Habit formation and labor supply[1]

Helmuth Cremer[2], Philippe De Donder[3],
Dario Maldonado[4] and Pierre Pestieau[5]

This version, May 2010

[1] We thank the editor and the referee for their detailed and constructive comments. Financial support from ANR "Retraite" (ANR-05-BLAN-0020) is gratefully acknowledged.

[2] Toulouse School of Economics (GREMAQ, IDEI and Institut universitaire de France).

[3] Toulouse School of Economics (GREMAQ-CNRS and IDEI).

[4] Universidad del Rosario (Department of Economics and CeiBA - Complejidad), Bogota.

[5] CREPP, HEC-Management School University of Liège; CORE, Université catholique de Louvain; PSE and CEPR.

Abstract

This paper analyzes the pattern of consumption taxes in a two period model with habit formation and myopia. Individuals' second period needs increase with their first period consumption. However, myopic individuals do not see this habit formation relation when they take their saving decision. The first-best solution is decentralized by a simple "Pigouvian" (paternalistic) consumption tax (along with suitable lump-sum taxes). In a second-best setting, when personalized lump-sum transfers are not available, consumption taxes may have conflicting paternalistic and redistributive effects. Taxes should discourage consumption of goods that entail negative externalities (unforeseen habits), but discourage less the consumption of goods that are proportionately consumed by individuals with high net social marginal utility of income. We also show that both myopic and farsighted individuals benefit more from the second-best policy when the proportion of myopic agents in society increases.

JEL codes: D91, H21, H55

Keywords: Habit formation, myopia, projection bias, postponing retirement.

1 Introduction

This paper studies the effect of (rational or myopic) habit formation on labor supply and particularly on the retirement decision. It also examines how this phenomenon affects the design of commodity taxation. Empirical macroeconomists studying time series models of aggregate consumption expenditure have long been puzzled by the significant role that lagged consumption plays in explaining current consumption, irrespective of the other variables (such as income, assets and interest rates) that enter the estimated equations. As early as 1952, T. M. Brown obtained impressive empirical results and suggested habit persistence as the explanation for his findings. Later on, Muellbauer (1988) has studied how to add habit persistence to the life-cycle model. He has distinguished two versions of habit formation, rational and myopic: in the former, consumers are aware of the relationship between current consumption and future tastes, while in the latter they are not and are "constantly being surprised even if they forecast their budget constraints accurately" (p. 50). Muellbauer (1988) studies the implications of these two versions of habit formation for the Euler equation and applies his model to US data. The evidence he obtains favors the hypothesis of myopic habit formation.

There is surprisingly little literature on the implications of (myopic or rational) habit formation for optimal public policies such as taxation. As a matter of fact, this theme of habit formation has been rediscovered quite recently by the literature on happiness. The puzzle emerging from these studies is that, although richer people tend to report being happier than poorer people in most countries, inter-country comparisons do not show that increases in aggregate income are associated with higher aggregate happiness. Looking at micro evidence based on experiments, authors such as Layard (2005) conclude that two phenomena are at play: what he calls rivalry (in assessing how happy they are, people make relative comparisons with their peers—this would explain why richer people feel happier than poorer people in any given country) and habit (people get accustomed to any consumption level, which would explain why on average richer countries do not seem happier than poorer countries). Layard (2005) goes further than his predecessors by challenging the profession to analyze the consequences of these two phenomena on optimal public policies, as the title of his contribution makes clear: "Rethinking public economics: the implications of rivalry and habit". He points out, among other things, that distortions arise with habit formation only if the "habituation effect is not foreseen" (p.8). There is other recent evidence for myopic habit formation, such as Fehr and Sych (2008) and Woittiez and Kapteyn (1998).

Our contribution relies on two concepts, habit formation and myopia, that are introduced in a two period model.[1] Individual labor supply is fixed and unitary in the first period; in the second one, it is endogenous and can be viewed as the retirement age. In the first period, individuals consume a certain fraction of their earnings. This provides some utility but creates future needs or habits in the second period. However, we assume that out of myopia or ignorance, individuals underestimate the extent of this habit formation. Consequently, when they reach the second period, they face unexpected consumption needs along with insufficient saving, which may force them to work longer than expected. This is in line with a recent paper by Maestas (2007) which finds that one out of five U.S. retirees unexpectedly return to work.

With habit formation and myopia, there is a case for government intervention since myopic agents come to regret their past decisions.[2] With identical individuals, it suffices for the government to induce more saving or to tax first period consumption. With individuals differing in earnings, and in the absence of lump sum transfers, our linear tax instruments play two roles: correction for myopia and redistribution. Taxes should discourage consumption of goods that entail negative externalities that are not internalized (unforeseen habits[3]) but discourage less the consumption of goods that are proportionately consumed by individuals with high net social marginal utility of income. We also show that both myopic and farsighted individuals benefit more from the second-best policy when the proportion of myopic agents in society increases.

Our paper is related to the literature on present-biased preferences, such as O'Donoghue and Rabin (2003 and 2006), Gruber and Koszegi (2001 and 2004), Diamond and Koszegi (2003) and Cremer *et al.* (2009). These papers also assume that individuals use time-inconsistent preferences and study how these inconsistencies can be mitigated through tax instruments. The main difference with our paper is that they assume that individuals do not care (or care too little) about all components of their future utility while we assume that

[1]The paper closest to ours is Diamond and Mirrlees (2000). Our approach differs from theirs in two ways: they focus on saving and not on labor supply and they do not examine the tax policy implications of habit formation.

[2]We measure aggregate welfare as the sum of individual's "long term" preferences —i.e., the preferences that individuals use in the second period of the game. Observe that since individuals regret their myopia when old, they are ultimately thankful to the planner for having influenced their choices. This is line with the "cautious" (O'Donoghue and Rabin, 1999) or "asymmetric" (Camerer *et al.*, 2003) paternalism advocated in the literature. Note that our approach assumes that the true parameter relating past and future consumption is known to the government. A different analysis would emerge if this parameter could only be inferred from observation of individual choices (as in Bernheim and Rangel, 2009).

[3]The psychology literature dubs this inability to forecast future preferences "projection bias".

individuals do not correctly forecast the effects of their current consumption on their future utility.

The rest of the paper is organized as follows. We present the basic model along with the laissez-faire and the first-best solutions (including decentralization) in Section 2. Section 3 is devoted to the second-best analysis, assuming that all individuals are myopic. Section 4 presents simulations of our model where society is composed of both myopic and farsighted individuals.

