Optimal Nonlinear Income Taxation with Habit Formation

Jang-Ting Guo*
University of California, Riverside

Alan Krause†
University of York

February 8, 2009

Abstract

It has recently been shown that incorporating "keeping up with the Joneses" preferences into a prototypical two-ability-type optimal nonlinear income tax model leads to higher marginal tax rates for both types of agents. In particular, the high-skill type faces a positive marginal tax rate, rather than zero as in the conventional case. In this paper, agents' utility functions are postulated to exhibit "habit formation in consumption" such that the prototypical two-ability-type optimal nonlinear income tax model becomes a dynamic analytical framework. We show that if the government can commit to its future fiscal policy, the presence of consumption habits does not affect the standard results on optimal marginal tax rates. By contrast, if the government cannot pre-commit, the high-skill type will face a negative marginal tax rate, while the low-skill type's marginal tax rate remains positive.

Keywords: Income Taxation; Habit Formation; Commitment.

JEL Classifications: H21, H24.

---

*Department of Economics, 4128 Sproul Hall, University of California, Riverside, CA, 92521, U.S.A., Phone: 1-951-827-1588, Fax: 1-951-827-5685, E-mail: guojt@ucr.edu.

†Corresponding Author. Department of Economics and Related Studies, University of York, Heslington, York, YO10 5DD, U.K., Phone: 44-(0)1904-433-572, Fax: 44-(0)1904-433-759, E-mail: ak519@york.ac.uk.

1 Introduction

There is now an extensive literature that examines the macroeconomic effects of "keeping up with the Joneses" and/or "habit formation in consumption" preferences within dynamic general equilibrium models; see recent examples of Carroll [2000], Ljungqvist and Uhlig [2000], Dupor and Liu [2003], Alonso-Carrera, et al. [2004], and Guo [2005]. A "keeping up with the Joneses" utility specification captures the idea that individuals may evaluate their own consumption relative to that of their neighbours, thus individual consumption imposes a negative externality on others. On the other hand, "habit formation in consumption" preferences imply that individuals evaluate their current consumption relative to their own past consumption, hence increases in current consumption will, *ceteris paribus*, reduce the utility from future consumption. The popularity of these preference formulations follows from their success in explaining various macroeconomic phenomena such as business cycle fluctuations (Lettau and Uhlig [2000]), the interrelations between savings and economic growth (Carroll, et al. [2000]) and the equity premium puzzle (Abel [1990, 1999], Constantinides [1990], Gali [1994] and Campbell and Cochrane [1999]), among others.

In the field of public economics, Aronsson and Johansson-Stenman [2008a] recently have introduced "keeping up with the Joneses" preferences to a prototypical two-ability-type Mirrlees/Stiglitz model of optimal nonlinear income taxation.[1] In the standard framework, agents derive utility only from their own consumption, and the government seeks to redistribute from the high-skill to the low-skill type. However, since an individual's skill type is assumed to be private information, the government cannot implement (the first-best) personalised lump-sum taxes. Instead, the government is constrained to use (the second-best) incentive-compatible taxation in which individuals are willing to reveal their skill types. In this case, it is well known that the optimal incentive-compatible tax system is characterised by a zero marginal tax rate on the high-skill type together with a positive marginal tax rate on the low-skill type, where the latter is used to distort the low-skill type's labour supply downwards to relax an incentive-compatibility

[1] See also the related papers by Oswald [1983] and Ireland [2001], among others.

constraint. By considering a "keeping up with the Joneses" utility function within this setting,[2] Aronsson and Johansson-Stenman [2008a] provide another justification for marginal tax-rate distortions, namely, a corrective motive to overturn the effects of a negative consumption externality. These authors show that "keeping up with the Joneses" preferences lead to higher optimal marginal tax rates than in the conventional case, hence both types of agents face positive marginal tax rates. The intuition for higher marginal tax rates is to distort the labour supply of both types downwards, thereby lowering aggregate income and consumption, which in turn attenuates the effects of negative consumption externalities.

This paper complements Aronsson and Johansson-Stenman's [2008a] analysis by incorporating "habit formation in consumption" preferences — where an agent's utility depends on the difference between her current consumption and a fraction of the past level of her own consumption — into a prototypical two-ability-type nonlinear income tax model. Since consumption habits turn the standard model into an intertemporal setting, we examine a dynamic economy with two time periods.[3] Although an infinite-horizon formulation might be considered more general, our two-period framework allows the results to be derived in a clear and intuitive manner. Moreover, the first period of our two-period model is sufficient to capture the essential problem of dynamic taxation. That is, when choosing its present tax policy, the government must consider how its choices affect its taxation possibilities in the future.[4]

Our main results can be summarised as follows: (i) If in the first period the government can commit to its second-period tax policy, habit formation in consumption exerts

[2] In particular, an agent's utility depends on the difference between her own consumption and a fraction of the contemporaneous level of average consumption in the economy.

[3] Berliant and Ledyard [2005], Apps and Rees [2006], and Brett and Weymark [2008] also work with two-period optimal nonlinear income tax models. In Berliant and Ledyard [2005], there is a continuum of types. In Apps and Rees [2006] and Brett and Weymark [2008], there are only two types; but there is a continuum of consumers of each type in Apps and Rees [2006], whereas there is a single consumer of each type in Brett and Weymark [2008]. Our model is therefore most closely related to that of Apps and Rees [2006], although methodologically our analysis is closer to Brett and Weymark [2008].

[4] In a recent paper, Aronsson and Johansson-Stenman [2008b] consider habit formation in a dynamic overlapping generations model, which is quite different to the prototypical Mirrlees/Stiglitz nonlinear income tax framework. They also consider only the case when the government can commit to its future tax policy, while our focus is on the case when the government cannot commit.

no impact on the standard results that the high-skill type should face a zero marginal tax rate and the low-skill type should face a positive marginal tax rate. This is because the consumers rationally consider the effects on their second-period utility when deciding their first-period consumption levels. Hence, there is no need for the government to implement marginal tax-rate distortions to correct the negative "internality" that an individual's first-period consumption imposes on their second-period utility. (ii) If the government cannot commit to its second-period tax policy, the presence of consumption habits lowers the marginal tax rate faced by the high-skill agents, meaning they now face a negative marginal tax rate, whereas habit formation raises the low-skill type's marginal tax rate, which means it remains positive. As in the commitment case, there is still no corrective motive for marginal tax-rate distortions; but when the government cannot pre-commit, consumption habits yield different optimal marginal tax rates to relax an incentive-compatibility constraint. Therefore, while Aronsson and Johansson-Stenman [2008a] show that "keeping up with the Joneses" preferences move the recommendations of optimal tax theory closer to actual practice by providing a rationale for a positive marginal tax rate on the highest-skill workers, we show that "habit formation in consumption" cannot help reconcile optimal tax theory with real-world fiscal systems.

