On the Optimal Marginal Rate of Income Tax

Gareth D. Myles
Institute for Fiscal Studies and University of Exeter
June 1999

Abstract: The paper shows that in the quasi-linear model of income taxation, the optimal marginal rate of tax can be calculated without needing to specify the utility of consumption. This result is used to investigate the qualitative behavior of the marginal rate. It is shown that every possible qualitative outcome may be achieved by appropriate selection of the skill distribution. Therefore the model does not place any \textit{a priori} restrictions on the behavior of the marginal rate and the constancy of findings in previous simulation analysis is a consequence of their restricted structures.

Keywords: Income taxation, optimal, marginal rate

JEL Classification: 022, 323

Acknowledgements: Thanks are due to Frank Page, John Weymark and an anonymous referee.

Correspondence: Department of Economics, University of Exeter, Exeter, EX4 4RJ, UK, email G.D.Myles@ex.ac.uk, phone 01392 264487, fax 01392 263242.
1. Introduction

The analysis of nonlinear income taxation pioneered by Mirrlees (1971) has characterized a number of properties that the optimal tax must possess (see e.g. Myles (1995) for a survey of these). The theoretical results though do not answer all the questions that are raised about income taxation. The most hotly-debated practical issue is the behavior of the marginal rate of tax, in particular whether the optimal income tax should be progressive - a property that the tax systems of all developed countries possess. The theory has so far not fully resolved this question. It is well-known that the marginal tax rate should be zero for the highest skill consumer, so the tax function cannot be progressive everywhere. But this end-point result provides no information on the behavior of the tax schedule on the interior of the skill distribution. This is a significant gap in our knowledge.

Given the relative paucity of theoretical results, numerical simulations have been used in order to provide insight into the overall structure of the tax function. These have combined a log-normal distribution of skill with either Cobb-Douglas (Mirrlees (1971)) or CES utility (Kanbur and Tuomala (1994)). These specifications generate what will be called here the “classical” optimal tax function: as a function of skill, the marginal rate of tax either first rises and then falls or is highest initially and then falls. Which behavior obtains depends on the degree of equity in social welfare and the standard deviation of the skill distribution (see Kanbur and Tuomala (1994)). These qualitative properties of the tax function have remained consistent throughout all simulation results that have been reported. Such consistency would be re-assuring, and suggestive that the income tax should always have these properties, were it not for the very narrow range of specifications that have been used to generate the results. For these results to have valuable policy implications it must be shown that they are robust to changes in specification.

The reason that the range of specifications is so limited can probably be found in the computational difficulties involved in solving the model. These are not insurmountable, but do suggest that an alternative approach would be better. Consequently, this paper adopts the approach of pursuing a computationally simpler model. An alternative approach is followed in Saez (1999).

---

1 For instance the practical relevance of the zero endpoint result has been assessed by considering its implications for consumers “close” to the top of the skill distribution.

2 Mirrless also concluded that the marginal tax rate was fairly constant but other simulations have since disproved this.

3 An alternative approach is followed in Saez (1999).
be found explicitly as function of consumers’ skills and is independent of the utility-of-consumption function. This makes it computationally simple to assess the effect of varying the skill distribution upon the optimal marginal rates.

What the analysis shows is the following result: any qualitative structure for the optimal tax function can be supported by some skill distribution. Expressed alternatively, except for the fact that the marginal rate cannot rise between the second to highest and highest skill consumers, there are no a priori restrictions on the qualitative properties of the optimal tax function. So the structure of the “classical” optimal tax function is just a consequence of the restricted set of simulation specifications and does not capture some deeper feature of optimal taxation. The model used here assumes utility is linear in labor supply. Diamond (1998) has already exploited a linear-in-consumption model to show tax rates may increase above the modal income for some skill distributions.

It should be stressed that although the results are developed here in the context of optimal income taxation, they are also applicable to nonlinear incentive schemes in general. The same properties of the quasi-linear model can easily be exploited in other contexts to provide similar insights.

2. Quasi-linearity and marginal tax rates
This section briefly introduces the model of income taxation with quasi-linear utility. Lollivier and Rochet (1983) applied this to a model with a continuum of consumers. The model with a finite numbers of consumers on which this paper is based is analyzed in detail in Weymark (1986a, b 1987).