2 First-Best and Decentralization

2.1 The model

The economy we consider consists of a continuum of individuals differing in their wage or productivity w. Each individual works one unit of time in the first period of his life and earns w. This income is divided into current consumption, c, and saving, s. In the second period, he works an amount of time ℓ, and earns $w\ell$.[4] We assume without loss of generality a zero interest rate, so that total second period income is equal to $w\ell + s$ and is devoted to second period consumption, d.

Individual utility is given by

$$U(c, d, \ell) = u(c) + v(d, c) - h(\ell),$$

where $v(d, c)$ is the utility from second period consumption that depends on first period consumption as well. We assume that u is strictly concave and h strictly convex, while $v(d, c)$ is strictly concave and increasing in d. Our habit formation assumption implies that $v_c < 0$ and $v_{dc} > 0$, so that previous period consumption reduces second period's utility and generates additional needs.[5]

Myopia is represented by the fact that in the first period of their life, individuals do not see this delayed effect of consumption. Consequently, when they choose saving and their projected retirement age they use a different expression for the second period utility, namely $v(d, 0)$. A farsighted individual would have a correct perception of such a habit formation, that is $v(d, c)$.[6]

[4] As is standard in the optimal taxation literature, we consider implicitly a linear production function with labor as its sole input. Individuals differ in the productivity of the labor they supply.

[5] An alternative specification is to assume $v_c > 0$, which implies that previous consumption brings status and hence additional utility.

[6] By restricting our setting to two periods, we reduce the self-control problem to its simplest expression. This is the price to pay for simplicity. Further, we assume that our

For the sake of simplicity, we assume that v is given by

$$v(d,c) = u\left(d - \alpha c\right),$$

where $\alpha = 0$ for myopic individuals in the first period of their life, while $\alpha = \bar{\alpha}$ is the true value of the parameter (used by myopic individuals in the second period, and by farsighted in both periods).[7]

In the remainder of this section, we first solve the individuals' optimization program in the absence of government intervention, and compare a myopic individual's consumption, saving and labor supply choices with the choices he would have made had he been farsighted. We then introduce government intervention and determine the first-best solution. We then analyze how this first-best allocation can be decentralized using consumption taxes together with individualized lump sum transfers.

2.2 Laissez-faire allocation

We first study the impact of myopic behavior on consumption and retirement decisions in the absence of government intervention, starting with the first period's optimization. Although the only decision taken by individuals in the first period is how much of their exogenous income to consume and to save, to solve this problem they must anticipate their second period's labor supply decision and their future marginal utility from consumption. With our formulation and using the budget constraint, the first period optimization problem of an individual with wage w can then be written as

$$\max_{s,\ell} u(w-s) + u(w\ell - \alpha w + (1+\alpha)s) - h(\ell),$$

where a myopic individual mistakenly uses $\alpha = 0$ while a farsighted would use the correct value of $\alpha = \bar{\alpha}$. The FOCs are given by

$$[s^e] \quad : \quad -u'(w-s^e) + (1+\alpha)u'(w\ell^p - \alpha w + (1+\alpha)s^e) = 0 \qquad (1)$$

$$[l^p] \quad : \quad wu'(w\ell^p - \alpha w + (1+\alpha)s^e) - h'(\ell^p) = 0. \qquad (2)$$

myopic agents are partially sophisticated as they acknowledge their perception error in the second period of their life.

[7]For simplicity, we assume that individuals are either farsighted (use the correct value of α in both periods) or fully myopic (use $\alpha = 0$ in the first period), but our model could accommodate any intermediate degree of myopia (i.e., individuals using any value of α such that $0 \leq \alpha < \bar{\alpha}$ in the first period). The analysis up to (and including) section 2.3 remains indeed unchanged with people differing in their degree of myopia as well as in their productivity. Also, we could generalize the model to a setting where habit formation differs from one individual to another (i.e., agents differ also in their value of $\bar{\alpha}$), but this would complexify the model without adding significantly to the intuition of the results.

The solution to these equations gives the *effective* value of saving s^e (and thus of consumption c^e) in the first period and the *planned* second period labor supply ℓ^p (and thus also the planned second period consumption d^p).

The appendix shows that saving increases with α, so that a myopic individual saves less and consumes more in the first period than a farsighted individual of the same wage w. This is intuitive, since the myopic individual under-estimates the needs that first period consumption creates later on. However, the sign of the derivative of ℓ^p with respect to α is ambiguous.[8] On the one hand, for any given first period consumption/saving level, the myopic individual under-estimates his second period marginal utility and thus plans to supply less labor than the farsighted individual. On the other hand, the myopic individual saves less than the far-sighted individual, which lowers his second period consumption for any given labor supply and drives him to plan to supply more labor than the farsighted.

Formally, solving simultaneously the FOCs (1) and (2) and denoting by c^e the effective first period consumption, we obtain that

$$\frac{u'(c^e)}{1+\alpha}w = h'(\ell^p). \tag{3}$$

For a given ℓ^p, the numerator of the left-hand side of (3) increases with α (as $c^e = w - s^e$ decreases with α) while the denominator is also increasing in α, so that it is not possible to sign the derivative of ℓ^p with respect to α.

In the second period, individuals choose their (effective) labor supply by solving

$$\max_{\ell} u\left[w\ell - \overline{\alpha}w + (1+\overline{\alpha})s^e\right] - h(\ell),$$

which yields the following first-order condition:

$$wu'\left[w\ell^e - \overline{\alpha}w + (1+\overline{\alpha})s^e\right] = h'(\ell^e). \tag{4}$$

The solution to this problem gives the effective second period consumption d^e and labor supply ℓ^e.

Observe that condition (4) is identical to (2) for farsighted individuals: their realized and planned labor supplies are identical. Myopic agents, however, have a different behavior. They realize in the second period that the true value of α is $\overline{\alpha}$ and that they did not save enough in the first period.