The remainder of this paper is organised as follows. Section 2 presents our two-period model economy. Section 3 examines optimal income taxation with commitment, and Section 4 examines optimal income taxation without commitment. Section 5 discusses the implications of considering a "catching up with the Joneses" preference specification in our model. Section 6 contains some concluding remarks. Proofs and other mathematical details are relegated to the Appendix.

2 The Model

We consider a simple two-period model with a continuum of consumers of unit measure, with a proportion $\phi \in (0, 1)$ being high-skill workers and the remaining $(1 - \phi)$ being low-skill workers. Consumption by a type i consumer ($i = 1, 2$) in period t ($t = 1, 2$) is denoted by c_i^t, and labour supply by type i in period t is denoted by l_i^t. Type i's wage in period t is denoted by w_i^t, and it is assumed that $w_2^1 > w_1^1$ and $w_2^2 > w_1^2$ so that type-1 consumers are low-skill workers and type-2 consumers are high-skill workers. Type i's pre-tax income in period t is denoted by $y_i^t = w_i^t l_i^t$.

The consumers' utility functions are given by $u(c_i^1) - v(l_i^1)$ in period 1, and $u(c_i^2 - \gamma c_i^1) - v(l_i^2)$ in period 2, where $u'(\cdot) > 0$, $u''(\cdot) < 0$, $v'(\cdot) > 0$ and $v''(\cdot) > 0$. It can be seen that consumption in period 1 sets a standard of living (or a "habit" level of consumption) which is used to evaluate the utility of consumption in period 2. The second-period utility function corresponds to the typical habit-formation model as specified, e.g., by Alonso-Carrera, et al. [2004]. The parameter $\gamma \in (0,1)$ measures the importance of first-period consumption in evaluating the second-period level of utility. Since $\gamma > 0$, a *ceteris paribus* increase in the first-period consumption reduces the second-period utility, thus a negative "internality" is present.

It is well known that in models with a finite number of types, the optimal income tax schedule may not be differentiable. We therefore follow the standard practice of deriving expressions for "implicit" marginal tax rates in terms of derivatives of the utility function. To this end, suppose the consumers face smooth nonlinear income tax functions $T^1(y_i^1)$ and $T^2(y_i^2)$ in periods 1 and 2, respectively. Then type i's behaviour can be described by the following programme:

$$\max_{c_i^1,\, l_i^1,\, c_i^2,\, l_i^2} u(c_i^1) - v(l_i^1) + \delta \left[u(c_i^2 - \gamma c_i^1) - v(l_i^2) \right], \tag{2.1}$$

subject to:

$$c_i^1 \leq y_i^1 - T^1(y_i^1), \tag{2.2}$$

$$c_i^2 \leq y_i^2 - T^2(y_i^2), \tag{2.3}$$

where $\delta \in (0,1]$ is the discount factor. In order to isolate the effects of habit formation, we assume that there are no savings. It is shown in the Appendix that the solution to programme (2.1) – (2.3) yields the following expressions for the marginal tax rates:

$$MTR_i^1 := \frac{\partial T^1(\cdot)}{\partial y_i^1} = 1 - \frac{v'(l_i^1)}{\left[u'(c_i^1) - \gamma\delta u'(c_i^2 - \gamma c_i^1)\right] w_i^1}, \tag{2.4}$$

and

$$MTR_i^2 := \frac{\partial T^2(\cdot)}{\partial y_i^2} = 1 - \frac{v'(l_i^2)}{u'(c_i^2 - \gamma c_i^1)w_i^2}, \tag{2.5}$$

where MTR_i^t denotes the marginal tax rate faced by type i in period t. The expression for the second-period marginal tax rate given in (2.5) corresponds, as in static settings, to one minus the marginal rate of substitution of pre-tax income for consumption.[5] This is because period 2 is the last period of our model. The first-period marginal tax rate, however, is not simply equal to one minus the marginal rate of substitution. The expression for the first-period marginal tax rate given in (2.4) is complicated by the dynamic nature of our model due to the presence of habit formation.

3 Optimal Income Taxation with Commitment

If the period-1 government can commit to its second-period tax policy, the government can be described as choosing "lifetime" tax treatments $\langle c_1^1, y_1^1, c_1^2, y_1^2\rangle$ and $\langle c_2^1, y_2^1, c_2^2, y_2^2\rangle$ for types 1 and 2, respectively, to maximise:[6]

$$(1-\phi)\left\{u(c_1^1) - v\left(\frac{y_1^1}{w_1^1}\right) + \delta\left[u(c_1^2 - \gamma c_1^1) - v\left(\frac{y_1^2}{w_1^2}\right)\right]\right\} + \phi\left\{u(c_2^1) - v\left(\frac{y_2^1}{w_2^1}\right) + \delta\left[u(c_2^2 - \gamma c_2^1) - v\left(\frac{y_2^2}{w_2^2}\right)\right]\right\}, \tag{3.1}$$

subject to:

$$(1-\phi)\left[y_1^1 - c_1^1\right] + \phi\left[y_2^1 - c_2^1\right] \geq 0, \tag{3.2}$$

$$(1-\phi)\left[y_1^2 - c_1^2\right] + \phi\left[y_2^2 - c_2^2\right] \geq 0, \tag{3.3}$$

$$u(c_2^1) - v\left(\frac{y_2^1}{w_2^1}\right) + \delta\left[u(c_2^2 - \gamma c_2^1) - v\left(\frac{y_2^2}{w_2^2}\right)\right] \geq u(c_1^1) - v\left(\frac{y_1^1}{w_2^1}\right) + \delta\left[u(c_1^2 - \gamma c_1^1) - v\left(\frac{y_1^2}{w_2^2}\right)\right], \tag{3.4}$$

[5] See, e.g., Stiglitz [1982].

[6] A lifetime tax treatment consists of pre-tax income and post-tax income (which is equal to consumption) in each period. The difference between pre-tax income and consumption is total taxes paid (or transfers received).

where (3.1) is a utilitarian social welfare function with the consumers' utility functions written in terms of the government's choice variables c_i^t and y_i^t, (3.2) and (3.3) are the government's first- and second-period budget constraints,[7] and (3.4) is type-2's incentive-compatibility constraint. As in static nonlinear income tax models, we assume that each consumer's skill type is private information. The government must therefore satisfy incentive-compatibility constraints to induce each type to choose their intended tax treatment, rather than "mimicking" the other type by choosing the other type's tax treatment. However, we omit the low-skill type's incentive-compatibility constraint, since we focus on what Stiglitz [1982] calls the "normal" case and what Guesnerie [1995] calls "redistributive equilibria", as the redistributive goals of the government imply that the high-skill type's incentive-compatibility constraint binds, whereas the low-skill type's incentive-compatibility constraint is always slack.[8]

It is shown in the Appendix that the solution to programme (3.1) − (3.4) yields:

Proposition 1 *When the government is able to commit to its second-period tax policy, optimal nonlinear income taxation under habit formation in consumption is characterised by*: $MTR_1^1 > 0$, $MTR_2^1 = 0$, $MTR_1^2 > 0$, *and* $MTR_2^2 = 0$.

