Working Papers
Institute of Mathematical Economics

Arbeiten aus dem
Institut für Mathematische Wirtschaftsforschung

Nr. 102
Wolfram F. Richter and J. Felix Hampe

On Income Tax Functions

Dezember 1980

Institut für Mathematische Wirtschaftsforschung
an der
Universität Bielefeld
Adresse/Address:
Universitätsstraße
4800 Bielefeld 1
Bundesrepublik Deutschland
Federal Republic of Germany
ON INCOME TAX FUNCTIONS

WOLFRAM F. RICHTER AND J. FELIX HAMPE

ABSTRACT:
The paper sheds some light on distributive properties of income tax functions. For this purpose the ARROW-PRATT notion of risk (or inequality) aversion is adapted to income taxation. JAKOBSSON's (1976) theorems concerning the residual progression of tax schedules are generalized. The question is then raised how progressive income tax must be to offset the dispersing effect of saving on the distribution of residual income. The answer will relate the elasticity of saving, the residual progression and the proportional inequality aversion. The paper is completed by a short discussion of historically effective German income tax formulas.

Adress:
Institute of Math. Economics
Postfach 8640
Universität
4800 BIELEFELD 1
F.R.G.
1. Introduction

This paper analyses distributive properties of income tax functions. We only consider tax schedules that are given in functional form $T(\cdot)$ such that $T(Y)$ denotes tax liability and $R(Y) = Y - T(Y)$ residual income for arbitrary income $Y \geq 0$.

The discussion will centre around three magnitudes, namely the residual progression (of $T$ at $Y$)
$$
\rho(Y,T) := \frac{YR'(Y)}{R(Y)}
$$
the proportional inequality aversion (of $T$ at $Y$)
$$
\alpha(Y,T) := \frac{-YR''(Y)}{R'(Y)} = \frac{YT''(Y)}{(1-T'(Y))}
$$
and
$$
\pi(Y,T) := \frac{d}{dy}(\frac{R(Y)}{R'(Y)}) = 1 + \frac{\alpha(Y,T)}{\rho(Y,T)}.
$$
(Primes indicate derivates.) It is the purpose of this paper to make explicit the specific type of distributional information carried by $\rho(\cdot,T)$, $\alpha(\cdot,T)$ and $\pi(\cdot,T)$. Thereby we are able to heavily rely on earlier work - in particular JAKOBSSON (1976) in the case of $\rho$ and PRATT (1964) in the case of $\alpha$.

We shall restate some of their results for the sake of a unified presentation.

Section 3 deals with $\alpha(\cdot,T)$. It seems as if tax schedules have never before been subject to a discussion in terms of
their inequality aversion. We shall argue that $\alpha$ measures the "degree of tax progression" as seen by the individual taxpayer. This follows by mere reinterpretation of the work of PRATT (1964) and ATKINSON (1970). $\alpha$ determines the distribution of residual income as drawn by the taxpaying unit, say, through time.

On the other hand $\rho$ can be said to measure the "degree of tax progression" as seen by society. $\rho$ determines the (personal) distribution of residual income across society. This observation is primarily due to JAKOBSSON (1976; see also KAKAWANI, 1976). However, JAKOBSSON's statements lack the (formal) precision which might be desirable and which is considered to be necessary for our purposes. Hence in section 4 we shall reproduce JAKOBSSON's main results with some formal corrections and slight extensions.

Section 5 collects arguments for EDGEWORTH tax functions that are characterized by a constant residual progression.

Section 6 introduces saving. We there deal with the question "how progressive taxation must be" to maintain the achieved level of personal income inequality if we allow for a rising propensity to save out of residual income. The answer will involve the elasticity of saving and function $\pi$ — thus
indirectly both measures of tax progression $\rho$ and $\alpha$. To avoid potential confusion let us stress that this approach to the determination of some "minimal degree of progression" has little to do with the optimal tax approach in the MIRRELES (1971) tradition which focuses on work incentives.

The proof of the main theorem 11 is given in section 7. Section 8 is devoted to a short discussion of historically effective German tax formulas in terms of $\rho, \alpha$ and $\pi$.

2. Edgeworth Tax Functions

Throughout this paper we assume that $T$ is sufficiently times continuously differentiable to justify the proofs. (Three times will be sufficient but not necessary.) Furthermore $T(0)=0$, $T'(Y)<1$ or equivalently $R(0)=0$ and $R'(Y)>0$ for all $Y\geq 0$.

Theorem 1: Let $a, b$ be constants, $a\neq 0$. The following conditions are equivalent:

a) $R(Y) = bY^a$ for all $Y > 0$;
b) $\rho(Y,T) = a$ for all $Y > 0$;
c) $\alpha(Y,T) = 1-a$ for all $Y > 0$;
d) $\pi(Y,T) = 1/a$ for all $Y > 0$. 
The proof is trivial. By this theorem the case of 
\( R(Y) = bY^a \) clearly deserves special interest. Let us call functions \( T, \) where \( T(Y) = Y - bY^a, \) EDGEWORTH tax functions as he will have been the first to discuss them explicitly (1919, p. 249).

3. **INEQUALITY AVERTION**

Let \( N = \{1, \ldots, n\}, n \geq 2 \) stand for some household whose members \( i \in N \) are all liable to pay tax from their income \( y_i. \) The income profile is \( y = (y_1, \ldots, y_n) \geq 0, \neq 0 \) and the income per capita \( \bar{y} = \Sigma y_i/n. \) The interpretation of \( N \) as some household is not the only one which makes good sense. One could equally think of \( y \) as a flow of income, where \( y_i \) is the income of some fixed individual at time \( i=1,\ldots,n. \) Probabilistic interpretations also present themselves.