From (4), it is easy to see that the optimal value of ℓ decreases with s. As we know that a myopic individual saves less than a farsighted one, we obtain

[8] If myopia would concern not only habit formation but the whole perception of the future (instead of $u(d,\alpha) - h(\ell)$ we would have $\alpha(u(d,c) - h(\ell))$), this would be different.

that the effective labor supply of a myopic individual is larger than the one of a farsighted individual. The intuition for this result is that a myopic individual under-estimates his second period needs and does not save enough in the first period. Consequently, he has to work more than planned, and also more than a farsighted individual. In other words, myopia leads to prolonged activity.[9]

The comparison of second period consumption between a myopic and a farsighted individual is also ambiguous. On the one hand, myopic agents save less, which drives them to consume less in second period. On the other hand, we have just shown that they work more than farsighted individuals, which leads them to consume more. The net impact of myopia on second period consumption depends upon the relative concavity of the utility function and of the utility derived from leisure. Expressing (4) in terms of second period consumption, we obtain that

$$wu'\left[d^e - \alpha w + \alpha s^e\right] - h'(\ell^e) = 0. \tag{5}$$

Recall that the only difference between a myopic and a farsighted agent in the second period is their amount of saving s^e. Applying the implicit function theorem to (5), and using the fact that $\ell^e = (d^e - s^e)/w$, we obtain that

$$\text{sign}(\frac{\partial d^e}{\partial \alpha}) = \text{sign}(\alpha w u''(d^e - \alpha w + \alpha s^e) + h''(\ell^e)),$$

where the sign of the first term in the right hand side is negative while that of the second term is positive. If $u''(.)$ is very large (in absolute term) compared to $h''(.)$, myopic agents work a lot more than the farsighted and end up consuming more in the second period. If $h''(.)$ is very large, the labor supply difference between myopic and farsighted agents is small enough that myopia leads to lower consumption levels in second period for the myopic.

We summarize the comparison of laissez-faire choices by myopic and by farsighted individuals earning the same wage rate w in the following Proposition.

Proposition 1 *In a laissez-faire situation, the comparison of the optimal choices of myopic and farsighted agents earning the same wage w runs as follows:*

[9] In a model with more than two periods one would have the additional issue of learning about habit formation. As time passes, individuals are forced to revise the age at which they plan to retire, so they may end up learning that they form habits and start to introduce this issue in their optimization problems. Studying the effect of learning about habit formation is out of the scope of this paper. Moreover, learning would only affect our qualitative results if it is very fast. If learning is not too fast, the qualitative implications of our model that individuals under save and end up retiring later than optimal will continue to hold.

(i) Myopic agents save less and consume more in the first period;

(ii) The myopic's planned labor supply may be larger or smaller than the (planned and effective) labor supply of the farsighted. It is larger when the marginal utility curve is close to flat;

(iii) The effective labor supply of a myopic is larger than the farsighted's;

(iv) Myopic agents consume more than farsighted agents in the second period if the concavity of the utility function is large compared to the convexity of the disutility from working, and vice versa.

2.3 First-best

The previous section has shown that myopic agents make choices that differ from the ones they would have made had they been farsighted. These different choices are due to mistakes made in the first period optimization, when they underestimate the habit formation process. They thus come to regret their first period decisions when they enter the second period and observe the true extent of habit formation. Based on this regret, there is a paternalistic case for government intervention. In this section, we assume that the social planner observes the productivity of each individual and their degree of myopia, but imposes to all the preferences of the farsighted individuals in both periods. Since all individuals share by assumption the same habit formation technology (i.e., they all exhibit the same $\bar{\alpha}$), and since we look at the first-best optimal allocation, this allocation does not depend on the individual's degree of myopia: individuals differing only in their degree of myopia will receive the same first-best allocation from the planner. The allocation received will typically depend on the individual's productivity, so we need to know how wages are distributed. We take w to be continuously distributed on $[w^-, w^+]$ according to $F(w)$.

We denote by $c(w)$, $d(w)$ and $\ell(w)$, respectively, the first period consumption, second period consumption and labor supply of an individual of productivity w. The social planner's problem consists in finding the schedules $c(w)$, $d(w)$ and $\ell(w)$ that maximize the integral of the individuals' utility (using $\alpha = \bar{\alpha}$ for all agents) under an aggregate budget constraint:

$$\max_{c(w),d(w),\ell(w)} \quad \int_{w^-}^{w^+} \{u(c(w)) + u(d(w) - \overline{\alpha}c(w)) - h(\ell(w))\} dF(w) + \lambda \int_{w^-}^{w^+} (w + w\ell(w) - c(w) - d(w))\, dF(w),$$

where λ denotes the Lagrange multiplier of the aggregate budget constraint.

This optimization program leads to the FOCs:

$$\begin{aligned}
[c(w)] &\ :\ u'(c^*(w)) - \overline{\alpha} u'(d^*(w) - \overline{\alpha} c^*(w)) = \lambda, && (6)\\
[d(w)] &\ :\ u'(d^*(w) - \overline{\alpha} c^*(w)) = \lambda, && (7)\\
[\ell(w)] &\ :\ h'(\ell^*(w)) = \lambda w && (8)
\end{aligned}$$

where the asterisk (*) denotes the optimal first-best consumption and labor supply schedules.

We then obtain the following proposition:

Proposition 2 *The first-best allocation is such that:*
(i) The planner equalizes marginal utility of consumption across consumption goods c and d for all individuals. The marginal utility of first period consumption c takes into account its full impact, including the one materializing in the second period;
(ii) The consumption (in the same period) of individuals of different productivities is equalized: $c^*(w) = c^*$ and $d^*(w) = d^*$;
(iii) Habit formation implies that consumption in the second period is higher than in the first period;
(iv) Labor supply increases with productivity: more productive individuals retire later than less productive ones.
Proof. *Statements (i) and (ii) follow directly from (6) and (7) while statement (iv) comes from (8). As to (iii), we obtain from (6) and (7) that*

$$u'(c^*(w)) = (1 + \overline{\alpha}) u'(d^*(w) - \overline{\alpha} c^*(w)),$$

from which we conclude that $u'(c^*(w)) > u'(d^*(w) - \overline{\alpha} c^*(w))$, so that $c^*(w) < d^*(w) - \overline{\alpha} c^*(w)$. This in turn means that $c^*(w) < d^*(w)$ as claimed. ■

2.4 Decentralization of the First-Best allocation

We now turn to the decentralization of the first-best allocation. Decentralization consists in designing taxes and transfers that will induce individuals to take the socially optimal consumption and labor supply decisions. As individual choices depend on the individuals' degree of myopia, we have to make an assumption on the distribution of myopia in the economy. We take the simplest framework and we assume from now on that all individuals are fully myopic (i.e., use $\alpha = 0$ in their first period choices).[10] We also assume in this

[10] This assumption is made for simplicity reasons. The assumption of a nul α is minor; what is less minor is to assume that everyone has the same myopia. Ideally, we would like

section, as is traditional, that the planner can observe individuals' incomes. Finally, we assume that the planner controls three instruments: a tax on first period consumption, τ_c, a tax on second period consumption, τ_d and a lump sum transfer $T(w)$ that, for the time being, may depend on wage.