The presence of habit formation therefore has no effect on the pattern of optimal marginal tax rate distortions when the government can pre-commit. The intuition is quite straightforward: since the consumers take into account how their first-period consumption affects their second-period utility, there is no need for the government to engage in corrective taxation. Hence, the only motive the government has to implement marginal tax-rate distortions is the same as that in the standard model, i.e., to distort the low-skill type's labour supply downwards to relax the high-skill type's incentive-compatibility constraint.

[7] As with the consumers, we do not permit the government to save.

[8] Aronsson and Johansson-Stenman [2008a] also make this assumption.

4 Optimal Income Taxation without Commitment

If the government cannot commit to its second-period tax policy, the government can use skill-type information revealed in period 1 to implement (the first-best) personalised lump-sum taxes in period 2. In this case, the government's behaviour in period 2 can be described as follows. Choose $\langle c_1^2, y_1^2 \rangle$ and $\langle c_2^2, y_2^2 \rangle$ to maximise:

$$(1-\phi)\left[u(c_1^2 - \gamma c_1^1) - v\left(\frac{y_1^2}{w_1^2}\right)\right] + \phi\left[u(c_2^2 - \gamma c_2^1) - v\left(\frac{y_2^2}{w_2^2}\right)\right], \tag{4.1}$$

subject to:

$$(1-\phi)\left[y_1^2 - c_1^2\right] + \phi\left[y_2^2 - c_2^2\right] \geq 0, \tag{4.2}$$

where (4.1) is the second-period social welfare function, and (4.2) is the government's second-period budget constraint. Since the government can identify the consumers, it does not face incentive-compatibility constraints. The solution to programme (4.1) – (4.2) yields the functions $c_1^2(\phi, \gamma, c_1^1, w_1^2, c_2^1, w_2^2)$, $y_1^2(\phi, \gamma, c_1^1, w_1^2, c_2^1, w_2^2)$, $c_2^2(\phi, \gamma, c_1^1, w_1^2, c_2^1, w_2^2)$ and $y_2^2(\phi, \gamma, c_1^1, w_1^2, c_2^1, w_2^2)$. Substituting these functions into (4.1) leads to the value function $W^2(\phi, \gamma, c_1^1, w_1^2, c_2^1, w_2^2)$.

The consumers and the government know that the government will solve programme (4.1) – (4.2) in period 2. Therefore, the government in period 1 can be described as choosing $\langle c_1^1, y_1^1 \rangle$ and $\langle c_2^1, y_2^1 \rangle$ to maximise:

$$(1-\phi)\left[u(c_1^1) - v\left(\frac{y_1^1}{w_1^1}\right)\right] + \phi\left[u(c_2^1) - v\left(\frac{y_2^1}{w_2^1}\right)\right] + \delta W^2(\phi, \gamma, c_1^1, w_1^2, c_2^1, w_2^2), \tag{4.3}$$

subject to:

$$(1-\phi)\left[y_1^1 - c_1^1\right] + \phi\left[y_2^1 - c_2^1\right] \geq 0, \tag{4.4}$$

$$u(c_2^1) - v\left(\frac{y_2^1}{w_2^1}\right) + \delta\left[u(c_2^2(\cdot) - \gamma c_2^1) - v\left(\frac{y_2^2(\cdot)}{w_2^2}\right)\right] \geq u(c_1^1) - v\left(\frac{y_1^1}{w_2^1}\right) + \delta\left[u(c_1^2(\cdot) - \gamma c_1^1) - v\left(\frac{y_1^2(\cdot)}{w_2^2}\right)\right]. \tag{4.5}$$

When choosing $\langle c_1^1, y_1^1 \rangle$ and $\langle c_2^1, y_2^1 \rangle$, the government considers how its choice will affect the level of social welfare attainable in period 2. The first-period social welfare function (4.3) therefore includes the second-period value function $W^2(\cdot)$. Equation (4.4) is the government's first-period budget constraint, while (4.5) is type 2's incentive-compatibility constraint. Since type-2 consumers know that they will be subjected to the first-best taxation in period 2 if they reveal their skill type in period 1, they must be offered a relatively favourable tax treatment in period 1 to compensate for the relatively unfavourable tax treatment they will receive in period 2.[9] Consequently, in order for a type-2 consumer to be willing to reveal her skill type in period 1, the utility she obtains from choosing $\langle c_2^1, y_2^1 \rangle$ in period 1 plus the utility from the first-best tax treatment $\langle c_2^2(\cdot), y_2^2(\cdot) \rangle$ that she must accept in period 2 must be greater than or equal to what she could obtain from choosing $\langle c_1^1, y_1^1 \rangle$ in period 1 plus the utility from the low-skill type's first-best tax treatment $\langle c_1^2(\cdot), y_1^2(\cdot) \rangle$ in period 2. That is, if a type-2 consumer chooses $\langle c_1^1, y_1^1 \rangle$ in period 1, she is announcing to the government that she is a low-skill worker and will therefore be treated as such in the second period.

It is shown in the Appendix that the solutions to programmes (4.1) – (4.2) and (4.3) – (4.5) together imply:

Proposition 2 *When the government is unable to commit to its second-period tax policy, optimal nonlinear income taxation under habit formation in consumption is characterised by*: $MTR_1^1 > 0$, $MTR_2^1 < 0$, $MTR_1^2 = 0$, *and* $MTR_2^2 = 0$.

Both types face zero marginal tax rates in period 2 simply because the first-best taxation is used in that period. To understand why type-2 consumers face a negative marginal tax rate in period 1, note that in period 2 the government will choose c_1^2 and c_2^2 such that $u'(c_1^2 - \gamma c_1^1) = u'(c_2^2 - \gamma c_2^1)$, taking c_1^1 and c_2^1 as given.[10] Now consider a marginal increase in c_2^1. Since $u(\cdot)$ is strictly concave, the marginal utility of c_2^2 becomes greater than that of c_1^2, so the government can raise the second-period social welfare by transferring a small amount of consumption from type-1 to type-2 consumers. Thus under the first-best taxation in period 2, an individual's utility is *increasing* in own

[9]Indeed, the favourable tax treatment offered high-skill consumers in period 1 to obtain skill-type information could be very costly from a social welfare point of view. Instead, it is theoretically possible that the government may prefer to pool some or all of the high-skill consumers with the low-skill consumers in period 1 so that skill-type information is not revealed, even though it is then constrained to use the second-best taxation in period 2. However, we focus only on the more economically interesting case in which the high-skill and low-skill consumers all receive different tax treatments.