We assume that tax laws generally prescribe to tax \( y_1, \ldots, y_n \) separately, one by one. Residual income then amounts to \( \Sigma R(y_i)/n \) on the average. There might be legal provisions that allow to average out incomes before tax. In that case residual income per capita amounts to \( R(\bar{y}). \) German tax laws, e.g., provide such a "privilege of splitting" for married couples. The marriage certificate gives the right to increase residual income per capita from \( R(y^c) := \Sigma R(y_i)/n \) to \( R(\bar{y}) = R(\Sigma y_i/n). \) \( y^c \) is the income before tax which, if
equally distributed among \( N \) yields the same residual income *per capita* as \( y \) without privilege of splitting. If, on the other hand, \( y \) stands for a flow of dated incomes \( R(\bar{y}) - R(y^0) \) measures the loss of residual income due to fluctuations through time. Hence it does not matter whether any legal provisions of averaging out \( y_1, \ldots, y_n \) do in fact exist. More favourable economic circumstances might as well render possible a steady flow of income. Yet, let us stick to the idea of household, and of the privilege of splitting. Thus let us call

\[
AGS(y,T) := \tilde{y} - y^0 \quad \text{the absolute gain from splitting and}
\]

\[
RGS(y,T) := (\tilde{y} - y^0)/\tilde{y} \quad \text{the relative gain from splitting.}
\]

These definitions have their counter-parts in PRATT (1964) and ATKINSON (1970). Confer these publications for additional interpretative comments.

Fix \( \tilde{y} > 0 \) and consider only such \( y \in \mathbb{R}^n \) holding

\[
\sum y_i/n = \tilde{y}.
\]

Let \( \sigma^2(y) = \frac{1}{n} \sum (y_i - y)^2 \) denote the variance of income. \( \sigma(\sigma^2(y)) \) is a term which tends to zero for \( \sigma^2(y) \to 0 \).

**Theorem 2 (PRATT, 1964):**

\[
AGS(y,T) = -\frac{1}{2} \sigma^2(y) R''(\bar{y})/R'(\bar{y}) + \sigma(\sigma^2(y)) \quad (1)
\]

\[
RGS(y,T) = \frac{1}{2} \sigma^2(\bar{y}^{-1} y) \alpha(\bar{y}, T) + \sigma(\sigma^2(\bar{y}^{-1} y)) \quad (2)
\]
Thus AGS(y,T) - and RGS mutatis mutandis - falls into two factors if the variance of income is "sufficiently small". The first factor, \( \frac{1}{2}\sigma^2(y) \), is independent of \( T(\cdot) \). The second factor, namely \(-R''(\bar{y})/R'(\bar{y})\), is only depending on the average income \( \bar{y} \) and not on \( y \) itself.

The degree of approximation is probably best illustrated by an example. The computation of RGS is based on the polynomial tax formula of §32a, German income tax, effective 1979/80: \( n=2 \), \( y_1 = 21600 \) and \( y_2 = 26400 \) yield RSP = 2.1709 E-3.

Computing the right-hand side of (2) yields
\[
\frac{1}{2} \times 0.433901 \times \frac{1}{100} = 2.1695 \text{ E-3}.
\]
The relative error is less than 1%.

It is not being suggested to use (1) and (2) for computing AGS and RGS. Both values can more directly be computed for arbitrary \( \sigma^2 \) by recurring to the official tax schedules which list tax liabilities \( T(Y_i) \) for increasing equidistant \( Y_i \geq 0 \). The error incurred by rounding off is in fact negligible for all practical purposes.

Theorem 2 becomes relevant when the personal tax load of two alternative tax functions \( T_1 \) and \( T_2 \) stands for comparison. (2) tells us that \( a(\cdot,T) \) is a cardinal measure of RGS(\( \cdot,T \)) - at least when income variances are small. This measure is still an ordinal one if \( \sigma^2 \) is allowed to take arbitrary values:
Theorem 3 (PRATT, 1964):

The following two conditions are equivalent in either the weak form (in which case the bracketed material has to be deleted) or in the strong form (indicated in brackets).

a) \( \text{RGS}(y, T_1) \leq [\leq] \text{RGS}(y, T_2) \)

for all \( y \in \mathbb{R}_+^n \) [that are non-equally distributed, i.e. \( y \in \mathbb{R}(1, \ldots, 1) \)].

b) \( \alpha(Y, T_1) \leq [\leq] \alpha(Y, T_2) \)

for all \( Y \) out of a dense\(^1\) subset of \( \mathbb{R}_+^n \).

The theorem remains true if, mutatis mutandis, attention is restricted to some fixed income interval. Furthermore in condition a) RGS may be replaced by AGS.

For a linear tax function \( T_1 \), i.e. \( T_1(Y) = a + bY \), we obtain a nice corollary. Then \( \text{AGS}(y, T_1) = 0 \). Note furthermore that

\[
T''(Y) > 0 \quad \text{iff} \quad -R''(Y)/R'(Y) > 0
\]

as \( R' > 0 \) by assumption.

