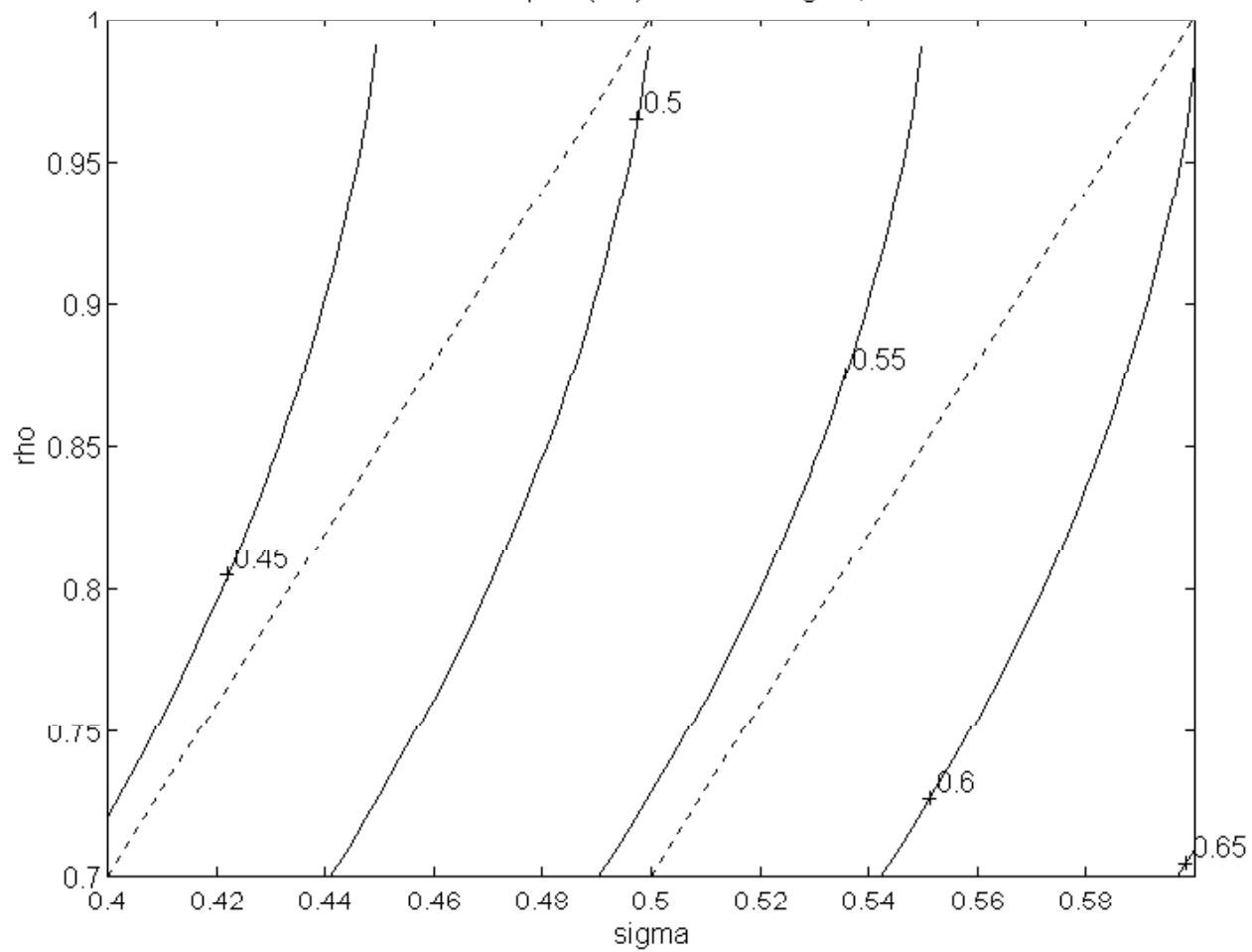
$$\lambda(x; \sigma, \rho) = \frac{E(x; \sigma, \rho)}{e^{[\text{sd}(\ln x)]^2/2}}$$