[10]See equations (A.24) and (A.26) in the Appendix.

first-period consumption.[11] It follows that the government can increase type-2's second-period utility by raising their first-period consumption. This is achieved through a negative marginal tax rate to distort type-2's first-period labour supply upwards, which in turn generates additional income that can be used to increase type-2's first-period consumption. The reason the government wants to raise the utility type-2 consumers obtain under the first-best taxation in period 2 is that it makes them more willing to reveal their skill type in period 1, i.e., it relaxes the incentive-compatibility constraint. Moreover, the increase in type 2's second-period utility necessitates a reduction in type 1's second-period utility, which further serves to make mimicking less attractive and relax the incentive-compatibility constraint.

The first-period marginal tax rate faced by type-1 consumers is positive as in the standard framework, but now there are three factors at work. First, the government wants to distort type-1 workers' first-period labour supply downwards for the same reasons as in the prototypical setting. Second, habit formation implies that the government wants to distort type 1's first-period labour supply downwards to lower their first-period income and consumption, since this reduces the utility they obtain in period 2 under the first-best taxation (recall that individual utility in period 2 is increasing in own first-period consumption). This makes mimicking less attractive, thus relaxing the incentive-compatibility constraint. Third, the decrease in type-1's second-period utility comes with an increase in type-2's second-period utility, making type-2 consumers more willing to reveal their type which again relaxes the incentive-compatibility constraint. As a result, all three factors act to make it optimal for the government to distort type 1's first-period labour supply downwards, which implies a positive marginal tax rate.

We conclude this section with two comments. First, the preceding discussion makes it clear that the only motive the government has to implement marginal tax-rate distortions is to relax an incentive-compatibility constraint. As in the commitment case, there is no role for any form of corrective taxation. Second, habit formation lowers the marginal tax

[11] While this might seem counter-intuitive, it is similar to the standard result that individual utility under the first-best taxation is decreasing in the wage rate, since all individuals receive the same level of consumption, but higher-wage individuals are required to work longer.

rate faced by the high-skill type, but raises the marginal tax rate faced by the low-skill type. Therefore, unlike Aronsson and Johansson-Stenman [2008a] with “keeping up with the Joneses” preferences, habit formation exerts opposite effects on the high-skill and low-skill type’s optimal marginal tax rates.

5 Discussion

In Aronsson and Johansson-Stenman [2008a], individuals evaluate their current consumption relative to *contemporaneous* average consumption in the economy. This class of preferences is known as “keeping up with the Joneses”. By contrast, “catching up with the Joneses” postulates that an individual’s current consumption is compared with *past* levels of average consumption. As a result, similar to the specification with habit formation, there exists intertemporal dependence between consumptions at different time periods. Here, we use our dynamic two-period model to examine the “catching up” case in which consumer i’s objective function is given by:

$$\max_{c_i^1,\, l_i^1,\, c_i^2,\, l_i^2} u(c_i^1) - v(l_i^1) + \delta \left[u(c_i^2 - \gamma \bar{c}^1) - v(l_i^2) \right], \tag{5.1}$$

where $\bar{c}^1$ represents the economy’s period-1 average consumption.

The main results can be summarised as follows:[12] (i) When the government can pre-commit, the standard optimal marginal tax rate results apply in period 2 simply because this is the last period and contemporaneous consumption does not generate an externality. In period 1, however, both the high-skill and low-skill types face higher marginal tax rates than in the prototypical model for the same corrective-motive reasons *a la* Aronsson and Johansson-Stenman [2008a]. Hence, it makes no difference whether one works with “catching up” or “keeping up with the Joneses” preferences when the government can commit to its future tax policy. (ii) When the government cannot pre-commit, both types of agents face a zero marginal tax rate in period 2 as this is the last period, the first-best taxation is possible, and contemporaneous consumption

[12] The detailed derivations for these results are available upon request.

does not generate an externality. In period 1, however, the marginal tax rates faced by both types cannot be signed. On the one hand, the government wants to impose higher marginal tax rates to reduce first-period consumption and correct the effects of consumption externalities as in Aronsson and Johansson-Stenman [2008a]. But on the other hand, the government wants to lower marginal tax rates to raise first-period consumption to relax an incentive-compatibility constraint for reasons similar as to why it does so under habit formation. Since the corrective and incentive motives for marginal tax-rate distortions work in opposite directions under "catching up with the Joneses" when the government cannot pre-commit, the optimal marginal tax rates cannot be signed.

6 Concluding Comments

By introducing "keeping up with the Joneses" preferences to a prototypical optimal nonlinear income tax model, Aronsson and Johansson-Stenman [2008a] have moved the recommendations of normative tax theory closer to the characteristics of real-world tax systems, in that a positive marginal tax rate on the highest-skill workers can be justified. In this paper, we have incorporated a closely-related class of preferences, known as "habit formation in consumption", into the standard two-ability-type optimal nonlinear income tax framework. Our analysis shows that consumption habits affects the conventional results only if the government cannot commit to its future tax policy, which raises the question as to whether governments can pre-commit in practice. On the one hand, one could argue that the commitment assumption is justified by the observation that real-world income tax schedules are not frequently redesigned.[13] On the other hand, the no-commitment assumption might be justified by the fact that the present government cannot impose binding constraints on future governments.[14] Moreover, the second-best nature of the standard optimal nonlinear income tax framework stems from the

[13] Gaube [2007] makes this argument.

[14] For example, Auerbach [2006] cites a proposal regarding the problem of the U.S. Social Security system's imbalance, which includes a tax increase to be made in 2045. As Auerbach suggests, it is highly unlikely that the government presiding in 2045 will feel constrained by a decision taken by another government over 40 years earlier.

assumption that skill types are private information. But taxation in earlier periods may result in this information being revealed, which would enable the government to implement the first-best (lump-sum) taxation in latter periods. Hence, ruling out lump-sum taxes in a dynamic nonlinear income tax model via a commitment assumption might be considered somewhat artificial. In any event, our results show that habit formation in consumption, regardless of whether the government can commit or not, cannot help reconcile optimal tax theory with the tax systems observed in practice. Unlike "keeping up with the Joneses" preferences, the feature of habit formation, if anything, moves the recommendations of optimal tax theory further away from the observed practice.