Theorem 4: The assertion

\( 0 < \text{AGS}(y, T) \) for all \( y \in \mathbb{R}_+^n \), \( \in \mathbb{R}(1, \ldots, 1) \)

is equivalent to

\( 0 < T''(Y) \) on a dense subset of \( \mathbb{R}_+^n \).\(^2\)
The literature does not unanimously agree on the "correct" definition of progressivity. Although the condition $\frac{d}{dY}(\frac{T(Y)}{Y}) > 0$ is generally accepted, $T''(Y) > 0$ is often offered as a possible alternative. See, for instance, SADKA (1976). SEIDL et al. (1970) even favour to call $T$ progressive at $Y$ if $\frac{d}{dY}(\frac{T(Y)}{Y}) > 0$ and $T''(Y) > 0$.

BÖS et al. (1977) take a similar viewpoint. The foregoing discussion strongly suggests to call $T$ (at $Y$)

- inequality averse \(^3\) if $T''(Y) > 0$ and
- progressive if $\frac{d}{dY}(\frac{T(Y)}{Y}) > 0$.

Developing $T$ at zero according to TAYLOR

$$0 = T(0) = T(Y) - YT'(Y) + \frac{1}{2}Y^2T''(Z)$$

(for some $Z \in [0,Y]$) yields the

**Remark:** If $T$ is inequality averse for all $Y > 0$ then $T$ is also progressive for all $Y > 0$. \(^4\)

The reversal is clearly false as linear tax functions show. A stronger counter-example is $T(Y) = Y(1-e^{-Y})$ which is progressive for all $Y \geq 0$ however, inequality preferring, i.e. $T''(Y) < 0$, for all $Y > 2$.

4. **Residual Progression**

In this section $y = (y_1, \ldots, y_n)$ denotes the income profile of the whole tax paying population. (In case that households are allowed to split in the sense of section 3 the members
appear in $y$ by their average income before tax.) Assume throughout that $y$ is increasingly ordered:

$$0 < y_1 \leq \ldots \leq y_n.$$ 

Let $x,z \in \mathbb{R}^n_+$ ; $x_1 \leq \ldots \leq x_n$ ; $z_1 \leq \ldots \leq z_n$. We say that $x$ is [strictly] LORENZ dominated by $z$ and write $x \preceq_L z$ $[x \preceq_L z]$, respectively if

$$\frac{\sum x_i}{\sum x_i} \leq [<] \frac{\sum z_i}{\sum z_i} \text{ for all } k=1,\ldots,n-1.$$ 

Sometimes we write $z \succeq_L x$ if $x \preceq_L z$ , also

$x =_L z$ if both $x \preceq_L z$ and $x \succeq_L z$.

In theorems 5 and 6 $x(y), x^j(y)$ denote vectors of residual incomes. E.g.:

$$x^j_i(y) := y_i - T_j(y_i) \quad (i \in \mathbb{N}, j=1,2).$$

Theorem 5 (JAKOBSSON, 1976):

The following two conditions are equivalent in weak or strong form:

a) $x^2(y) \preceq_L x^1(y)$ $[x^2(y) <_L x^1(y)$ , respectively$]$ for all $y$ $[y \in \mathbb{R}(1,\ldots,1)]$ ;

b) $\rho(Y,T_2) > [>] \rho(Y,T_1)$ for all $Y$ out of a dense subset of $\mathbb{R}^n_+$. 

The theorem strengthens JAKOBSSON's proposition 1. Careful
inspection of his proof yields our assertion.

Let us choose \( T_2 \) to be proportional, i.e. \( T_2(Y) = bY \). Then \( x^2(y) = \mathcal{L} y \) and \( \rho(Y, T_2) = 1 \). Note furthermore that \( \rho(Y, T_1) < 1 \) iff \( T_1 \) is progressive at \( Y \).

Thus we obtain the following corollary (strong form):

**Theorem 6 (JAKOBSSON, 1976):**

\[ y \preceq \mathcal{L} x(y) \text{ for all } y \in \mathbb{R}(1, \ldots, 1) \text{ iff } T \text{ is progressive on a dense set in } \mathbb{R}_+ . \]

Note the structural parallel between theorems 3, 4 and 5, 6.

5. **ARGUMENTS FOR EDGEWORTH TAX FUNCTIONS**

\( \alpha \) and \( \rho \), respectively, induce quasi-orderings (reflexive and transitive) on the set of all tax functions:

\[ T_2 \preceq_{\alpha} T_1 \iff \alpha(\cdot, T_2) \geq \alpha(\cdot, T_1) , \]

\[ T_2 \preceq_{\rho} T_1 \iff \rho(\cdot, T_2) \leq \rho(\cdot, T_1) . \]

By theorems 3 and 5 these quasi-orderings allow some straightforward economic interpretation. They are both equal, \( \preceq_{\alpha} = \preceq_{\rho} \), if they are restricted to EDGEWORTH tax functions. This follows from theorem 1:
1 - a_2 = \alpha(\cdot, T_2) \geq \alpha(\cdot, T_1) = 1 - a_1 \iff \\
a_2 = \rho(\cdot, T_2) \leq \rho(\cdot, T_1) = a_1 ,

where \( R_i(Y) = b_i Y \). In this restricted case \( \gamma(\alpha) = \gamma(\rho) \) is even complete, thus defining an ordering.

The distributive effects of parameter variations are easy to assess in the case of EDGOWERTH tax functions. However, simplification cannot be an ultimate goal in normative economics.