The basis of the model is that utility is quasi-linear in labor supply so

\[ U = u(x_i^i) - \ell_i^i = \frac{z_i^i}{s_i^i}, \]  

(1)

where \( x_i \) is consumption of consumer \( i \), \( z_i \) is pre-tax income and \( s_i \) is the level of skill. The marginal rate of substitution for \( i (MRS_i) \) is equal to \( 1/u's_i \), so that it is consistent with the requirements of agent monotonicity. With a weighted utilitarian welfare function and a tax policy that is purely redistributive, the choice of an optimal tax function is equivalent to the government choosing an allocation \{\( x_i, z_i \}\} for each consumer \( i = 1, \ldots, N \) to solve the following program

\[ 4 \] This assumption is not necessary.
Program 1: \[
\max \left\{ \sum_{i=1}^{N} \mu_i \left[ u(x^i) - \frac{z^i}{s^i} \right] \right\},
\]
subject to: (i) \[
\sum_{i=1}^{N} x^i = \sum_{i=1}^{N} z^i,
\]
(ii) \[
u(x^i) - \frac{z^i}{s^i} \geq u(x^{i'}) - \frac{z^{i'}}{s^{i'}}, \text{ all } i, i',
\]
where (i) is the budget constraint and (ii) the incentive compatibility constraints.

The general solution for the continuum version of this model is given in Lollivier and Rochet (1983). They provide a characterization of the optimal consumption function from which the tax rate could be inferred. However it is more direct to work from the results of Weymark (1986a, b) for the finite case. These show that Program 1 is equivalent to

Program 2: \[
\max \left\{ \sum_{i=1}^{N} \beta_i u(x^i) - \sum_{i=1}^{N} x^i \right\},
\]
where \(\beta_i := s^i + \left[ i - \sum_{n=1}^{i} \lambda_n \right] \left[ s^{i+1} - s^i \right], \lambda_i := \frac{\mu_i}{s^i}, \) and \(s^{N+1}\) is an arbitrary number. With no bunching, the solution to Program 2 is described by
\[
\beta_i u'(x^i) = 1, \quad i = 1, \ldots, N.
\]
(2)

As Weymark (1986b) noted, the tax function is kinked at the location of each consumer, so the marginal tax rate is not formally defined at these points. However, it is possible to take the gradient of the indifference curve as determining an implicit marginal rate of tax. Doing this, the marginal tax rate facing consumer \(i\) (\(MTR_i\)) is
\[
MTR_i := 1 - MRS_i = 1 - \frac{\beta_i}{s^i},
\]
(3)
using (2). Hence the marginal tax rate is defined by the skill distribution and is independent of the utility of consumption.

In undertaking the calculations reported in Section 3 the normalization of Weymark (1986a, b) is adopted so \(\sum_{i=1}^{N} \lambda_i = N\). For a weighted utilitarian social welfare function, the weights \(\mu_i\) can be selected to ensure that this does not impose any additional restriction. However, an unweighted utilitarian social welfare function is used in the calculations with \(\mu_i = 1\) for all \(i\). The joint effect of this choice and the
normalization is to place an additional restriction on the permissible distributions of ability. Relaxing this would simply make it easier to find ability distributions that generate the required patterns of tax rates. The reason for adopting the normalization is to permit direct use of the necessary and sufficient condition for no-bunching in Theorem 2 of Weymark (1986b). This is the requirement that $\beta_1 < \beta_2 < \ldots < \beta_n$ and $\beta_2 > 0$. The results reported in Table 1 are for skill distributions that satisfy this condition which justifies the use of the necessary condition in (2).

3. Numerical results

The numerical results reported in this section are constructed for a five-consumer economy. The number of consumers was chosen as the minimum necessary to exhibit a sufficiently interesting set of results. Before describing the results in detail, it is worth discussing what these are aiming to achieve.

The aim of the paper, as noted in the Introduction, is to investigate the qualitative behavior of the marginal rate of tax. In this respect, after ranking consumers by income, a progressive tax system would have a marginal rate which increased from one consumer to the next. More generally, starting with the lowest income consumer (who is also the lowest skill via incentive compatibility), qualitatively the marginal tax rate can be either higher or lower for the next skill level. The same is true in passing from the second skill level to the third. With five consumers and the fact that the marginal rate is zero at the top, eight possible qualitative patterns can arise. These are illustrated in Figures 1 and 2 where “i” denotes an increase in the tax rate between consumers and “d” denotes a decrease.