Faced with these taxes and transfers, the first period problem of a myopic individual of wage w is to maximize:

$$u(c) + u(d) - h(\ell) - \mu_1 \left[c(1+\tau_c) + d^p(1+\tau_d) - w(1+\ell) - T(w) \right],$$

where μ_1 is the Lagrange multiplier associated with his budget constraint. Consequently, effective first period consumption c^e and planned second period consumption d^p and labor supply ℓ^p must satisfy

$$[c^e] \;:\; u'(c^e) = \mu_1 (1+\tau_c), \tag{9}$$
$$[d^p] \;:\; u'(d^p) = \mu_1 (1+\tau_d), \tag{10}$$
$$[\ell^p] \;:\; h'(\ell^p) = \mu_1 w. \tag{11}$$

In the second period, the individual's problem is to maximize

$$u(d - \bar{\alpha} c^e) - h(\ell) - \mu_2 \left[d(1+\tau_d) - s^e - T(w) - w\ell \right],$$

where μ_2 is the Lagrange multiplier (which is different from μ_1) associated with the individual's second period budget constraint. The effective second period consumption and labor supply must satisfy the FOCs which are given by

$$[d^e] \;:\; u'(d^e - \bar{\alpha} c^e) = \mu_2 (1+\tau_d), \tag{12}$$
$$[\ell^e] \;:\; h'(\ell^e) = \mu_2 w. \tag{13}$$

To achieve the first-best, the planner needs to induce the (myopic) individuals to save the appropriate amount. From there on, the choice of retirement age will be optimal. To obtain the "right" (first-best) level of saving, we combine (6), (7),(9) and (10), keeping in mind from Proposition 1*(iii)* that $c^*(w) = c^*$ and that $d^*(w) = d^*$, to obtain

$$\frac{1+\tau_c}{1+\tau_d} = \frac{(1+\bar{\alpha})\, u'(d^* - \bar{\alpha} c^*)}{u'(d^*)}.$$

to consider a society in which people differ in both productivity and myopia with some correlation between the two. But this would make the analysis much more complicated than it already is. In Section 4, we study numerically the case where a fraction λ of society is composed of myopic individuals.

Proposition 3 *In order to implement the first-best allocation, the planner needs only a tax on the first period consumption supplemented by the lump sum tax transfers $T(w)$. More specifically, one needs*

$$\tau_c = \frac{(1+\bar{\alpha})\, u'\left(d^* - \bar{\alpha} c^*\right) - u'\left(d^*\right)}{u'\left(d^*\right)} > 0 \qquad \text{with } \tau_d = 0. \tag{14}$$

In words, to decentralize the first-best solution, we need a Pigouvian tax on first period consumption. With heterogenous individuals, decentralization also calls for individualized transfers $T(w)$ in order to equalize marginal utilities across agents. Alternatively, one can also subsidize second period consumption to decentralize the first-best, but in that case one also must subsidize labor supply to leave the second period consumption/labor trade-off undistorted.

Finally, observe that, if society were composed of both myopic and farsighted individuals, and if the type of individuals were observable, the planner would not want to impose consumption taxes on farsighted agents but only on myopic individuals. Decentralization in a mixed society would then imply the same consumption taxes as those discussed in this section for myopic individuals together lump-sum transfers for both types of individuals. We treat numerically the case where individual myopia is not observable by the government in Section 4.

In the next Section, we study the case where the planner cannot observe the productivity level of individuals in a society made exclusively of myopic agents.

3 Second-best

As we have just seen, the planner can achieve the first-best optimum with tax τ_c and individualized transfers $T(w)$. This requires that the planner is able to observe an individual's type (wage and first period income) and to condition individual lump sum transfers upon it. This is an especially strong assumption, which we now lift in order to move to the second-best realm where we assume that the lump sum transfer is the same for all individuals and is denoted by T. Furthermore, we restrict our attention to linear tax instruments on expenditures c and d. The government budget constraint then implies[11]

$$T = \tau_c Ec + \tau_d Ed.$$

In such a setting, we expect that the two taxes will play two roles: a corrective Pigouvian role (positive for τ_c, negative for τ_d) and a redistributive role (the taxes are used to finance the lump sum transfer). Given that both c and d are normal goods (increasing with w), such taxes are indeed redistributive.[12]

We continue considering a society where all individuals are equally myopic; we leave the case of a society where myopic and farsighted individuals coexist to Section 4. In the first period, an individual with productivity w maximizes:

$$u(c) + u\left(\frac{1}{1+\tau_d}(w + w\ell + T) - \left(\frac{1+\tau_c}{1+\tau_d}\right)c\right) - h(\ell). \tag{15}$$

This yields the effective level of c^e and a planned value of second period labor supply ℓ^p. Both are functions of tax instruments and yield a planned value for d, d^p. *Ex post*, given c^e, they maximize

$$u\left(\frac{1}{1+\tau_d}(w + w\ell + T) - \left(\frac{1+\tau_c}{1+\tau_d} + \bar{\alpha}\right)c^e\right) - h(\ell), \tag{16}$$

to determine the effective levels of second period consumption and labor supply d^e and ℓ^e which are different from the planned ones d^p and ℓ^p.

As above the social planner chooses the tax instruments τ_c, τ_d and T with the *preferences* of (hypothetical) farsighted individuals in his objective function but taking into account the actual demand and supply functions of myopic agents. The Lagrangian expression associated with the problem of the social planner is thus given by

$$\begin{aligned} \mathcal{L} \;=\; & E\left\{u(c^e) + u\left(\frac{1}{1+\tau_d}(w + w\ell^e + T) - \left(\frac{1+\tau_c}{1+\tau_d} + \bar{\alpha}\right)c^e\right) - h(\ell^e)\right\} \\ & -\lambda E\left(T - \tau_c c^e - \tau_d d^e\right), \end{aligned}$$

[11]From now on, E denotes the expectation operator. For any expression x, we have

$$E(x) = \int_{w^-}^{w^+} x(w)dF(w).$$

[12]The fact that we do not tax wage income is without loss of generality. From the individual's budget constraint, a linear income tax is equivalent to a uniform tax on expenditures. We can thus normalize the tax on income to zero and put all the taxes on the expenditure side. This argument also shows that the endogeneity of labor supply is a crucial ingredient to our second-best analysis. When ℓ is constant, a tax on income has not deadweight loss and consequently the taxes on c and d do not produce any distortions.

where λ is the multiplier associated with the aggregate revenue constraint and c^e, d^e and ℓ^e now represent the individual choices given the policy instruments.

