

UNIVERSITY *of York*



Discussion Papers in Economics

No. 15/21

**On Redistributive Taxation under the Threat of
High-Skill Emigration**

Alan Krause

Department of Economics and Related Studies
University of York
Heslington
York, YO10 5DD

On Redistributive Taxation under the Threat of High-Skill Emigration

Alan Krause*
University of York

19 October 2015

Abstract

The increasing international mobility of high-skill individuals is often seen as posing a threat to domestic social welfare, by limiting the ability of governments to tax these individuals and redistribute to the poor. In this note, we examine a simple dynamic nonlinear income tax model without commitment. In this setting, it is shown that the threat of emigration by high-skill individuals facilitates redistribution and increases social welfare in the short run, and has no effect on social welfare over the long run.

Keywords: nonlinear taxation; migration; commitment.

JEL Classifications: H21; H24; F22.

*Department of Economics and Related Studies, University of York, Heslington, York, YO10 5DD, U.K. E-mail: alan.krause@york.ac.uk.

1 Introduction

Optimal tax analyses typically assume that individuals cannot emigrate to avoid domestic taxation. Such an assumption is, however, increasingly viewed as being unrealistic, especially as it relates to high-skill individuals. When high-skill individuals are immobile, redistributive taxation must take into consideration that these individuals may change their labour supply along the intensive margin. That is, high-skill individuals may work less, thus reducing the amount of income available for redistribution. When high-skill individuals are internationally mobile, they have the additional option of changing their labour supply along the extensive margin, i.e., they may emigrate. An often-employed method to capture the threat of emigration is to introduce type-dependent participation constraints into the optimal tax problem; see, e.g., Osmundsen (1999), Krause (2009a), and Simula and Trannoy (2010, 2012). Naturally, these additional constraints reduce the level of social welfare attainable.

In this note, we introduce the threat of emigration by high-skill individuals into a dynamic nonlinear income tax model without commitment. As in the related literature, this threat is captured by introducing a participation constraint for high-skill individuals. However, we show that the introduction of the high-skill type's participation constraint actually facilitates redistribution and increases social welfare in the short run, and has no effect on social welfare over the long run. The intuition is as follows. In period 1 the government does not know each individual's skill type, and therefore implements standard (incentive-compatible) nonlinear income taxation which induces individuals to reveal their types. However, high-skill individuals know that if they reveal their type in period 1, they lose their information advantage and will be subjected to first-best taxation from period 2 onwards. This means that high-skill individuals must be offered a very attractive tax treatment in period 1 to reveal their type, to compensate them for having to face first-best taxation thereafter. Accordingly, the ability of the government to redistribute in period 1 is severely limited. However, the threat of emigration reduces the extent of redistribution possible under first-best taxation, meaning that high-skill individuals require less compensation in period 1 to reveal their type. This enables

the government to implement more redistribution in period 1, which correspondingly increases first-period social welfare. The threat of emigration limits redistribution and reduces social welfare from period 2 onwards, but optimal taxation balances the short-run benefits against the long-run costs. Thus the threat of emigration has no effect on the level of social welfare summed over all periods.

In terms of previous results, the paper most closely related to ours is that by Leite-Monteiro (1997). He also finds that increased international mobility may enhance redistribution, but his model is entirely different to ours. In Leite-Monteiro's model, the government can always implement first-best personalised lump-sum taxes, and changes in the country's skill composition *after* migration takes place is what makes enhanced redistribution a possibility. By contrast, in our model the participation constraint ensures that no one migrates, so the skill composition remains unchanged. Our note is also related to the literature on dynamic Mirrlees (1971) nonlinear income taxation without commitment, e.g., Roberts (1984), Apps and Rees (2006), Brett and Weymark (2008a), Krause (2009b), Berliant and Ledyard (2014), and Guo and Krause (2011, 2013, 2014, 2015a, 2015b). In dynamic Mirrlees models, the question arises as to whether the government can or cannot commit to not using skill-type information revealed by individuals in earlier periods when it implements taxation in latter periods. A common theme in the literature is the highlighting of how different and counter-intuitive optimal policy can be when the commitment assumption is relaxed. The present note provides another example of a result that, at first glance, appears quite counter-intuitive.

The remainder of the note is organised as follows. Section 2 outlines the model and the structure of optimal taxation. Section 3 presents and discusses our result, while Section 4 concludes. The proof of our result is contained in an appendix.