The structural simplification of the EDGOWERTH situation could be regarded as a "loss of degrees of freedom" for policy making. Politicians may sometimes wish to disconnect variations of \( \alpha \) and \( \rho \) . The latest German tax reform may serve as an example ex post. (We are not claiming that the effect was planned as such.) Confer charts 2 and 3, below. Comparing T81 and T79 we may conclude that in the income range [28000, 42000] progressivity has been increased in terms of \( \alpha \) but decreased in terms of \( \rho \).

Unfortunately, little is known about the structural relationship of \( \gamma(\alpha) \) and \( \gamma(\rho) \) for non-EDGOWERTH tax functions. Let us consider EDGOWERTH tax functions in more detail. Which normative reasons could make one to choose them?
Constant proportional inequality aversion

An important implication of theorem 3 is:

Theorem 7 (Pratt, 1964):

\[ \alpha(Y,T) \text{ is decreasing [increasing or constant] in } Y \text{ if and only if for all } y \]

\[ \text{RGS}(\lambda y, T) \text{ is decreasing [increasing or constant] in } \lambda \in \mathbb{R}, > 0. \]

The requirement that \( \text{RGS}(\lambda y, T) \) should be independent of the income level measured by \( \lambda \) is appealing, though probably not compelling. Even if \( \alpha(\cdot, T) = \text{const} \) were normatively imperative small variations of the values taken by \( \alpha(Y,T) \) should not be overstated in practice. As has been argued before the uncompromising realization of \( \alpha(\cdot, T) = \text{const} \) could unduly interfere with competing political goals.

Constant residual progression

The following theorem parallels theorem 7. It is proved upon application of theorem 5 to \( T_1(Y) = T(Y) \) and \( T_2(Y) = \frac{1}{\lambda} T(\lambda Y) \) for arbitrary \( Y \) and \( \lambda \).

Theorem 8:

\[ \rho(Y,T) \text{ is decreasing [increasing or constant] in } Y \]

if and only if for all \( y \)
\( x(\lambda y) \) is increasing [decreasing or constant] in \( \lambda \) with respect to the LORENZ quasi-ordering \( \leq_L \).

The version concerning a constant residual progression was first proved by JAKOBSSON (1976). However, his proof does not rely on theorem 5. His argument makes use of CHAUCHY's functional equation. JAKOBSSON suggests to interpret \( \lambda \) as a measure of "inflation". An EDGEWORTH tax function is then called for if the redistributive effect of inflation is to be neutralized. Note that the same argument could formally be advanced in connection with theorem 7. However, it would be less convincing as the distributive effect there only refers to isolated households. The normative strength of a constancy requirement seems to be more convincing in the case of theorem 8.

**Sacrifice approach**

Nineteenth century's economists mainly tried to justify progressive taxation by inflicting an equal sacrifice on all taxpayers. An equal absolute, proportional or marginal sacrifice does not yield the desired result. There is, however, a non-classical sacrifice concept suggested by RICHTER (1980) which implies "moderate progression" for all utility functions \( U: \mathbb{R}_+ \rightarrow \mathbb{R} \) holding \( U(0)=0 \), \( U > 0 \), \( U'' < 0 \). This sacrifice concept is given by

\[
\frac{U(Y)U'(Y-T(Y))}{[U(Y-T(Y))]^2} = \text{const.} \quad (3)
\]

A normative, game-theoretic justification for condition (3) was presented by RICHTER and SELTEN (1980). For any fixed \( U \) (3)
implicitly defines some tax function $T = T_U$ on some open interval $\sigma = \sigma_U$ in $\mathbb{R}_+$. Irrespective of the considered $U$ we obtain

**Theorem 9 (RICHTER, 1980):**

$$1 > T'(Y) > T(Y)/Y$$

for all $Y \in \sigma$,

where the latter inequality is subject to the condition $0 < T(Y)/Y \leq 1/2$.

Inserting iso-elastic utility functions $U(Y) = Y^{\varepsilon}/\varepsilon$ into (3) yields the EDEGWORTH tax function

$$R_\varepsilon(Y) = Y - T_\varepsilon(Y) = bY^{1+\varepsilon}.$$

The range of the residual progression $\varepsilon/(1+\varepsilon)$ deserves special notice. The axiomatization of (3) makes use of $U'' \leq 0$ and $U(0) = 0$. Hence $\varepsilon$ is restricted to the interval $(0,1]$ and consequently $\varepsilon/(1+\varepsilon)$ to $(0,\frac{1}{2}]$. We thus only obtain a sacrifice theoretic justification of EDGEWORTH tax function if the residual progression belongs to the interval $(0,.5]$. The upper boundary, .5, is particularly noteworthy as it remains below empirical suggestions.

By using data from the American Federal Income Tax of 1917 EDGEWORTH derived concrete proposals as to the choice of $a$ and $b$ in $R(Y) = bY^a$. Thus he suggested .967, .952, and
.946 for \( a \) (p.250-252).

GENSER presents seven alternative specifications of \( R(Y) = bY^a \). Six of them are obtained by fitting best - according to various criteria - to the Austrian Income Tax of 1975. In all these cases \( a \) exceeds .828. There is a single proposal which comes .5 a bit nearer by suggesting \( a = .6554 \). However, that case does not pretend to fit the effective Austrian Income Tax. Whereas the effective average tax rate is 14.2% for some yearly income \( \bar{Y} \) of 50000 S the parameter value \( a = .6554 \) is obtained by requiring \( T(\bar{Y})/\bar{Y} = .05 \), i.e., by deliberate deviation of the status-quo.