![Figure 1: Changes in Tax Rate](image-url)
Out of the eight possibilities, existing simulations have found the patterns 6, 7 and 8 - these are the “classical” optimal tax functions described in the Introduction. This would be a clear guide to policy if it was correct that these were the only qualitative forms of tax function that could arise. In fact, what is now shown is that patterns 1 - 5 can also arise. Consequently, since the model can generate every qualitative form of tax function, no \textit{a priori} restrictions can be placed on it whatsoever. Expressed alternatively, the fact that only the “classical” form has emerged in previous simulations is not a reflection of something deeper but just a consequence of the assumptions.

Table 1 reports the optimal marginal tax rates for five different distributions of ability. In each case, the skill level of consumer 5 is implied by the other four \textit{via} the normalization rule. It should be noted that the relaxation of the assumption of a utilitarian social welfare function could only serve to enhance the conclusions.

As already noted, these results show that any pattern of optimal tax rates that is theoretically possible can be achieved for some distribution of ability. The model does not place any restriction on the pattern that emerges except that, since the high ability faces a zero rate and the rate must be nonnegative, it cannot rise in going from the second highest to highest ability.
### Table 1: Optimal Marginal Tax Rates

<table>
<thead>
<tr>
<th></th>
<th>Ability</th>
<th>$s_1$</th>
<th>$s_2$</th>
<th>$s_3$</th>
<th>$s_4$</th>
<th>$s_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>i,i,i</td>
<td>.65</td>
<td>.75</td>
<td>1</td>
<td>1.4</td>
<td>2.42</td>
</tr>
<tr>
<td></td>
<td>Marginal tax rate (%)</td>
<td>.08</td>
<td>.29</td>
<td>.35</td>
<td>.425</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>d,i,i</td>
<td>.8</td>
<td>.82</td>
<td>.83</td>
<td>1.1</td>
<td>2.4</td>
</tr>
<tr>
<td></td>
<td>Marginal tax rate (%)</td>
<td>.0063</td>
<td>.0059</td>
<td>.219</td>
<td>.689</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>d,i,d</td>
<td>.9</td>
<td>.92</td>
<td>.93</td>
<td>1.1</td>
<td>1.22</td>
</tr>
<tr>
<td></td>
<td>Marginal tax rate (%)</td>
<td>.0025</td>
<td>.0021</td>
<td>.05</td>
<td>.02</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>i,d,i</td>
<td>.65</td>
<td>.75</td>
<td>1.05</td>
<td>1.2</td>
<td>2.92</td>
</tr>
<tr>
<td></td>
<td>Marginal tax rate (%)</td>
<td>.08</td>
<td>.35</td>
<td>.12</td>
<td>.94</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>d,d,i</td>
<td>.85</td>
<td>.91</td>
<td>.94</td>
<td>.95</td>
<td>1.64</td>
</tr>
<tr>
<td></td>
<td>Marginal tax rate (%)</td>
<td>.012</td>
<td>.009</td>
<td>.004</td>
<td>.27</td>
<td>0</td>
</tr>
</tbody>
</table>

4. Conclusions
The paper has considered at the qualitative properties of the optimal marginal rate of income taxation based on the observation that quasi-linearity of utility allows this to be determined without specifying the utility of consumption. Linked with the fact that the number of consumers is finite, this allows a break from the restrictive set of distributions and preferences adopted in previous simulations.

The results demonstrate that the model is capable of generating all qualitative patterns of marginal tax rates. The continual emergence of the “classical” tax function in previous studies can therefore be seen as just an artifact of their restrictive structure. The paper used a utilitarian social welfare function but a more general social welfare function would simply add another degree of freedom and make it easier to construct examples with the properties required.

In conclusion, the inference from previous simulations that the optimal tax function may belong to a narrow class has been shown to be invalid. The model does not restrict the qualitative structure in any way beyond that established in the existing theoretical results. It certainly does not provide any a priori restriction on the structure of marginal rates since these can behave in any way over the population.
The main message of the paper must be that it simulations are to provide any guide to policy, they must be based on real data - there are no general properties waiting to be discovered using artificial data.

**References**