7 Appendix

Derivation of Equations (2.4) and (2.5)

The Lagrangian corresponding to programme (2.1) – (2.3) can be written as:

$$L = u(c_i^1) - v(l_i^1) + \delta\left[u(c_i^2 - \gamma c_i^1) - v(l_i^2)\right] \\ + \ \alpha^1\left[w_i^1 l_i^1 - T^1(w_i^1 l_i^1) - c_i^1\right] + \alpha^2\left[w_i^2 l_i^2 - T^2(w_i^2 l_i^2) - c_i^2\right], \quad \text{(A.1)}$$

where $\alpha^1 \geq 0$ and $\alpha^2 \geq 0$ are Lagrange multipliers. The relevant first-order conditions can be written as:

$$u'(c_i^1) - \gamma\delta u'(c_i^2 - \gamma c_i^1) - \alpha^1 = 0, \quad \text{(A.2)}$$

$$-v'(l_i^1) + \alpha^1 w_i^1\left[1 - \frac{\partial T^1(\cdot)}{\partial y_i^1}\right] = 0, \quad \text{(A.3)}$$

$$\delta u'(c_i^2 - \gamma c_i^1) - \alpha^2 = 0, \quad \text{(A.4)}$$

$$-\delta v'(l_i^2) + \alpha^2 w_i^2\left[1 - \frac{\partial T^2(\cdot)}{\partial y_i^2}\right] = 0. \quad \text{(A.5)}$$

Straightforward manipulation of (A.2) and (A.3) yields equation (2.4), while straightforward manipulation of (A.4) and (A.5) yields equation (2.5). ■

Proof of Proposition 1

The relevant first-order conditions corresponding to programme (3.1) – (3.4) are:

$$(1-\phi-\theta_2)\left[u'(c_1^1)-\gamma\delta u'(c_1^2-\gamma c_1^1)\right]-\lambda^1(1-\phi)=0, \tag{A.6}$$

$$-(1-\phi)v'\left(\frac{y_1^1}{w_1^1}\right)\frac{1}{w_1^1}+\lambda^1(1-\phi)+\theta_2 v'\left(\frac{y_1^1}{w_2^1}\right)\frac{1}{w_2^1}=0, \tag{A.7}$$

$$(\phi+\theta_2)\left[u'(c_2^1)-\gamma\delta u'(c_2^2-\gamma c_2^1)\right]-\lambda^1\phi=0, \tag{A.8}$$

$$-(\phi+\theta_2)v'\left(\frac{y_2^1}{w_2^1}\right)\frac{1}{w_2^1}+\lambda^1\phi=0, \tag{A.9}$$

$$(1-\phi-\theta_2)\delta u'(c_1^2-\gamma c_1^1)-\lambda^2(1-\phi)=0, \tag{A.10}$$

$$-(1-\phi)\delta v'\left(\frac{y_1^2}{w_1^2}\right)\frac{1}{w_1^2}+\lambda^2(1-\phi)+\theta_2\delta v'\left(\frac{y_1^2}{w_2^2}\right)\frac{1}{w_2^2}=0, \tag{A.11}$$

$$(\phi+\theta_2)\delta u'(c_2^2-\gamma c_2^1)-\lambda^2\phi=0, \tag{A.12}$$

$$-(\phi+\theta_2)\delta v'\left(\frac{y_2^2}{w_2^2}\right)\frac{1}{w_2^2}+\lambda^2\phi=0, \tag{A.13}$$

where $\lambda^1 \geq 0$ is the multiplier on the government's first-period budget constraint (3.2), $\lambda^2 \geq 0$ is the multiplier on the government's second-period budget constraint (3.3), and $\theta_2 \geq 0$ is the multiplier on type 2's incentive-compatibility constraint (3.4).

Dividing (A.13) by (A.12) and rearranging yields:

$$\frac{v'(l_2^2)}{u'(c_2^2-\gamma c_2^1)w_2^2}=1, \tag{A.14}$$

which using (2.5) establishes that $MTR_2^2=0$. Similarly, dividing (A.9) by (A.8) and rearranging yields:

$$\frac{v'(l_2^1)}{\left[u'(c_2^1)-\gamma\delta u'(c_2^2-\gamma c_2^1)\right]w_2^1}=1, \tag{A.15}$$

which using (2.4) establishes that $MTR_2^1=0$.

Using (A.10) and (A.11), we obtain:

$$(1-\phi-\theta_2)\delta u'(c_1^2-\gamma c_1^1)=(1-\phi)\delta v'\left(\frac{y_1^2}{w_1^2}\right)\frac{1}{w_1^2}-\theta_2\delta v'\left(\frac{y_1^2}{w_2^2}\right)\frac{1}{w_2^2}. \tag{A.16}$$

Because $w_2^2 > w_1^2$ and $v(\cdot)$ is strictly convex:

$$(1-\phi)\delta v'\left(\frac{y_1^2}{w_1^2}\right)\frac{1}{w_1^2}-\theta_2\delta v'\left(\frac{y_1^2}{w_2^2}\right)\frac{1}{w_2^2} > (1-\phi)\delta v'\left(\frac{y_1^2}{w_1^2}\right)\frac{1}{w_1^2}-\theta_2\delta v'\left(\frac{y_1^2}{w_1^2}\right)\frac{1}{w_1^2}. \quad \text{(A.17)}$$

Therefore, (A.16) and (A.17) imply that:

$$(1-\phi-\theta_2)\delta u'(c_1^2-\gamma c_1^1) > (1-\phi-\theta_2)\delta v'\left(\frac{y_1^2}{w_1^2}\right)\frac{1}{w_1^2}. \quad \text{(A.18)}$$

Using (A.12) and (A.10), it follows that $1-\phi-\theta_2 > 0$. Hence, (A.18) can be rearranged to yield:

$$1 > \frac{v'(l_1^2)}{u'(c_1^2-\gamma c_1^1)w_1^2}, \quad \text{(A.19)}$$

which using (2.5) establishes that $MTR_1^2 > 0$.

Using (A.6) and (A.7), we obtain:

$$(1-\phi-\theta_2)\left[u'(c_1^1)-\gamma\delta u'(c_1^2-\gamma c_1^1)\right] = (1-\phi)v'\left(\frac{y_1^1}{w_1^1}\right)\frac{1}{w_1^1}-\theta_2 v'\left(\frac{y_1^1}{w_2^1}\right)\frac{1}{w_2^1}. \quad \text{(A.20)}$$

Because $w_2^1 > w_1^1$ and $v(\cdot)$ is strictly convex:

$$(1-\phi)v'\left(\frac{y_1^1}{w_1^1}\right)\frac{1}{w_1^1}-\theta_2 v'\left(\frac{y_1^1}{w_2^1}\right)\frac{1}{w_2^1} > (1-\phi)v'\left(\frac{y_1^1}{w_1^1}\right)\frac{1}{w_1^1}-\theta_2 v'\left(\frac{y_1^1}{w_1^1}\right)\frac{1}{w_1^1}. \quad \text{(A.21)}$$

Therefore, (A.20) and (A.21) imply that:

$$(1-\phi-\theta_2)\left[u'(c_1^1)-\gamma\delta u'(c_1^2-\gamma c_1^1)\right] > (1-\phi-\theta_2)v'\left(\frac{y_1^1}{w_1^1}\right)\frac{1}{w_1^1}. \quad \text{(A.22)}$$