Finally, R.PAULY (1979; see also PAULY et al., 1979) worked out several tax proposals for the F.R.G. referring to 1979. For certain income intervals in the range from 20000 DM to 50000 DM yearly income he bases his proposals on \( R(Y) = bY^a \) with \( a > .75 \).

All these proposals are well substantiated by empirical facts (Confer also chart 2. The residual progression \( \rho(Y,T) \) of various depicted German tax formulas never reached the range (0,.5].) They are, however, not supported by the normative approach (3).
The dichotomy between policy and theory has some noteworthy parallel in the theory of the measurement of income inequality.

Let \( y \) stand for the vector of personal incomes. ATKINSON (1970) suggested to measure income inequality by means of

\[
I(y,A) := 1 - \frac{nA^{-1}(\frac{1}{n}A(y_i))}{\sum_i y_i}.
\]

ATKINSON does not say much about the interpretation of the function \( A \). (See also the critique by SEN, 1973.) If we chose \( A \) to be the residual income function \( R \) \( I(y,R) \) would formally equate \( RGS(y,T) \) serving here as a measure of residual income inequality in the sense of ATKINSON. If \( T \) were the tax schedule effectuated by some planner or political party then \( a(\cdot,T) \) would reveal their inequality aversion and hence their ideas about distributive justice. According to ATKINSON there are good reasons to choose iso-elastic functions

\[
A_a(y) = y^{a/a} \quad (a \in \mathbb{R})
\]

for \( A \). (An axiomatization of the corresponding class of inequality measures \( \{I(\cdot,A_a) \mid a \leq 1\} \) was given by KOLM, 1975, and BORK et al., 1978.)
The interesting point is that whenever ATKINSON applies his measure \( I(\cdot, A_a) \) to make international comparisons of income inequality he selects values of \( a \) which never exceed 1/2. (E.g. \( a = 1/2, 0, -1/2 \), etc. Cf. p. 259, 1970, or p. 48, 1975.) These values of \( a \) could not be justified if \( A_a \) were chosen to approximate the empirical residual income functions discussed above.

6. **The Effect of Saving on the Distribution of Residual Income**

Theorems 3 and 5 tell us the distributive effects of reducing the residual progression or increasing the inequality aversion. They tell us nothing about the "optimal degree of progression". We shall now turn to this question.

The sacrifice approach of section 5 gave some hint however vague and controversial with view to tax policy. More clear-cut answers might be expected from the income theoretic approach to taxation. However, connecting ones arguments with the LORENZ curve may become a slippery slope. Accepting once that strict LORENZ domination is a definite welfare improvement - for which since the work of ROTHSCCHILD and STIGLITZ (1973) much can be said - there is no getting round that residual incomes should best be equalized. On the other hand it can seriously be doubted that many people are in fact prepared to bear such extremal
consequences. Equality will not be considered a prerequisite to true distributive equity.

In this section saving is introduced as an element of dynamic change. Without taxation a rising propensity to save tends to increase income inequality through time. We shall pose the question how progressive income tax must be to at least offset this distributive effect of saving. The answer will relate the elasticity of saving, the residual progression, and the proportional inequality aversion.

Let \( y \) - as before - denote the vector of personal incomes before tax - non-equally distributed and increasingly ordered. \( y_i \) comprises all taxable earnings, including interests on capital. We assume that there is an equilibrium in the sense that without saving the same income profile \( y \) would develop in all succeeding periods. By that we implicitly assume a constant population. \( x^0 = x(y) \) with \( x_i^0 = R(y_i) = y_i - T(y_i) \) is the residual income vector of period zero.

We then take regard to the fact that individuals save out of their residual income \( R \). (It is convenient to use the same symbol \( R \) for functions and values of these functions. No confusion should arise.) Let the saving function be given by
Cross section studies suggest

\[ \frac{d}{dR} \left( \frac{S(R)}{R} \right) > 0. \quad (4) \]

In the short run \( S(R) < 0 \) is possible, even realistic. From a long run point of view \( S(R) < 0 \) does not make much sense as zero income is subsistence level. On the other hand Duesenberry (1949) convincingly argues that in the long run saving does not only depend on \( R \) itself but also on the distribution of residual income.

As we would like to stick to the specification \( S(R) \) we just disregard individuals whose savings are negative. Hence we assume \( S(R) > 0 \) for all \( R \geq R(y_1) > 0 \). Then (4) is equivalent to

**Assumption:** \( RS'(R)/S(R) > 1 \) for all \( R \geq R(y_1) \). \( (4') \)

Individual \( i \) is saving \( S(R(y_i)) \) in period one. Hence his income before tax in period one is \( y_i + rS(R(y_i)) \) where \( r \) stands for the fixed market rate of interest. The residual income vector of period one is thus given by \( x^1 \) where

\[ x^1_i = R(y_i + rS(R(y_i))) = R(y_i + rS(x^0_i)). \]

We are going to compare \( x^1 \) and \( x^0 \) with respect to LORENZ domination.
The following two theorems are strict corollaries of theorem 12. They deal with special tax functions thus preparing for the general case.

**Theorem 10 (linear taxation):**

\[ c > 0, \ b > 0. \quad R(Y) = c + bY \text{ implies } x^0 \succeq_L x^1. \]

Hence linear taxation is not able to maintain the level of residual income inequality achieved in periode zero. Saving according to \((4')\) makes residual incomes disperse. Legal tax schedules generally become linear from some high income upwards. "The rich then tend to become even richer by the mere fact of saving".