2 A Simple Model

We consider an infinite-horizon model with a unit measure of individuals and with the following timing. In period 1 the government knows there are $\phi \in (0, 1)$ high-skill individuals and $(1 - \phi)$ low-skill individuals in the economy, but it does not know

any individual's skill type. The government therefore implements standard second-best (incentive-compatible) nonlinear income taxation, under which each individual is willing to reveal their type. Then, from period 2 onwards, the government knows each individual's skill type, and is tempted to use this information to implement first-best redistributive taxation. However, high-skill individuals have the option of emigrating, so redistribution is limited by the high-skill type's participation constraint. For simplicity we assume that individuals do not save or borrow, so the only link between periods is the revelation and use of skill-type information.

Specifically, the government in period t (where $t \geq 2$) solves the following problem. Choose tax treatments $\langle c_L^t, y_L^t \rangle$ and $\langle c_H^t, y_H^t \rangle$ for the low-skill and high-skill individuals, respectively, to maximise:

$$(1 - \phi) \left\{ u(c_L^t) - \frac{y_L^t}{w_L} \right\} + \phi \left\{ u(c_H^t) - \frac{y_H^t}{w_H} \right\} \quad (2.1)$$

subject to:

$$(1 - \phi) [y_L^t - c_L^t] + \phi [y_H^t - c_H^t] \geq 0 \quad (2.2)$$

$$u(c_H^t) - \frac{y_H^t}{w_H} \geq V_H \quad (2.3)$$

where c_i^t is type i 's consumption (or post-tax income) in period t , $y_i^t = w_i l_i^t$ is type i 's pre-tax income in period t , with w_i denoting type i 's wage rate and l_i^t denoting type i 's labour supply in period t . It is assumed that $w_H > w_L > 0$ and that wages remain constant over time. Equation (2.1) is a utilitarian social welfare function, where $u(\cdot)$ is increasing and strictly concave and the individuals' utility function is quasi-linear in labour.¹ Equation (2.2) is the government's budget constraint,² and equation (2.3) is the high-skill type's participation constraint. High-skill individuals can obtain a utility level of V_H by emigrating, which is their reservation utility and is assumed to remain constant

¹The literature on the comparative statics of optimal nonlinear income taxes (e.g., Weymark (1987), Brett and Weymark (2008b, 2011), and Simula (2010)) has shown that results are generally obtainable only when the utility function is quasi-linear. As we make use of comparative statics methods, we also assume that the utility function is quasi-linear.

²We assume that the government cannot save or borrow, and that its revenue requirement is zero. Thus taxation is implemented only for redistributive purposes.

over time. The solution to programme (2.1) – (2.3) yields functions for the choice variables, in particular $c_L^t(\phi, w_L, w_H, V_H)$ and $y_L^t(\phi, w_L, w_H, V_H)$. Let $W^t(\phi, w_L, w_H, V_H)$ denote the value function associated with programme (2.1) – (2.3), which represents the level of social welfare attainable in period t .

In period 1 the government does not know each individual’s skill type, and therefore implements second-best (incentive-compatible) nonlinear income taxation. It chooses tax treatments $\langle c_L^1, y_L^1 \rangle$ and $\langle c_H^1, y_H^1 \rangle$ for the low-skill and high-skill individuals, respectively, to maximise:

$$(1 - \phi) \left\{ u(c_L^1) - \frac{y_L^1}{w_L} \right\} + \phi \left\{ u(c_H^1) - \frac{y_H^1}{w_H} \right\} \quad (2.4)$$

subject to:

$$(1 - \phi) [y_L^1 - c_L^1] + \phi [y_H^1 - c_H^1] \geq 0 \quad (2.5)$$

$$u(c_H^1) - \frac{y_H^1}{w_H} + \sum_{t=2}^{\infty} \delta^{t-1} V_H \geq u(c_L^1) - \frac{y_L^1}{w_H} + \sum_{t=2}^{\infty} \delta^{t-1} \left[u(c_L^t(\cdot)) - \frac{y_L^t(\cdot)}{w_H} \right] \quad (2.6)$$