We first focus on the choice of T. The FOC is given by[13]

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial T} &= E\left\{\left[u'\left(c^e\right)-u'\left(d^e-\bar{\alpha}c^e\right)\left(\frac{1+\tau_c}{1+\tau_d}+\bar{\alpha}\right)\right]\frac{\partial c^e}{\partial T}+u'\left(d^e-\bar{\alpha}c^e\right)\frac{1}{1+\tau_d}\right.\\ &\quad\left.-\lambda\left(1-\tau_c\frac{\partial c^e}{\partial T}-\tau_d\frac{\partial d^e}{\partial T}\right)\right\}=0.\end{aligned}$$

Using (9) and (10) we obtain the following proposition.

Proposition 4 *In a second-best setting where the planner awards the same lump sum transfer to all agents irrespective of their income, the optimal level of this transfer is given by*

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial T} &= E\left[\frac{1}{1+\tau_d}u'\left(d^e-\bar{\alpha}c^e\right)-\Delta\frac{\partial c^e}{\partial T}-\lambda\left(1-\tau_c\frac{\partial c^e}{\partial T}-\tau_d\frac{\partial d^e}{\partial T}\right)\right],\\ &= \lambda E\left(b-1\right)=0.\end{aligned}$$

where

$$\Delta \equiv \left(\frac{1+\tau_c}{1+\tau_d}+\bar{\alpha}\right)u'\left(d^e-\bar{\alpha}c^e\right)-\frac{1+\tau_c}{1+\tau_d}u'\left(d^p\right)>0,$$

and

$$b=\frac{1}{\lambda}\left(\frac{1}{1+\tau_d}u'\left(d^e-\bar{\alpha}c^e\right)-\Delta\frac{\partial c^e}{\partial T}\right)+\tau_c\frac{\partial c^e}{\partial T}+\tau_d\frac{\partial d^e}{\partial T}.$$

The term Δ reflects the cost of myopia in terms of ex post utility. It tends to 0 when α is not equal to 0 but tends to $\bar{\alpha}$. The term b is quite standard in the linear taxation literature; it is what Atkinson and Stiglitz (1980) call the net social marginal valuation of income. It is measured in terms of government revenue. It is net in the sense that the effect of a lump sum transfer includes the direct effect on individual utility but also the indirect effect on tax revenue.

Using this notation, we can get the two other FOCs:

$$\frac{\partial \mathcal{L}}{\partial \tau_c}=E\left[\frac{-c^e}{1+\tau_d}u'\left(d^e-\bar{\alpha}c^e\right)-\Delta\frac{\partial c^e}{\partial \tau_d}+\lambda\left(c^e+\tau_c\frac{\partial c^e}{\partial \tau_c}+\tau_d\frac{\partial d^e}{\partial \tau_c}\right)\right]=0,$$

and

$$\frac{\partial \mathcal{L}}{\partial \tau_d}=E\left[\frac{-d^e}{1+\tau_d}u'\left(d^e-\bar{\alpha}c^e\right)-\Delta\frac{\partial c^e}{\partial \tau_d}+\lambda\left(d^e+\tau_c\frac{\partial c^e}{\partial \tau_d}+\tau_d\frac{\partial d^e}{\partial \tau_d}\right)\right]=0.$$

[13] Note that the FOCs do not involve derivatives of ℓ because for ℓ the envelope theorem applies. As a consequence of individuals' myopia, we cannot use the envelope theorem for the choice of first period consumption.

Replacing all Marshallian price effects by their equivalent decomposition in Hicksian price effects and income effects (the Slutsky equation) these FOCs can be written as

$$\begin{aligned} -\operatorname{cov}(b, c^e) - E\left[\frac{\Delta}{\lambda}\frac{\partial \tilde{c}^e}{\partial \tau_c} - \tau_c \frac{\partial \tilde{c}^e}{\partial \tau_c} - \tau_d \frac{\partial \tilde{d}^e}{\partial \tau_c}\right] &= 0, \\ -\operatorname{cov}(b, d^e) - E\left[\frac{\Delta}{\lambda}\frac{\partial \tilde{c}^e}{\partial \tau_d} - \tau_c \frac{\partial \tilde{c}^e}{\partial \tau_d} - \tau_d \frac{\partial \tilde{d}^e}{\partial \tau_d}\right] &= 0, \end{aligned}$$

where $\widetilde{c}^e$ and $\widetilde{d}^e$ are the compensated demand functions. To get more intuition, let us first consider the case when only one of the consumption taxes (either τ_c or τ_d) is available.

Proposition 5 *In a second-best setting where the planner uses a single tax on consumption together with a lump sum transfer, the optimal value of the consumption tax is given*

$$\tau_c = \frac{\operatorname{cov}(b, c^e)}{E\frac{\partial \tilde{c}^e}{\partial \tau_c}} + \frac{E\frac{\Delta}{\lambda}\frac{\partial \tilde{c}^e}{\partial \tau_c}}{E\frac{\partial \tilde{c}^e}{\partial \tau_c}} \qquad \textit{when } \tau_d = 0, \tag{17}$$

or

$$\tau_d = \frac{\operatorname{cov}(b, d^e)}{E\frac{\partial \tilde{d}^e}{\partial \tau_d}} + \frac{E\frac{\Delta}{\lambda}\frac{\partial \tilde{c}^e}{\partial \tau_d}}{E\frac{\partial \tilde{d}^e}{\partial \tau_d}} \qquad \textit{when } \tau_c = 0. \tag{18}$$

First of all, if $\Delta = 0$, namely if there is no myopia, we only have the first part of these formulas, that is standard in optimal consumption tax with heterogenous individuals. The numerator reflects the redistributive objective; it is negative as the covariance between consumption and the marginal utility of income is negative. This term would be zero with identical individuals or without concern for redistribution (linear utility). The denominator is also negative and reflects the efficiency effect (deadweight loss). The tax and thus redistribution will be larger if the (compensated) demand for c or d is inelastic.[14]

Note that if the first part of (17) were equal to zero, we would end up with an expression very similar to (14). That is:

$$\tau_c = \frac{\Delta}{\lambda} > 0.$$

[14] As usual, the expressions we obtain for the optimal taxes are *rules* rather than the optimal *levels*. This is standard in the optimal taxation literature.