**Theorem 11 (EDGEBWORTHian taxation):** Let \( R(Y) = bY^a \) and

\[
\frac{RS'(R)}{S(R)} \geq [\leq] \quad \frac{1}{a} \quad \text{for all } R \in [R(y_1), R(y_n)].
\]

Then \( x^1 \preceq_L x^0 \) [\( x^1 \succeq_L x^0 \), respectively].

(A technical remark: \( S(R) \leq R \) requires that \( RS'(R)/S(R) \) converges to 1 for \( R \to \infty \). The range of \( R \) thus has to be restricted if \( RS'(R)/S(R) \geq 1/a > 1 \) is to hold.)

According to theorem 10 it is sufficient to choose

\[ RS'(R)/S(R) \leq 1/a \quad \text{(for all } R) \]

if we intend to ensure \( x^1 \succeq_L x^0 \). However, this turns out to be a severe condition.
We learned above that the empirical studies of EDGEO\textsc{worth}, GENS\textsc{er}, and PAUL\textsc{y} suggest values for \( 1/a \) which (with one noted exception) all remain below \( 4/3 = 1.33 \). Later we shall take a closer look at the German saving function of the mid-seventies. The data give some evidence that the critical value \( 1.33 \) is exceeded by \( RS'(R)/S(R) \) at least for wide income ranges. The conclusion then is that - ignoring all other economic forces of change except saving - an EDGEO\textsc{worth} tax function with \( a = .75 \) is not sufficiently progressive to maintain the achieved level of income inequality.

We now turn to the general case. Theorem 11 holds true in four alternative versions indicated by the symbol \( \circ \) which has to be replaced by one of the relations \( \leq, \geq, \prec \) or \( \succ \). Accordingly \(\circ_L\) stands for \( \leq_L, \geq_L \), etc. We employ the shortened notation

\[
RS'(R)/S(R) \mid_Y = R(Y)S'(R(Y))/S(R(Y))
\]

here and at similar places.

Theorem 12 (General case):

For all \( Y \in [y_1, y_n] \) and all \( Z \in [Y, Y + rS(R(Y))] \)

\[
RS'(R)/S(R) \mid_Y \circ_{\circ L} \pi(Z, T).
\]  \hspace{1cm} (5)

Then \( x^0 \circ_{\circ L} x^1 \).

\( R(Y) = c + bY \) implies \( \pi(Y, T) = 1 \). Hence theorem 10 follows from
Theorem 12 by assumption \( (4') \). In case of \( R(Y) = bY^a \)
\( \pi(Y,T) = 1/a \). (See theorem 1.) That proves theorem 11. The
proof of theorem 12 is relegated to section 7.

Chart 4 depicts \( \pi(\cdot, T) \) for various tax formulas that with
the one exception of T136 have all been effective in Germany.

Let us now give a rough empirical estimate of \( RS'(R)/\bar{S}(R) \).
Data about savings are difficult to obtain. In what follows
we completely rely on the official survey on incomes
and consumption carried out for Germany in 1973. This survey
shows data about the empirical saving function denoted \( \tilde{S} \).
See table, below. The point-elasticity \( RS'(R)/\tilde{S}(R) \) will be
approximated by arc-elasticities. Writing
\[
R_{-1} \bigg| y_i := R(y_{i-1}), \quad R_{+1} \bigg| y_i := R(y_{i+1})
\]
we define
\[
\eta_-(R) := \frac{\tilde{S}(R) - \tilde{S}(R_{-1})}{R - R_{-1}} \cdot \frac{R_{-1}}{\tilde{S}(R)}
\]
\[
\eta_+(R) := \frac{\tilde{S}(R) - \tilde{S}_+(R)}{R_{+1} - R} \cdot \frac{R_{+1}}{\tilde{S}(R)}
\]

Note that if \( \bar{S} \) were convex we should obtain
\[
\eta_-(R) \leq R\tilde{S}'(R)/\bar{S}(R) \leq \eta_+(R).
\]
The data, however, show that \( \bar{S} \) is not convex \( (\bar{S}'' > 0) \) everywhere.
The elasticity of saving (F.R.G., 1973)

<table>
<thead>
<tr>
<th>Household net income; monthly; from - to (DM)</th>
<th>(1) mean (DM)</th>
<th>(2) mean savings (DM)</th>
<th>(3) ( \hat{S}(R) )</th>
<th>(4) ( \eta_-(R) )</th>
<th>(5) ( \eta_+(R) )</th>
<th>(4) + (5) / 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>below 600</td>
<td>510</td>
<td>9</td>
<td>-</td>
<td>7.524</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>600 - 800</td>
<td>751</td>
<td>41</td>
<td>2.432</td>
<td>2.714</td>
<td>2.573</td>
<td></td>
</tr>
<tr>
<td>800 - 1000</td>
<td>967</td>
<td>73</td>
<td>1.962</td>
<td>2.019</td>
<td>1.991</td>
<td></td>
</tr>
<tr>
<td>1000 - 1200</td>
<td>1177</td>
<td>105</td>
<td>1.708</td>
<td>1.997</td>
<td>1.853</td>
<td></td>
</tr>
<tr>
<td>1200 - 1500</td>
<td>1452</td>
<td>154</td>
<td>1.680</td>
<td>1.418</td>
<td>1.549</td>
<td></td>
</tr>
<tr>
<td>1500 - 1800</td>
<td>1791</td>
<td>205</td>
<td>1.314</td>
<td>2.033</td>
<td>1.674</td>
<td></td>
</tr>
<tr>
<td>1800 - 2500</td>
<td>2298</td>
<td>323</td>
<td>1.656</td>
<td>2.068</td>
<td>1.862</td>
<td></td>
</tr>
<tr>
<td>2500 - 5000</td>
<td>3430</td>
<td>652</td>
<td>1.529</td>
<td>2.816</td>
<td>2.173</td>
<td></td>
</tr>
<tr>
<td>5000 - 15000</td>
<td>6993</td>
<td>2559</td>
<td>1.463</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>


\( \hat{S}(R) \) will clearly depend on the grouping of incomes. For that reason the \( \hat{S}(R) \)-entries for the extremal incomes 510 and 6993 are particularly unreliable. The values \( \eta_-(510) = 7.524 \) and \( \eta_-(6993) = 1.463 \) should therefore best be ignored.