Hence, (A.22) can be rearranged to yield:

$$1 > \frac{v'(l_1^1)}{\left[u'(c_1^1)-\gamma\delta u'(c_1^2-\gamma c_1^1)\right]w_1^1}, \quad \text{(A.23)}$$

which using (2.4) establishes that $MTR_1^1 > 0$. ■

Proof of Proposition 2

The first-order conditions corresponding to programme (4.1) – (4.2) are:

$$(1-\phi)u'(c_1^2-\gamma c_1^1)-\lambda^2(1-\phi)=0, \tag{A.24}$$

$$-(1-\phi)v'\left(\frac{y_1^2}{w_1^2}\right)\frac{1}{w_1^2}+\lambda^2(1-\phi)=0, \tag{A.25}$$

$$\phi u'(c_2^2-\gamma c_2^1)-\lambda^2\phi=0, \tag{A.26}$$

$$-\phi v'\left(\frac{y_2^2}{w_2^2}\right)\frac{1}{w_2^2}+\lambda^2\phi=0, \tag{A.27}$$

$$(1-\phi)\left[y_1^2-c_1^2\right]+\phi\left[y_2^2-c_2^2\right]=0, \tag{A.28}$$

where $\lambda^2 \geq 0$ is the multiplier on the government's second-period budget constraint (4.2). Dividing (A.25) by (A.24) and rearranging yields:

$$\frac{v'(l_1^2)}{u'(c_1^2-\gamma c_1^1)w_1^2}=1, \tag{A.29}$$

while dividing (A.27) by (A.26) and rearranging yields:

$$\frac{v'(l_2^2)}{u'(c_2^2-\gamma c_2^1)w_2^2}=1, \tag{A.30}$$

which using (2.5) establish that $MTR_1^2=0$ and $MTR_2^2=0$.

The relevant first-order conditions corresponding to programme (4.3) – (4.5) can be written as:

$$(1-\phi-\theta_2^1)u'(c_1^1)+\delta\frac{\partial W^2(\cdot)}{\partial c_1^1}+\theta_2^1\gamma\delta u'(c_1^2-\gamma c_1^1)-\lambda^1(1-\phi)$$
$$+\theta_2^1\delta\left[u'(c_2^2-\gamma c_2^1)\frac{\partial c_2^2(\cdot)}{\partial c_1^1}-v'\left(\frac{y_2^2}{w_2^2}\right)\frac{1}{w_2^2}\frac{\partial y_2^2(\cdot)}{\partial c_1^1}\right]-\theta_2^1\delta\left[u'(c_1^2-\gamma c_1^1)\frac{\partial c_1^2(\cdot)}{\partial c_1^1}-v'\left(\frac{y_1^2}{w_2^2}\right)\frac{1}{w_2^2}\frac{\partial y_1^2(\cdot)}{\partial c_1^1}\right]=0, \tag{A.31}$$

$$-(1-\phi)v'\left(\frac{y_1^1}{w_1^1}\right)\frac{1}{w_1^1}+\lambda^1(1-\phi)+\theta_2^1 v'\left(\frac{y_1^1}{w_2^1}\right)\frac{1}{w_2^1}=0, \tag{A.32}$$

$$(\phi+\theta_2^1)u'(c_2^1)+\delta\frac{\partial W^2(\cdot)}{\partial c_2^1}-\theta_2^1\gamma\delta u'(c_2^2-\gamma c_2^1)-\lambda^1\phi$$

$$+\theta_2^1\delta\left[u'(c_2^2-\gamma c_2^1)\frac{\partial c_2^2(\cdot)}{\partial c_2^1}-v'\left(\frac{y_2^2}{w_2^2}\right)\frac{1}{w_2^2}\frac{\partial y_2^2(\cdot)}{\partial c_2^1}\right]-\theta_2^1\delta\left[u'(c_1^2-\gamma c_1^1)\frac{\partial c_1^2(\cdot)}{\partial c_2^1}-v'\left(\frac{y_1^2}{w_2^2}\right)\frac{1}{w_2^2}\frac{\partial y_1^2(\cdot)}{\partial c_2^1}\right]=0, \tag{A.33}$$

$$-(\phi+\theta_2^1)v'\left(\frac{y_2^1}{w_2^1}\right)\frac{1}{w_2^1}+\lambda^1\phi=0, \tag{A.34}$$

where $\lambda^1 \geq 0$ is the multiplier on the government's first-period budget constraint (4.4), and $\theta_2^1 \geq 0$ is the multiplier on the incentive-compatibility constraint (4.5). To derive expressions for $\partial W^2(\cdot)/\partial c_1^1$ and $\partial W^2(\cdot)/\partial c_2^1$, note that the Lagrangian corresponding to programme (4.1) – (4.2) can be written as:

$$L=(1-\phi)\left[u(c_1^2-\gamma c_1^1)-v\left(\frac{y_1^2}{w_1^2}\right)\right]+\phi\left[u(c_2^2-\gamma c_2^1)-v\left(\frac{y_2^2}{w_2^2}\right)\right]+\lambda^2\left[(1-\phi)\left[y_1^2-c_1^2\right]+\phi\left[y_2^2-c_2^2\right]\right]. \tag{A.35}$$

By the Envelope Theorem:

$$\frac{\partial W^2(\cdot)}{\partial c_1^1}=\frac{\partial L(\cdot)}{\partial c_1^1}=-\gamma(1-\phi)u'(c_1^2-\gamma c_1^1), \tag{A.36}$$

$$\frac{\partial W^2(\cdot)}{\partial c_2^1}=\frac{\partial L(\cdot)}{\partial c_2^1}=-\gamma\phi u'(c_2^2-\gamma c_2^1). \tag{A.37}$$

Substituting (A.37) into (A.33) and combining the result with (A.34) yields:

$$(\phi+\theta_2^1)\left[u'(c_2^1)-\gamma\delta u'(c_2^2-\gamma c_2^1)\right]=(\phi+\theta_2^1)v'\left(\frac{y_2^1}{w_2^1}\right)\frac{1}{w_2^1}$$

$$-\theta_2^1\delta\left[u'(c_2^2-\gamma c_2^1)\frac{\partial c_2^2(\cdot)}{\partial c_2^1}-v'\left(\frac{y_2^2}{w_2^2}\right)\frac{1}{w_2^2}\frac{\partial y_2^2(\cdot)}{\partial c_2^1}\right]+\theta_2^1\delta\left[u'(c_1^2-\gamma c_1^1)\frac{\partial c_1^2(\cdot)}{\partial c_2^1}-v'\left(\frac{y_1^2}{w_2^2}\right)\frac{1}{w_2^2}\frac{\partial y_1^2(\cdot)}{\partial c_2^1}\right]. \tag{A.38}$$