Similarly, if the first part of (18) were equal to zero the optimal tax would be

$$\tau_d = \frac{\Delta}{\lambda} \frac{\partial \tilde{c}^e}{\partial \tau_d} \bigg/ \frac{\partial \tilde{d}^e}{\partial \tau_d} < 0.$$

If we assume that c and d have the same redistributive pattern (same covariance between marginal utility of income and consumption) and the same price elasticity, two reasonable assumptions, one can state that $\tau_d < \tau_c$ if the two taxes are used alone.

Let us now turn to the case when the two taxes are used together.

Proposition 6 *In a second-best setting where the planner taxes both first period and second period consumption together with a lump sum transfer, the optimal value of the consumption taxes is given by:*

$$\tau_c = \frac{\operatorname{cov}(b, c^e)\, E\frac{\partial \tilde{d}^e}{\partial \tau_d} - \operatorname{cov}(b, d^e)\, E\frac{\partial \tilde{d}^e}{\partial \tau_c}}{E\frac{\partial \tilde{c}^e}{\partial \tau_c}\ E\frac{\partial \tilde{d}^e}{\partial \tau_d} - E\frac{\partial \tilde{c}^e}{\partial \tau_d}\ E\frac{\partial \tilde{d}^e}{\partial \tau_c}} + \frac{E\frac{\Delta}{\lambda}\left[\frac{\partial \tilde{c}^e}{\partial \tau_c} E\frac{\partial \tilde{d}^e}{\partial \tau_d} - \frac{\partial \tilde{c}^e}{\partial \tau_d} E\frac{\partial \tilde{d}^e}{\partial \tau_c}\right]}{E\frac{\partial \tilde{c}^e}{\partial \tau_c}\ E\frac{\partial \tilde{d}^e}{\partial \tau_d} - E\frac{\partial \tilde{c}^e}{\partial \tau_d}\ E\frac{\partial \tilde{d}^e}{\partial \tau_c}}, \quad (19)$$

and

$$\tau_d = \frac{\operatorname{cov}(b, d^e)\, E\frac{\partial \tilde{c}^e}{\partial \tau_c} - \operatorname{cov}(b, c^e)\, E\frac{\partial \tilde{c}^e}{\partial \tau_d}}{E\frac{\partial \tilde{c}^e}{\partial \tau_c}\ E\frac{\partial \tilde{d}^e}{\partial \tau_d} - E\frac{\partial \tilde{c}^e}{\partial \tau_d}\ E\frac{\partial \tilde{d}^e}{\partial \tau_c}} - \frac{E\frac{\Delta}{\lambda}\left[\frac{\partial \tilde{c}^e}{\partial \tau_c} E\frac{\partial \tilde{c}^e}{\partial \tau_d} - \frac{\partial \tilde{c}^e}{\partial \tau_d} E\frac{\partial \tilde{c}^e}{\partial \tau_c}\right]}{E\frac{\partial \tilde{c}^e}{\partial \tau_c}\ E\frac{\partial \tilde{d}^e}{\partial \tau_d} - E\frac{\partial \tilde{c}^e}{\partial \tau_d}\ E\frac{\partial \tilde{d}^e}{\partial \tau_c}}. \quad (20)$$

These are very complex formulas. Note that in the case where cross derivatives $\partial \tilde{d}^e / \partial \tau_c$ and $\partial \tilde{c}^e / \partial \tau_d$ are negligible, we return to expressions (17) and (18). In other words, what makes these formulas (19) and (20) different is the different cross effects.

Note that the expressions for τ_c and τ_d have the same denominator which measures the inefficiencies introduced by the tax system. This denominator is equal to the determinant of the Slutsky matrix and consequently we can expect it to be positive. Focusing on the numerators, they include a positive equity effect (the numerator of the first fraction) and a corrective effect (the numerator of the second fraction). Both effects are intuitive and very similar to what happens when only one of the taxes is present. The first effect (positive for both τ_c and τ_d) is related to equity since the covariances measure inequality in consumption. With cross-price effects, we have to take into account covariances between marginal utility of income and consumption in both periods for both taxes. The second effect is related to myopia since it is proportional to Δ. The presence of cross-price effects makes it difficult to sign this term.

An interesting case emerges when the form of the utility function implies demand functions which exhibit multiplicative separability.

Proposition 7 *Suppose (compensated) demand functions can be written as*

$$\widetilde{c}^e = \gamma_c(w) \times \beta_c(\tau_c, \tau_d) \qquad \textit{and} \qquad \widetilde{d}^e = \gamma_d(w) \times \beta_d(\tau_c, \tau_d). \tag{21}$$

In this case, in a second-best setting where the planner taxes both first period and second period consumption together with a lump sum transfer, the optimal value of the consumption taxes is given by

$$\tau_c = \frac{\operatorname{cov}(b, c^e)\, E\frac{\partial \tilde{d}^e}{\partial \tau_d} - \operatorname{cov}(b, d^e)\, E\frac{\partial \tilde{d}^e}{\partial \tau_c}}{E\frac{\partial \tilde{c}^e}{\partial \tau_c}\; E\frac{\partial \tilde{d}^e}{\partial \tau_d} - E\frac{\partial \tilde{c}^e}{\partial \tau_d}\; E\frac{\partial \tilde{d}^e}{\partial \tau_c}} + \frac{E\frac{\Delta}{\lambda}\frac{\partial \tilde{c}^e}{\partial \tau_c}}{E\frac{\partial \tilde{c}^e}{\partial \tau_c}},$$

and

$$\tau_d = -\frac{\operatorname{cov}(b, c^e)\, E\frac{\partial \tilde{c}^e}{\partial \tau_d} - \operatorname{cov}(b, d^e)\, E\frac{\partial \tilde{c}^e}{\partial \tau_c}}{E\frac{\partial \tilde{c}^e}{\partial \tau_c}\; E\frac{\partial \tilde{d}^e}{\partial \tau_d} - E\frac{\partial \tilde{c}^e}{\partial \tau_d}\; E\frac{\partial \tilde{d}^e}{\partial \tau_c}}.$$

In this special case, only the tax on first period consumption includes a Pigouvian term and the taxation of second period's consumption is only used for redistribution. The Pigouvian term is positive, meaning that the tax on the first period's consumption is higher in the presence of myopic habit formation than when this type of behavior is absent.