where equation (2.4) is the first-period utilitarian social welfare function, and equation (2.5) is the government’s first-period budget constraint. Equation (2.6) is the high-skill type’s incentive-compatibility constraint, with $\delta \in (0, 1)$ denoting the individuals’ discount factor. If a high-skill individual chooses $\langle c_H^1, y_H^1 \rangle$ in period 1, they are revealing their type to the government, and will then receive the high-skill type’s reservation utility, V_H , in each period $t \geq 2$. Therefore, a high-skill individual will be willing to choose $\langle c_H^1, y_H^1 \rangle$ in period 1 if the utility they obtain from this tax treatment, plus the utility they then receive from period 2 onwards, is greater than or equal to the utility they could obtain by pretending to be a low-skill individual. That is, if a high-skill individual chooses $\langle c_L^1, y_L^1 \rangle$ in period 1, they are announcing to the government that they are low skill, and will therefore receive the low-skill type’s tax treatment, $\langle c_L^t, y_L^t \rangle$, in each period $t \geq 2$. We omit the low-skill type’s incentive-compatibility constraint, because we make the standard assumption that the redistributive goals of the government create an incentive for high-skill individuals to mimic low-skill individuals, but not vice

versa.³ We also omit the high-skill type’s first-period participation constraint, as we assume that it is their incentive-compatibility constraint that binds.⁴ The solution to programme (2.4)–(2.6) yields the value function $W^1(\phi, w_L, w_H, V_H, \delta)$, which represents the level of social welfare attainable in the first period.

3 The Threat of Emigration and Social Welfare

It is shown in the appendix that:

Proposition *In our dynamic nonlinear income tax model without commitment, the threat of high-skill emigration increases social welfare in period 1 ($\partial W^1(\cdot)/\partial V_H > 0$), decreases social welfare in period $t \geq 2$ ($\partial W^t(\cdot)/\partial V_H < 0$), and has no effect on social welfare summed over the model’s horizon ($\partial W^1(\cdot)/\partial V_H + \sum_{t=2}^{\infty} \delta^{t-1} \partial W^t(\cdot)/\partial V_H = 0$).*

The proposition implies that the threat of high-skill emigration facilitates redistribution and increases social welfare in the short run, and has no effect on the level of social welfare attainable over the long run. Social welfare is increased in period 1 because the threat of emigration relaxes the incentive-compatibility constraint, i.e., it becomes cheaper for the government to obtain skill-type information. In general, high-skill individuals must be offered an attractive tax treatment in period 1 to reveal their type, as compensation for having to face first-best taxation from period 2 onwards. Therefore, the ability of the government to redistribute in period 1 is severely limited. However, the threat of emigration also restricts the ability of the government to redistribute after high-skill individuals have revealed their type. This means that the government in period 1 can offer high-skill individuals a less attractive tax treatment, which in turn means more redistribution and a higher level of social welfare. In each period from period 2

³In fact, high-skill individuals have a stronger incentive to mimic when the government cannot commit than when it can commit. This makes it theoretically possible that ‘pooling’ taxation may be optimal; see, e.g., Brett and Weymark (2008a), Krause (2009b), and Guo and Krause (2015a, 2015b). In this note, however, we focus only on ‘separating’ taxation, under which high-skill and low-skill individuals always receive different tax treatments.

⁴To assume otherwise, i.e., that the high-skill type’s first-period participation constraint is binding and their incentive-compatibility constraint is slack, would effectively mean that the government never faces an information problem. The government would therefore be implementing first-best taxation in every period, albeit subject to a participation constraint.

onwards, the threat of emigration reduces social welfare, because the government cannot implement unrestricted first-best redistribution. Over the long run, however, the threat of emigration has no effect on social welfare, because optimal taxation balances the short-run benefits *vis-a-vis* relaxation of the incentive-compatibility constraint against the long-run costs of the binding participation constraints.

4 Concluding Comments

In this note we have shown, using a simple dynamic nonlinear income tax model without commitment, that the threat of high-skill emigration increases social welfare in the short run and has no effect on social welfare over the long run. While this result may be viewed as being primarily of theoretical interest, it also provides a new example of how restrictions on policy instruments that are clearly welfare-reducing when the government can commit, may in fact be beneficial when the government cannot commit.

5 Appendix

Proof of the Proposition

The Lagrangian corresponding to programme (2.1) – (2.3) is:

$$\begin{aligned} \mathcal{L}^t = & (1 - \phi) \left\{ u(c_L^t) - \frac{y_L^t}{w_L} \right\} + \phi \left\{ u(c_H^t) - \frac{y_H^t}{w_H} \right\} \\ & + \lambda^t \left\{ (1 - \phi) [y_L^t - c_L^t] + \phi [y_H^t - c_H^t] \right\} + \alpha_H^t \left\{ u(c_H^t) - \frac{y_H^t}{w_H} - V_H \right\} \end{aligned} \quad (\text{A.1})$$

where $\lambda^t > 0$ and $\alpha_H^t > 0$ are Lagrange multipliers. The first-order conditions are:

$$(1 - \phi)u'(c_L^t) - \lambda^t(1 - \phi) = 0 \quad (\text{A.2})$$

$$-(1 - \phi)\frac{1}{w_L} + \lambda^t(1 - \phi) = 0 \quad (\text{A.3})$$

$$\phi u'(c_H^t) - \lambda^t \phi + \alpha_H^t u'(c_H^t) = 0 \quad (\text{A.4})$$

$$-\phi \frac{1}{w_H} + \lambda^t \phi - \alpha_H^t \frac{1}{w_H} = 0 \quad (\text{A.5})$$

$$(1 - \phi) [y_L^t - c_L^t] + \phi [y_H^t - c_H^t] = 0 \quad (\text{A.6})$$

$$u(c_H^t) - \frac{y_H^t}{w_H} - V_H = 0 \quad (\text{A.7})$$

By the Envelope Theorem:

$$\frac{\partial W^t(\cdot)}{\partial V_H} = \frac{\partial \mathcal{L}^t(\cdot)}{\partial V_H} = -\alpha_H^t = -\phi \left[\frac{w_H}{w_L} - 1 \right] < 0 \quad (\text{A.8})$$

where the last equality follows from (A.3) and (A.5). This shows that $\partial W^t(\cdot)/\partial V_H < 0$ for all $t \geq 2$.

The Lagrangian corresponding to programme (2.4) – (2.6) is:

$$\begin{aligned} \mathcal{L}^1 = & (1 - \phi) \left\{ u(c_L^1) - \frac{y_L^1}{w_L} \right\} + \phi \left\{ u(c_H^1) - \frac{y_H^1}{w_H} \right\} + \lambda^1 \{ (1 - \phi) [y_L^1 - c_L^1] + \phi [y_H^1 - c_H^1] \} \\ & + \theta_H^1 \left\{ u(c_H^1) - \frac{y_H^1}{w_H} + \frac{\delta V_H}{1 - \delta} - u(c_L^1) + \frac{y_L^1}{w_H} - \frac{\delta u(c_L^t(\cdot))}{1 - \delta} + \frac{\delta y_L^t(\cdot)}{(1 - \delta)w_H} \right\} \end{aligned} \quad (\text{A.9})$$

where $\lambda^1 > 0$ and $\theta_H^1 > 0$ are Lagrange multipliers, and use has been made of:

$$\sum_{t=2}^{\infty} \delta^{t-1} V_H = \frac{\delta V_H}{1 - \delta} \quad \text{and} \quad \sum_{t=2}^{\infty} \delta^{t-1} \left[u(c_L^t(\cdot)) - \frac{y_L^t(\cdot)}{w_H} \right] = \frac{\delta}{1 - \delta} \left[u(c_L^t(\cdot)) - \frac{y_L^t(\cdot)}{w_H} \right] \quad (\text{A.10})$$

The first-order conditions on y_L^1 and y_H^1 are, respectively:

$$-(1 - \phi) \frac{1}{w_L} + \lambda^1 (1 - \phi) + \theta_H^1 \frac{1}{w_H} = 0 \quad (\text{A.11})$$

$$-\phi \frac{1}{w_H} + \lambda^1 \phi - \theta_H^1 \frac{1}{w_H} = 0 \quad (\text{A.12})$$

By the Envelope Theorem:

$$\frac{\partial W^1(\cdot)}{\partial V_H} = \frac{\partial \mathcal{L}^1(\cdot)}{\partial V_H} = \frac{\delta \theta_H^1}{1 - \delta} \left[1 - u'(c_L^t) \frac{\partial c_L^t(\cdot)}{\partial V_H} + \frac{1}{w_H} \frac{\partial y_L^t(\cdot)}{\partial V_H} \right] = \frac{\delta \theta_H^1}{1 - \delta} \left[1 - \frac{1}{w_L} \frac{\partial c_L^t(\cdot)}{\partial V_H} + \frac{1}{w_H} \frac{\partial y_L^t(\cdot)}{\partial V_H} \right] \quad (\text{A.13})$$

where the last equality makes use of (A.2) and (A.3). By applying the Implicit Function

Theorem and Cramer's Rule to (A.2) – (A.7), it can be shown that:

$$\frac{\partial c_L^t(\cdot)}{\partial V_H} = 0 \quad \text{and} \quad \frac{\partial y_L^t(\cdot)}{\partial V_H} = \frac{\phi w_H}{1 - \phi} \quad (\text{A.14})$$

Equation (A.13) can then be simplified to:

$$\frac{\partial W^1(\cdot)}{\partial V_H} = \frac{\partial \mathcal{L}^1(\cdot)}{\partial V_H} = \frac{\delta \phi}{1 - \delta} \left[\frac{w_H}{w_L} - 1 \right] > 0 \quad (\text{A.15})$$

where use has been made of (A.11) and (A.12). This shows that $\partial W^1(\cdot)/\partial V_H > 0$.