In words, λ is the ratio of the mean $E(x)$ to what we might guess to be the mean on the basis of lognormality and $\text{sd}(\ln x)$. This ratio is one if F were indeed the lognormal distribution (ie, if $\rho = 1$) and greater than one otherwise. Hence the ratio can also be interpreted as a metric of the skewness of F relative to the lognormal, as I do in the main text. The two figures on the next pages display the level curves of $\text{sd}(\ln x)$ and λ in the $[\sigma, \rho]$ plane. The parameter space Γ is bounded in both figures as dashed lines, showing how Γ requires $\text{sd}(\ln x) \in [0.47, 0.56]$ and $\lambda \in [1.00, 1.08]$. Interestingly, $\sigma = 0.5$ attains both the min and max values for λ .

contour map: skew relative to lognormal as funct of sigma, rho



contour map: sd(ln x) as funct of sigma, rho



VII. Appendix II: Nonunitary Elasticities of Substitution

@ available on request @

VIII. References

- Alesina, Alberto and Dani Rodrik. "Distributive Politics and Economic Growth." *Quarterly Journal of Economics*. 109(2), May 1994: 465-90.
- Barro, Robert J. and Chaipat Sahasakul. "Measuring the Average Marginal Tax Rate from the Individual Income Tax." *Journal of Business*. 56, October 1983: 419-52.
- Becker, Gary S. *A Treatise on The Family*. Cambridge, MA: Harvard University Press, 1993.
- Becker, Gary S. "A Theory of Competition among Pressure Groups for Political Influence." *Quarterly Journal of Economics*. 98(3), August 1983: 371-400.
- Becker, Gary S. and Casey B. Mulligan. "Is Voting Rational or Instrumental?" Working paper, University of Chicago, October 1999.
- Becker, Gary S. and Casey B. Mulligan. "Deadweight Costs and the Size of Government." NBER Working paper #6789, November 1998.
- Benabou, Roland. "Inequality and Growth." *NBER Macroeconomics Annual 1996*. Cambridge, MA: M.I.T. Press, 1996a.
- Brennan, H. Geoffrey and James M. Buchanan. *The Power to Tax: Analytical Foundations of a Fiscal Constitution*. Cambridge, MA: Cambridge University Press, 1980.
- Brennan, Geoffrey and Loren Lomasky. "Institutional Aspects of Merit Good Analysis." *Finanzarchiv*. 41(2), 1983: 183-206.
- Browning, Edgar K. "Why the Social Insurance Budget Is Too Large in a Democracy." *Economic Inquiry*. 13(3), Sept. 1975: 373-88.
- Buchanan, James M. and Roger Congleton. "Proportional and Progressive Income Taxation with Utility-Maximizing Governments" *Public Choice*. 34(2), 1979: 217-30.
- Feldstein, Martin. "Tax Avoidance and the Deadweight Loss of the Income Tax." NBER working paper #5055, March 1995.
- Fudenberg, Drew and Jean Tirole. *Game Theory*. Cambridge, MA: MIT Press, 1991.
- Galasso, Vincenzo and J.I. Conde Ruiz. "Positive Arithmetic of the Welfare State." CEPR

- Discussion Paper no. 2202, August 1999.
- Hochman, Harold M., and James D. Rogers. "Pareto Optimal Redistribution." *American Economic Review* . 59(4), Part 1, September 1969: 542-57.
- Krusell, Per, Vincenzo Quadrini, and Jose-Victor Rios-Rull. "Are Consumption Taxes Really Better Than Income Taxes." *Journal of Monetary Economics*. 37(3), June 1996: 475-503.
- Krusell, Per and Jose-Victor Rios-Rull. "On the Size of U.S. Government: Political Economy in the Neoclassical Growth Model." *American Economic Review*. 89(5), December 1999: 1156-81.
- Lindert, Peter. "What Limits Social Spending?" *Explorations in Economic History*. 33, 1996: 1-34.
- McClellan, Mark and Jonathan Skinner. "The Incidence of Medicare." NBER Working paper #6013, April 1997.
- Meltzer, Allan H. and Scott F. Richard. "A Rational Theory of the Size of Government." *Journal of Political Economy*. 89(5), October 1981: 914-27.
- Meltzer, Allan H. and Scott F. Richard. "Tests of a Rational Theory of the Size of Government." *Public Choice*. 41(3), 1983: 403-18.
- Mirrlees, James A. "An Exploration in the Theory of Optimum Income Taxation." *Review of Economic Studies*. 38(114), April 1971: 175-208.
- Mulligan, Casey B. "Galton vs the Human Capital Approach to Inheritance." *Journal of Political Economy*. 107(6), December 1999: S184-224.
- Mulligan, Casey B. *Parental Priorities and Economic Inequality*. Chicago: University of Chicago Press, 1997.
- Mulligan, Casey B. and Tomas J. Philipson. "Paternalism and Government Intervention: Public Finance in Reverse." manuscript, University of Chicago, November 2000.
- Musgrave, R.A., K.E. Case, and H. Leonard. "The Distribution of Fiscal Burdens and Benefits." *Public Finance Quarterly*. 2, 1974: 269-311.
- Nelson P. "Redistribution and the Income of the Median Voter." *Public Choice*. 98(1-2), January 1999: 187-94.
- Olsen, Edgar O. "Pareto-Desirable Redistribution in Kind: Comment." *American Economic Review*. 70(5), Dec. 1980: 1028-31.
- OECD. *Social Expenditure Database*. 1997.