This can be stated in terms of the so-called principle of targeting (Sandmo, 1975) which says that to correct for the consequences of externalities Pigouvian terms appear only in the expressions for the taxes of the externality generating goods. The *expressions* for the other goods are not affected.[15] One can see myopic behavior as generating an externality from one self to another one. The reason that the principle of targeting found by Sandmo does not hold in the general model in this paper is that it applies only to "atmosphere externalities", *i.e.,* to situations where the externality depends only on aggregate consumption. When externalities are not of the atmosphere type, the principle of targeting does not apply as happens in the general formulation in this paper (unless we have multiplicative compensated demand functions as specified by (21)).

The analysis up to now has assumed that everyone in society was myopic. We now move to the case of heterogeneous societies where myopic coexist with farsighted agents. The analytical treatment of this problem is indeed extremely complex and the ensuing formulas are not very intuitive, so that we rather provide numerical simulations.

[15] But the *level* of the tax rates is affected.

4 Numerical simulations

In this section, we assume that a proportion λ of individuals are myopic (all with $\alpha = 0$) while the remaining fraction $1 - \lambda$ are farsighted. We assume that $\overline{\alpha} = 1/4$, $u(x) = \ln(x)$ and that $h(x) = x^2/2$. Table 1 presents the results obtained when all individuals have the same wage, $w = 2$. Like in the theoretical sections, social welfare is given by the sum of indivuals utilities (utilitarian welfare function).

The first two numerical columns report the laissez-faire allocation (the first one for a farsighted agent, the second one for a myopic) defined in Subection 2.2. We know from Proposition 2 that first-period consumption c is larger for the myopic agent (since he saves less), that effective labor supply ℓ is larger for the myopic, and that utility U is larger for the farsighted. The comparison of planned labor supply of the two types of agents was ambiguous in general. With the functional forms we have chosen, the two effects of myopia on planned labor supply cancel out: farsighted and myopic agents have exactly the same planned labor supply, ℓ^p. Since farsighted agents consume less (and hence save more) in the first period than myopics and plan the same amount of labor supply in the second period, their planned second-period consumption d^p is larger than for myopic individuals. Even though myopics end up working more than farsighted agents, their second-period consumption d is lower (since their saving is lower). As a matter of fact, with our assumption that $\alpha = 0$, myopic agents do not save at all in the laissez-faire. When society consists only of farsighted agents ($\lambda = 0$), the laissez-faire outcome is also the first-best solution; recall that all individuals have the same income for the time being. If society is composed only of myopics ($\lambda = 1$), the first-best allocation corresponds to the laissez-faire allocation *of farsighted individuals*. The last column of Table 1 shows that the first-best allocation can be decentralized using a tax on first-period consumption only. This result (which is a special case of Proposition 3) is intuitive: by choosing optimally τ_c, the planner induces the myopics to consume the first-best amount in the first-period. Agents then make mistakes when assessing their *planned* second-period labor supply and consumption levels. However, this is of no relevance for effective labor supply and consumption. Once individuals have reached their second period, they realize their mistake and take the correct decisions. This is because they then return to their long run preferences and because they have been induced to save the optimal amount.

Intermediate columns in Table 1 report results for mixed societies. As the proportion of myopics increases, the optimal value of τ_c increases from zero (when $\lambda = 0$) to 38% (when $\lambda = 1$), while the tax on second-period

	Laissez-faire		Second-Best						
λ			0.25		0.5		0.75		1
τ_c	0		0.13		0.23		0.31		0.38
τ_d	0		0		0		0		0
Type	F	M	F	M	F	M	F	M	M
c	1.600	2.000	1.500	1.831	1.435	1.726	1.390	1.655	1.600
d^p	2.400	2.000	2.444	2.069	2.482	2.124	2.516	2.169	2.209
ℓ^p	1.000	1.000	0.966	0.966	0.942	0.942	0.922	0.922	0.905
d	2.400	2.266	2.444	2.304	2.482	2.338	2.516	2.369	2.400
ℓ	1.000	1.133	0.966	1.084	0.942	1.049	0.922	1.023	1.000
U	0.663	0.620	0.665	0.631	0.671	0.641	0.678	0.652	0.663
$\ell - \ell^p$		0.133		0.117		0.107		0.100	0.095
$d - d^p$		0.266		0.234		0.215		0.201	0.190

Table 1: Numerical results with a single wage level

consumption, τ_d, remains at zero. This is an interesting result: although, with mixed populations, the tax on first-period consumption is no longer sufficient to reach the first-best allocation anymore, no gain in aggregate welfare is attained by taxing second-period consumption. Observe that the relationships detailed above for the laissez-faire situation between the levels of c, d^p, ℓ^p, d, ℓ and U of both types of agents all remain valid at the second-best allocation, whatever the proportion λ of myopics. One striking feature of Table 1 is that the utility of *both* types increases with λ. This is not surprising for myopic individuals: as their proportion in the economy is raised, the second-best taxation policy becomes increasingly tailored to their needs, with the first-best allocation attained when $\lambda = 1$. Farsighted agents also benefit from being pooled with myopic agents because myopics have a larger first-period consumption, and hence pay more tax than farsighted agents of the same wage level. Farsighted individuals then benefit from the redistribution from myopic consumers, and Table 1 shows that this redistribution effect is larger than the utility cost created by the distortive τ_c for the farsighted, and is increasing with λ because a larger value of λ increases the optimal value of τ_c. In other words, even in a simple world with identical wages, the optimal consumption tax has redistributive effects (from myopics towards farsighted agents) in addition to efficiency effects (negative for farsighted individuals as the tax distorts their choice away from the optimum, and positive for myopics as the tax helps them internalize their myopia).

As λ and τ_c increase, first-period consumption c decreases for both types of agents (and is then smaller than the first-best level of 1.6 for the farsighted agents, while remaining larger for the myopics). Beyond discouraging first-period consumption, the tax τ_c decreases labor supply (both planned ℓ^p and effective ℓ) and increases second-period consumption (also, both planned d^p and effective d) for both types of agents. Interestingly, the difference between planned and effective labor supply of myopics, $\ell - \ell^p$, and between planned and effective second period consumption of myopics, $d - d^p$, both decrease as λ increases. This result is intuitive as it shows that, as the proportion of myopics is raised, the second-best optimal policy manages to increasingly alleviate the myopia problem, as measured by the difference between planned and effective variables. The fact that the utility level attained by the myopics is always lower than the one reached by farsighted agents, for any value of $0 < \lambda < 1$, is consistent with the fact that second-best policy is too crude an instrument to perfectly correct for the myopia of agents in heterogenous societies.

We have also performed the same numerical computations (available upon request) with a larger value of $\bar{\alpha}$, and we obtain the same qualitative results (with larger values of the optimal τ_c, compared with Table 1, for any $0 < \lambda \leq 1$).