Dividing both sides of (A.38) by $(\phi+\theta_2^1)\left[u'(c_2^1)-\gamma\delta u'(c_2^2-\gamma c_2^1)\right]$ and rearranging yields:

$$1=\frac{v'(l_2^1)}{\left[u'(c_2^1)-\gamma\delta u'(c_2^2-\gamma c_2^1)\right]w_2^1}-\frac{\theta_2^1\delta u'(c_2^2-\gamma c_2^1)}{(\phi+\theta_2^1)\left[u'(c_2^1)-\gamma\delta u'(c_2^2-\gamma c_2^1)\right]}\left[\frac{\partial c_2^2(\cdot)}{\partial c_2^1}-\frac{\partial y_2^2(\cdot)}{\partial c_2^1}\right]$$

$$+\frac{\theta_2^1\delta}{(\phi+\theta_2^1)\left[u'(c_2^1)-\gamma\delta u'(c_2^2-\gamma c_2^1)\right]}\left[u'(c_1^2-\gamma c_1^1)\frac{\partial c_1^2(\cdot)}{\partial c_2^1}-v'\left(\frac{y_1^2}{w_2^2}\right)\frac{1}{w_2^2}\frac{\partial y_1^2(\cdot)}{\partial c_2^1}\right], \tag{A.39}$$

where use has been made of (A.26) and (A.27). Using (2.4), equation (A.39) can be simplified to:

$$MTR_2^1 = \frac{-\theta_2^1 \delta u'(c_2^2 - \gamma c_2^1)}{(\phi + \theta_2^1)\left[u'(c_2^1) - \gamma \delta u'(c_2^2 - \gamma c_2^1)\right]} \left[\frac{\partial c_2^2(\cdot)}{\partial c_2^1} - \frac{\partial y_2^2(\cdot)}{\partial c_2^1}\right]$$
$$+ \frac{\theta_2^1 \delta}{(\phi + \theta_2^1)\left[u'(c_2^1) - \gamma \delta u'(c_2^2 - \gamma c_2^1)\right]} \left[u'(c_1^2 - \gamma c_1^1)\frac{\partial c_1^2(\cdot)}{\partial c_2^1} - v'\left(\frac{y_1^2}{w_2^2}\right)\frac{1}{w_2^2}\frac{\partial y_1^2(\cdot)}{\partial c_2^1}\right]. \quad \text{(A.40)}$$

We now show that $\frac{\partial c_2^2(\cdot)}{\partial c_2^1} - \frac{\partial y_2^2(\cdot)}{\partial c_2^1} > 0$, $\frac{\partial c_1^2(\cdot)}{\partial c_2^1} < 0$ and $\frac{\partial y_1^2(\cdot)}{\partial c_2^1} > 0$ which from (A.40) establishes that $MTR_2^1 < 0$. Application of the Implicit Function Theorem and Cramer's Rule to (A.24) – (A.28) yields:

$$\frac{\partial c_2^2(\cdot)}{\partial c_2^1} = \frac{\gamma \phi^2 u''(c_2^2 - \gamma c_2^1) v''\left(\frac{y_2^2}{w_2^2}\right)\frac{1}{w_2^2 w_2^2}(1-\phi)\left[u''(c_1^2 - \gamma c_1^1) - v''\left(\frac{y_1^2}{w_1^2}\right)\frac{1}{w_1^2 w_1^2}\right]\left[1 - \phi(2-\phi)\right]}{\mid A \mid}$$
$$+ \frac{(1-\phi)^2 u''(c_1^2 - \gamma c_1^1) v''\left(\frac{y_1^2}{w_1^2}\right)\frac{1}{w_1^2 w_1^2}\gamma \phi^3 u''(c_2^2 - \gamma c_2^1)}{\mid A \mid} > 0, \quad \text{(A.41)}$$

$$\frac{\partial y_2^2(\cdot)}{\partial c_2^1} = \frac{(1-\phi)^2 u''(c_1^2 - \gamma c_1^1) v''\left(\frac{y_1^2}{w_1^2}\right)\frac{1}{w_1^2 w_1^2}\gamma \phi^3 u''(c_2^2 - \gamma c_2^1)}{\mid A \mid} > 0, \quad \text{(A.42)}$$

$$\frac{\partial c_1^2(\cdot)}{\partial c_2^1} = \frac{(1-\phi)^2 v''\left(\frac{y_1^2}{w_1^2}\right)\frac{1}{w_1^2 w_1^2} v''\left(\frac{y_2^2}{w_2^2}\right)\frac{1}{w_2^2 w_2^2}\gamma \phi^3 u''(c_2^2 - \gamma c_2^1)}{\mid A \mid} < 0, \quad \text{(A.43)}$$

$$\frac{\partial y_1^2(\cdot)}{\partial c_2^1} = \frac{(1-\phi)^2 u''(c_1^2 - \gamma c_1^1) v''\left(\frac{y_2^2}{w_2^2}\right)\frac{1}{w_2^2 w_2^2}\gamma \phi^3 u''(c_2^2 - \gamma c_2^1)}{\mid A \mid} > 0, \quad \text{(A.44)}$$

where A is the Hessian associated with (A.24) – (A.28):

$$A = \begin{bmatrix} (1-\phi)u''(c_1^2 - \gamma c_1^1) & 0 & 0 & 0 & -(1-\phi) \\ 0 & -(1-\phi)v''\left(\frac{y_1^2}{w_1^2}\right)\frac{1}{w_1^2 w_1^2} & 0 & 0 & 1-\phi \\ 0 & 0 & \phi u''(c_2^2 - \gamma c_2^1) & 0 & -\phi \\ 0 & 0 & 0 & -\phi v''\left(\frac{y_2^2}{w_2^2}\right)\frac{1}{w_2^2 w_2^2} & \phi \\ -(1-\phi) & 1-\phi & -\phi & \phi & 0 \end{bmatrix}, \quad \text{(A.45)}$$

and the determinant of A is given by:

$$\mid A \mid = (1-\phi)^2 u''(c_1^2 - \gamma c_1^1) v''\left(\frac{y_1^2}{w_1^2}\right) \frac{\phi^3}{w_1^2 w_1^2}\left[u''(c_2^2 - \gamma c_2^1) - v''\left(\frac{y_2^2}{w_2^2}\right)\frac{1}{w_2^2 w_2^2}\right]$$
$$+ \phi^2 u''(c_2^2 - \gamma c_2^1) v''\left(\frac{y_2^2}{w_2^2}\right)\frac{1}{w_2^2 w_2^2}(1-\phi)\left[u''(c_1^2 - \gamma c_1^1) - v''\left(\frac{y_1^2}{w_1^2}\right)\frac{1}{w_1^2 w_1^2}\right][1-\phi(2-\phi)] > 0. \tag{A.46}$$

Therefore, from (A.41) and (A.42) it follows that $\frac{\partial c_2^2(\cdot)}{\partial c_2^1} - \frac{\partial y_2^2(\cdot)}{\partial c_2^1} > 0$.