Despite of various legitimate objections the values above should not be all too far from the "true" elasticity of saving function \( R S'(R) / S(R) \). Let us whence relate the data to \( \pi(-,T) \) where we choose T75 for T. See chart 4. T75 is the tax function that was effective in Germany from 1975 to 1977. It was passed in 1974, i.e. only one year
after the savings data were collected. The tax function of the year 1973, namely T65, had been effective since 1965. T75 will best reflect the targets of income policy prevailing in 1973.

Note that
\[
\max_Y \pi(Y, T75) = \pi(29000, T75) = 1.391,
\]
a value which exceeds those of the table only once in the case of 1.314! Hence there are good reasons to infer that \( \theta > \) is the version of theorem 12 characterizing West-Germany during the mid-seventies. T75 was not sufficiently progressive to offset the dispersing effect of saving - at least for wide income ranges.

7. **Proof of Theorem 12**

Careful inspection and completion of a lemma proved by JAKOBSSON (1976) yields:

**Lemma:** Let \( 0 < x^0_1 < \ldots < x^0_n \), \( 0 < x^1_1 < \ldots < x^1_n \) and
\[
x^1_i/x^1_{i-1} \leq x^0_i/x^0_{i-1} \quad \text{for all } i=1, \ldots, n
\]
[and \( < \) for at least one \( i \)].

Then \( x^1 \gtrless_L x^0 \) [\( x^1 \gtrless_L x^0 \), respectively].
Proof of theorem 12: (5) \(\iff\)

\[
\frac{S'(R)R}{S(R)} \bigg|_Y \equiv \frac{d}{dY}(\frac{R}{R'}) \bigg|_Z
\]

\[= 1 + rS'(R)R' \bigg|_Y \equiv \frac{R'(Y)}{R(Y)} \left[ \frac{R(Y)}{R'(Y)} + rS(R(Y)) \frac{d}{dY}(\frac{R}{R'}) \bigg|_Z \right]
\]

\[= \frac{R'(Y)}{R(Y)} \frac{R}{R'} \bigg|_{Y+rS(R(Y))}
\]

where \(Z\) is appropriately chosen to justify the latter equality according to TAYLOR. Hence

\[
\frac{d}{dY} \log R(Y + rS(R(Y)))
\]

\[= \frac{R'}{R} \bigg[ 1 + rS'(R)R' \bigg] \bigg|_{Y+rS(R(Y))}
\]

\[\equiv \frac{R'(Y)}{R(Y)} = \frac{d}{dY} \log R(Y).
\]

If \(y_{j-1} < y_j\), integration of the extreme left-hand and right-hand terms yields

\[
\log R(y_j + rS(R(y_j))) \frac{R(y_j)}{R(y_{j-1} + rS(R(y_{j-1}))}
\]

\[
\log \frac{R(y_j)}{R(y_{j-1})}
\]

i.e. \(\frac{x_j}{y_j} \bigg|_{x_j=1} \equiv \frac{x_j}{y_j} \bigg|_{x_j=1} \). Here, equality would trivially hold in case of \(y_{j-1} = y_j\). Theorem 12 follows by the above lemma.
8. Charts

The following charts refer to historically effective tax functions of West-Germany - with one exception: T136 is a proposal for 1981 made by KARL-BRAUER-INSTITUT des Bundes der Steuerzahler e.V. (Tarifreform vorrangig, Wiesbaden, Feb. 1980). T65 became effective in 1965, T75 in 1975 and so forth. The list is not complete. There was some T78 which, however, differed little from T75. The computations are based on the mathematical formulas of §32a, German Income Tax Law. (Cf. Bundesgesetzblatt I: 1964, p.894; 1974, p.2195; 1979, p.757; 1980, p.1382.) The argument \( Y \) of \( T(\cdot) \) thus stands for taxable income. As such it may differ from income before tax by various personal allowances which here are all left out of account. Legal provisions concerning the rounding off of \( Y \) and \( T(Y) \) have also been neglected.

T65 is piecewise defined as a polynomial of degree 3 and less (in \( Y \)). T75, T79, T81 are piecewise representable by polynomials of degree 4 and less. T136 is not polynomial.

For \( Y \in [4200, 104200] \) it is of the form

\[
T(Y) = a_0(Z(Y))^{1.36} \text{ where } Z(Y) = (Y-4200)/1000.
\]

Charts 1 to 4 show \( T, \alpha(\cdot,T), \rho(\cdot,T) \) and \( \pi(\cdot,T) \).
The curves are computed for values of $Y$ that are multiples of 1000.

The tax functions are grouped into two sets: \{T65, T75, T81\} and \{T79, T136, T81\}. The first set is to illustrate long-run changes. The latter one is to shed some light on the latest amendment to tax laws.