Finally, equation (A.8) implies that:

$$\sum_{t=2}^{\infty} \delta^{t-1} \frac{\partial W^t(\cdot)}{\partial V_H} = \frac{-\delta \phi}{1 - \delta} \left[\frac{w_H}{w_L} - 1 \right] \quad (\text{A.16})$$

which with (A.15) shows that $\partial W^1(\cdot)/\partial V_H + \sum_{t=2}^{\infty} \delta^{t-1} \partial W^t(\cdot)/\partial V_H = 0$. ■

References

- [1] Apps, P. and R. Rees (2006), “Repeated Optimal Nonlinear Income Taxation”, *mimeo*.
- [2] Berliant, M. and J. Ledyard (2014), “Optimal Dynamic Nonlinear Income Taxes with No Commitment”, *Journal of Public Economic Theory*, 16, 196-221.
- [3] Brett, C. and J. Weymark (2008a), “Optimal Nonlinear Taxation of Income and Savings without Commitment”, *Vanderbilt University Working Paper*, No. 08-W05.
- [4] Brett, C. and J. Weymark (2008b), “The Impact of Changing Skill Levels on Optimal Nonlinear Income Taxes”, *Journal of Public Economics*, 92, 1765-1771.
- [5] Brett, C. and J. Weymark (2011), “How Optimal Nonlinear Income Taxes Change when the Distribution of the Population Changes”, *Journal of Public Economics*, 95, 1239-1247.
- [6] Guo, J-T. and A. Krause (2011), “Optimal Nonlinear Income Taxation with Habit Formation”, *Journal of Public Economic Theory*, 13, 463-480.
- [7] Guo, J-T. and A. Krause (2013), “Optimal Nonlinear Taxation of Income and Education Expenditures”, *Oxford Economic Papers*, 65, 74-95.
- [8] Guo, J-T. and A. Krause (2014), “Optimal Dynamic Nonlinear Income Taxation under Loose Commitment”, *Macroeconomic Dynamics*, 18, 1403-1427.
- [9] Guo, J-T. and A. Krause (2015a), “Dynamic Nonlinear Income Taxation with Quasi-Hyperbolic Discounting and No Commitment”, *Journal of Economic Behavior and Organization*, 109, 101-119.
- [10] Guo, J-T. and A. Krause (2015b), “Dynamic Income Taxation without Commitment: Comparing Alternative Tax Systems”, *Economic Modelling*, 47, 319-326.
- [11] Krause, A. (2009a), “Education and Taxation Policies in the Presence of Countervailing Incentives”, *Economica*, 76, 387-399.
- [12] Krause, A. (2009b), “Optimal Nonlinear Income Taxation with Learning-by-Doing”, *Journal of Public Economics*, 93, 1098-1110.
- [13] Leite-Monteiro, M. (1997), “Redistributive Policy with Labour Mobility Across Countries”, *Journal of Public Economics*, 65, 229-244.
- [14] Mirrlees, J. (1971), “An Exploration in the Theory of Optimum Income Taxation”, *Review of Economic Studies*, 38, 175-208.
- [15] Osmundsen, P. (1999), “Taxing Internationally Mobile Individuals—A Case of Countervailing Incentives”, *International Tax and Public Finance*, 6, 149-164.

- [16] Roberts, K. (1984), “The Theoretical Limits to Redistribution”, *Review of Economic Studies*, 51, 177-195.
- [17] Simula, L. (2010), “Optimal Nonlinear Income Tax and Nonlinear Pricing: Optimality Conditions and Comparative Static Properties”, *Social Choice and Welfare*, 35, 199-220.
- [18] Simula, L. and A. Trannoy (2010), “Optimal Income Tax under the Threat of Migration by Top-Income Earners”, *Journal of Public Economics*, 94, 163-173.
- [19] Simula, L. and A. Trannoy (2012), “Shall we keep the Highly Skilled at Home? The Optimal Income Tax Perspective”, *Social Choice and Welfare*, 39, 751-782.
- [20] Weymark, J. (1987), “Comparative Static Properties of Optimal Nonlinear Income Taxes”, *Econometrica*, 55, 1165-1185.