-
- Pechman, Joseph A. *Who Paid the Taxes 1965-85*. Washington, D.C.: Bookings Institution, 1985.
- Pechman, Joseph A. and Benjamin A. Okner. *Who Bears the Tax Burden?* Washington, D.C.: Bookings Institution, 1974.
- Peltzman, Sam. "The Growth of Government." *Journal of Law and Economics*. 23(2), Oct. 1980: 209-87.
- Perotti, Roberto. "Growth, Income Distribution and Democracy: What the Data Say." *Journal of Economic Growth*. 1(2), June 1996: 149-87.
- Persson, Torsten and Guido Tabellini. "Is Inequality Harmful for Growth?" *American Economic Review*. 84(3), June 1994: 600-21.
- Reynolds, Morgan O. and Eugene Smolensky. *Public expenditures, taxes, and the distribution of income : the U.S., 1950, 1961, 1970*. New York : Academic Press, 1977.
- Roberts, Russell. "A positive analysis of the design of government transfer programs." PhD Dissertation, Univ of Chicago, December 1981. [discusses median voter's tax incidence?]
- Ruggles, P. and M. O'Higgins. "The Distribution of Public Expenditures and Taxes Among Households in the United Kingdom." *Review of Income and Wealth*. 1981: 298-326.
- Saint-Paul, Gilles and Thierry Verdier. "Inequality, Redistribution, and Growth: A Challenge to the Conventional Political Economy." *European Economic Review*. 40, 1996: 719-28.
- Slemrod, Joel. "Methodological Issues in Measuring and Interpreting Taxable Income Elasticities." *National Tax Journal*. 51(4), December 1998: 773-88.
- Stephenson, E. Frank. "Average Marginal Tax Rates Revisited." *Journal of Monetary Economics*. 41, 1998: 389-409.
- Stigler, George J. *The Irregularities of Regulation*. Midlothian, Scotland: The David Hume Institute, 1986.
- Tabellini, Guido. "A Positive Theory of Social Security." mimeo IGIER, Milano, Italy, February 1992.
- de Tocqueville, Alexis. "Government of the Democracy in America." *Democracy in America*. Book 1, Chapter 13, reprinted by the University of Virginia, <http://xroads.virginia.edu/~HYPER/DETOC/>, 1835.
- United States Bureau of the Census. "Historical Income Tables, (Table) F-11. Age of Householder - (All Races) by Median and Mean Income: 1947 to 1999." published 13

December 2000; <<http://ftp.census.gov/hhes/income/histinc/f11.html>>

United States Bureau of Economic Analysis. *National Income and Product Accounts.*

<http://www.bea.doc.gov/bea/dn/nipaweb/>

United States Council of Economic Advisers. *Economic Report of the President 1995.* Washington, DC: Government Printing Office, February 1995.

United States Office of Management and Budget. *Historical Tables, Budget of the United States Government, Fiscal Year 2001.* Washington, DC: Government Printing Office, 2001.

United States Social Security Administration. *Annual Statistical Supplement, 2000.* Washington, DC: Government Printing Office, 2000.

United States Social Security Administration. *Social Security Programs Throughout the World - 1995.* Washington, DC: Government Printing Office, 1995.