Chart 3 tells us that progressivity - as measured by the residual progression - has been increased over time for middle and high incomes. There also is a recognizable horizontal shift which partially offsets the general increase of nominal incomes due to inflation and growth. T136 somewhat drops out of the general line. The progressivity burden is shifted to medium-high incomes.

Chart 2 seems to present a rather unsystematic picture. Such a first impression is, however, not correct. The "teeth" in the curves must be thought away. They are due to the habit of defining tax functions piecewise. The inequality aversion generally receives no attention by policy makers. Thus people will have made no effort "to smooth out" $\alpha(\cdot, T)$ as they will certainly have done with $\phi(\cdot, T)$. 
Let us ignore low incomes as there incomes may considerably differ from taxable incomes. Thus ignoring low incomes we observe a decreasing proportional inequality aversion for all tax functions except T136. If we interpret RGS as the relative gain from splitting in case of married couples a decreasing \( \alpha(\cdot, T) \) is at least more in line with the ability-to-pay principle than an increasing \( \alpha(\cdot, T) \). If, however, RGS is to measure the loss of residual income due to fluctuations through time the argument could be reversed. In any case the sharp kink of T136 at \( Y = 104200 \) will allow no convincing justification.

In fact, T136 tells us an interesting lesson to learn. Chart 1 could make one believe that there is no essential difference between T136 on one side and T79, T81 on the other. However, charts 2, 3, and 4 reveal severe distributive differences.

The similarity of charts 2 and 4 might be striking. The similarity was stressed by the choice of appropriate scales. The effect is yet merely accidental. It can easily be rationalized by the equality of

\[
\pi(\cdot, T) = 1 + \alpha(\cdot, T)/\rho(\cdot, T).
\]

For middle and high incomes \( \rho(Y, T) \) is almost constant in \( Y \) for all considered tax functions. For such incomes \( \pi(\cdot, T) \) approximately is an affine transformation of \( \alpha(\cdot, T) \). In the lower income range the fluctuations of \( \rho(\cdot, T) \) are swallowed by a vanishing \( \alpha(\cdot, T) \).
Again, T136 drops out of the general line in chart 4. A rising $\pi(\cdot, T)$ for growing incomes is not favourable with view to theorem 11. The elasticity of saving necessarily has to decrease to one - at least asymptotically. Hence, if we liked $\pi(\cdot, T)$ to dominate $RS'(R)/S(R)$ in the sense of theorem 11 a decreasing $\pi(\cdot, T)$ would seem to be advocated.

We considerably benefited from hints given by O.Börs, R.Pethig, and E.Schlicht. Thanks to all of them.

**FOOTNOTES**

1) $\sigma$ is dense in $\mathbb{R}_+$ if every $Y \in \mathbb{R}_+ \setminus \sigma$ is a limit point of $\sigma$.

2) $T$ is known to be strictly convex iff $T'' > 0$ on a dense subset. (Cf. KATZNER, p. 189.)

3) Pratt would add "strictly".

4) Relying on footnote 2) one could prove progressivity of $T$ for all $Y > 0$ by only assuming $T'' > 0$ on a dense subset of $\mathbb{R}_+$.

**REFERENCES**


Richter, W.F., 1980, Taxation according to ability to pay, Disc. Paper, IMW, University of Bielefeld.


Chart 1: Selected income tax functions
Chart 2: Proportional inequality aversion
Chart 3: Residual progression
Chart 4: $\pi(T) = \frac{d}{dT}(R/R')$
"WIRTSCHAFTSTHEORETISCHE ENTSCHEIDUNGSFORSCHUNG"

A series of books published by the Institute of Mathematical Economics, University of Bielefeld.

Wolfgang Rohde
Ein spieltheoretisches Modell eines Terminmarktes (A Game Theoretical Model of a Futures Market)
The model takes the form of a multistage game with imperfect information and strategic price formation by a specialist. The analysis throws light on theoretically difficult empirical phenomena.

Vol. 1  176 pages  price: DM 24,80

Klaus Binder
Oligopolistische Preisbildung und Markteintritte (Oligopolistic Pricing and Market Entry)
The book investigates subgame perfect equilibrium points of a three-stage game model of oligopoly with decisions on entry, on expenditures for market potential and on prices.

Vol. 2  132 pages  price: DM 22,80

Karin Wagner
Ein Modell der Preisbildung in der Zementindustrie (A Model of Pricing in the Cement Industry)
A location theory model is applied in order to explain observed prices and quantities in the cement industry of the Federal Republic of Germany.

Vol. 3  170 pages  price: DM 24,80

Rolf Stoecker
Experimentelle Untersuchung des Entscheidungsverhaltens im Bertrand-Oligopol (Experimental Investigation of Decision-Behavior in Bertrand-Oligopoly Games)
The book contains laboratory experiments on repeated supergames with two, three and five bargainers. Special emphasis is put on the end-effect behavior of experimental subjects and the influence of altruism on cooperation.

Vol. 4  197 pages  price: DM 28,80

Angela Klopsteck
Eingeschränkt rationale Marktprozesse (Market processes with Bounded Rationality)
The book investigates two stochastic market models with bounded rationality, one model describes an evolutionary competitive market and the other an adaptive oligopoly market with Markovian interaction.

Vol. 5  104 pages  price: DM 29,80

Orders should be sent to:
Pfeffersche Buchhandlung, Alter Markt 7, 4800 Bielefeld 1, West Germany.