To show that MTR_1^1 is positive, use (A.31) and (A.32) to obtain

$$(1-\phi-\theta_2^1)\left[u'(c_1^1) - \gamma\delta u'(c_1^2 - \gamma c_1^1)\right] = (1-\phi-\theta_2^1) v'\left(\frac{y_1^1}{w_1^1}\right)\frac{1}{w_1^1} + \theta_2^1\left[v'\left(\frac{y_1^1}{w_1^1}\right)\frac{1}{w_1^1} - v'\left(\frac{y_1^1}{w_2^1}\right)\frac{1}{w_2^1}\right]$$
$$-\theta_2^1 \delta u'(c_2^2 - \gamma c_2^1)\left[\frac{\partial c_2^2(\cdot)}{\partial c_1^1} - \frac{\partial y_2^2(\cdot)}{\partial c_1^1}\right] + \theta_2^1\delta\left[u'(c_1^2 - \gamma c_1^1)\frac{\partial c_1^2(\cdot)}{\partial c_1^1} - v'\left(\frac{y_1^2}{w_2^2}\right)\frac{1}{w_2^2}\frac{\partial y_1^2(\cdot)}{\partial c_1^1}\right], \tag{A.47}$$

where use has been made of (A.36), as well as (A.26) and (A.27). Using (2.4), equation (A.47) can be simplified to:

$$MTR_1^1 = \frac{\theta_2^1}{(1-\phi-\theta_2^1)\left[u'(c_1^1) - \gamma\delta u'(c_1^2 - \gamma c_1^1)\right]}\left[v'\left(\frac{y_1^1}{w_1^1}\right)\frac{1}{w_1^1} - v'\left(\frac{y_1^1}{w_2^1}\right)\frac{1}{w_2^1}\right]$$
$$-\frac{\theta_2^1 \delta u'(c_2^2 - \gamma c_2^1)}{(1-\phi-\theta_2^1)\left[u'(c_1^1) - \gamma\delta u'(c_1^2 - \gamma c_1^1)\right]}\left[\frac{\partial c_2^2(\cdot)}{\partial c_1^1} - \frac{\partial y_2^2(\cdot)}{\partial c_1^1}\right]$$
$$+\frac{\theta_2^1\delta}{(1-\phi-\theta_2^1)\left[u'(c_1^1) - \gamma\delta u'(c_1^2 - \gamma c_1^1)\right]}\left[u'(c_1^2 - \gamma c_1^1)\frac{\partial c_1^2(\cdot)}{\partial c_1^1} - v'\left(\frac{y_1^2}{w_2^2}\right)\frac{1}{w_2^2}\frac{\partial y_1^2(\cdot)}{\partial c_1^1}\right]. \tag{A.48}$$

The first term in (A.48) is positive, since $w_2^1 > w_1^1$ and $v(\cdot)$ is strictly convex. Using techniques similar to those used above, it can be shown that $\frac{\partial c_2^2(\cdot)}{\partial c_1^1} < 0$ and $\frac{\partial y_2^2(\cdot)}{\partial c_1^1} > 0$, and therefore the second term in (A.48) is positive. Likewise, it can be shown that $\frac{\partial c_1^2(\cdot)}{\partial c_1^1} > 0$, $\frac{\partial y_1^2(\cdot)}{\partial c_1^1} > 0$, and $\frac{\partial c_1^2(\cdot)}{\partial c_1^1} - \frac{\partial y_1^2(\cdot)}{\partial c_1^1} > 0$, and using equations (A.24) and (A.25) we obtain $u'(c_1^2 - \gamma c_1^1) = v'\left(\frac{y_1^2}{w_1^2}\right)\frac{1}{w_1^2} > v'\left(\frac{y_1^2}{w_2^2}\right)\frac{1}{w_2^2}$. Therefore, the last term in (A.48) is also positive. ■

References

[1] Abel, A [1990], "Asset Prices under Habit Formation and Catching Up with the Joneses", *American Economic Review*, 80, 38-42.

[2] Abel, A [1999], "Risk Premia and Term Premia in General Equilibrium", *Journal of Monetary Economics*, 43, 3-33.

[3] Alonso-Carrera, J., J. Caballe and X. Raurich [2004], "Consumption Externalities, Habit Formation and Equilibrium Efficiency", *Scandinavian Journal of Economics*, 106, 231-251.

[4] Apps, P. and R. Rees [2006], "Repeated Optimal Nonlinear Income Taxation", *mimeo.*

[5] Aronsson, T. and O. Johansson-Stenman [2008a], "When the Joneses' Consumption Hurts: Optimal Public Good Provision and Nonlinear Income Taxation", *Journal of Public Economics*, 92, 986-997.

[6] Aronsson, T. and O. Johansson-Stenman [2008b], "Positional Concerns with Multiple Reference Points: Optimal Income Taxation and Public Goods in an OLG Model", *University of Gothenburg Working Paper*, No. 304.

[7] Auerbach, A [2006], "Budget Windows, Sunsets, and Fiscal Control", *Journal of Public Economics*, 90, 87-100.

[8] Berliant, M. and J. Ledyard [2005], "Optimal Dynamic Nonlinear Income Taxes with No Commitment", *mimeo.*

[9] Brett, C. and J. Weymark [2008], "Optimal Nonlinear Taxation of Income and Savings without Commitment", *Vanderbilt University Working Paper*, 08-W05.

[10] Campbell, J. and J. Cochrane [1999], "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behaviour", *Journal of Political Economy*, 107, 205-251.

[11] Carroll, C [2000], "Solving Consumption Models with Multiplicative Habits", *Economics Letters*, 68, 67-77.

[12] Carroll, C., J. Overland and D. Weil [2000], "Saving and Growth with Habit Formation", *American Economic Review*, 90, 341-355.

[13] Constantinides, G [1990], "Habit Formation: A Resolution of the Equity Premium Puzzle", *Journal of Political Economy*, 98, 519-543.

[14] Dupor, B. and W. Liu [2003], "Jealousy and Equilibrium Overconsumption", *American Economic Review*, 93, 423-428.

[15] Gali, J [1994], "Keeping Up with the Joneses: Consumption Externalities, Portfolio Choice, and Asset Prices", *Journal of Money, Credit, and Banking*, 26, 1-8.

[16] Gaube, T [2007], "Optimum Taxation of Each Year's Income", *Journal of Public Economic Theory*, 9, 127-150.

[17] Guesnerie, R [1995], *A Contribution to the Pure Theory of Taxation*, Cambridge University Press.

[18] Guo, J [2005], "Tax Policy under Keeping Up with the Joneses and Imperfect Competition", *Annals of Economics and Finance*, 6, 25-36.

[19] Ireland, N [2001], "Optimal Income Tax in the Presence of Status Effects", *Journal of Public Economics*, 81, 193-212.

[20] Lettau, M. and H. Uhlig [2000], "Can Habit Formation be Reconciled with Business Cycles Facts?", *Review of Economic Dynamics*, 3, 79-99.

[21] Ljungqvist, L. and H. Uhlig [2000], "Tax Policy and Aggregate Demand Management under Catching Up with the Joneses", *American Economic Review*, 90, 356-366.

[22] Oswald, A [1983], "Altruism, Jealousy and the Theory of Optimal Non-Linear Taxation", *Journal of Public Economics*, 20, 77-87.

[23] Stiglitz, J [1982], "Self-Selection and Pareto Efficient Taxation", *Journal of Public Economics*, 17, 213-240.