We now turn to the case where agents differ in wages. Table 2 reports numerical results for the same functional forms as in Table 1, but where the distribution of wages is bell-shaped and symmetrical around 2 —i.e., the average wage in Table 2 is equal to the uniform wage in Table 1.[16]

For simplicity, we assume a zero correlation between income and myopia, so that the income distribution is the same for both types of agents. Table 2 reports the average (over all income levels) consumption, labor supply and utility levels for myopic and farsighted agents, according to the proportion of both types in the economy. From the laissez-faire situation, we obtain that consumption and labor supply levels are linear in wage w, so that their average levels (reported in the first two numerical columns) correspond to the levels reported in Table 1. Utility is not linear but concave in w, so that average utility is lower than the utility of the average-income individual. Contrary to Table 1, redistribution across income levels calls for public intervention even

[16] Income is distributed on the interval [0,4] according to a Beta distribution with parameters (2,2).

	Laissez-faire		Second-Best							
λ			0	0.25		0.5		0.75		1
τ_c	0		0.60	0.76		0.88		0.96		1.04
τ_d	0		0.60	0.56		0.56		0.52		0.52
Type	F	M	F	F	M	F	M	F	M	M
c	1.600	2.000	1.390	1.323	1.616	1.286	1.553	1.257	1.501	1.470
d^p	2.400	2.000	2.085	2.154	1.823	2.193	1.871	2.250	1.935	1.973
ℓ^p	1.000	1.000	0.680	0.665	0.665	0.647	0.647	0.642	0.642	0.629
d	2.400	2.266	2.085	2.154	1.970	2.193	2.007	2.250	2.065	2.097
ℓ	1.000	1.133	0.680	0.665	0.775	0.647	0.749	0.642	0.736	0.719
U	0.383	0.340	0.537	0.551	0.512	0.567	0.532	0.584	0.553	0.573
$\ell - \ell^p$		0.133			0.110		0.102		0.094	0.090
$d - d^p$		0.266			0.147		0.136		0.129	0.123

Table 2: Numerical results with symmetrical distribution of wages

in a homogenous farsighted society, as can be seen from the third numerical column of Table 2. Recall that the social welfare function is utilitarian and that individual utility functions are stricly concave. The second-best tax is 60%, both on first- and second-period consumption. The tax is the same in both periods since there is no reason to distort consumption choices across periods for farsighted agents. As λ increases, τ_c increases while τ_d slightly decreases (we can show that total tax proceeds increase with λ): redistribution across income levels increases using the tax on first-period consumption rather than on second-period consumption. This is intuitive since, as shown in Table 1, τ_c helps the increasingly large group in the economy internalize their myopia problem. As in Table 1, the gap between effective and planned variables for myopics decreases with λ. Both tax rates are much larger than in Table 1, whatever the value of λ, which is a reflection of the added reason to tax consumption in Table 2, namely redistribution across income levels. As in Table 1, average utility increases for both myopics and farsighted agents when λ increases, due to the redistribution from myopics to farsighted agents of the same wage levels, but the average utility of myopics is always smaller than the average utility of farsighted agents for any $0 < \lambda < 1$.

Interestingly, and in contrast to Table 1, the average utility in a myopic society ($\lambda = 1$) is, at 0.573, larger than the average utility in a farsighted society ($U = 0.537$ when $\lambda = 0$). To understand this result, we have to keep in mind that the average utility reported is not the utility of the average-income individual. The larger redistribution observed in a myopic-only society (compared to the farsighted-only case) trumps the distortion effect with our functional forms, resulting in the counter intuitive result of a larger average utility for myopics than for farsighted agents. In this sense, myopia allows the social planner to reach a larger value of its objective since it allows for more redistribution at the optimum.

Finally, we have performed numerical simulations (available upon request) with the same assumptions, but with a positively skewed distribution of income, keeping the average income constant at 2.[17] The qualitative results reported in Table 2 remain unchanged, the main difference being larger values of both τ_c and τ_d for any value of λ, which is intuitive since skewing the income distribution to the right increases the case for redistributive taxation.

The main results are summarized in the following proposition.

Proposition 8 *Assume that a proportion λ of individuals are myopic (all with $\alpha = 0$) while the remaining fraction $1 - \lambda$ are farsighted, that $\overline{\alpha} = 1/4$, $u(x) = \ln(x)$, and that $h(x) = x^2/2$.*

[17] Income is distributed on the support [0,6] according to a Beta(2,4) distribution.

(A) When all individuals have the same wage ($w = 2$) we have:

(i) In a mixed society ($0 < \lambda < 1$), the first-best allocation can no longer be decentralized by consumption taxes only.

(ii) The second best calls for a tax on first-period consumption only; its rate increasing with λ.

(iii) As λ increases, the second-best optimal policy manages to increasingly alleviate the myopia problem, as measured by the difference between planned and effective variables. The utility of both types of agents increases with λ, with farsighted agents benefitting from redistribution from (high consumption) myopics.

(B) When indivisuals differ also in income we have:

(i) Public intervention is called for even in homogeneous societies ($\lambda = 0$ or $\lambda = 1$). With a symmetrical and bell-shaped distribution of income around the same mean as in (A), with no correlation between income and myopia, it takes the form of the same tax rate on both first- and second-period consumption (which is equivalent to a linear income tax).

(ii) As λ increases, the tax on first-period (resp., second) consumption increases (resp., decreases) and the second-best solution decreases the difference between planned and effective variables, as in (A) above. Furthermore, the utility of both types of agents increases with λ, as in (A).

5 Conclusion

This paper analyzes the pattern of consumption taxes in a two period model with habit formation and myopia. The needs individuals have in the second period increase with their first period consumption. However, the myopic individuals do not see this habit formation relation when they take their saving decision.

Our main results are as follows. In the laissez-faire situation, myopic individuals consume more in the first period than if they were farsighted and end up working more in the second period. Whether they consume more or less in the second period than if they were farsighted depends on the comparison of the concavity of the utility derived from consumption and from leisure. When individuals differ in productivity, the first-best allocation calls for more consumption in the second period than in the first but equalizes consumption across individuals in each period. This allocation can be decentralized using a tax on first period consumption together with individualized lump sum transfers. If these transfers are not available, we show how the classical formula determining the second-best uniform lump sum transfer is modified by the introduction of myopia: a term reflecting the cost of myopia in terms of ex post utility has to be added in the computation of the usual Atkinson-Stiglitz (1980)'s net marginal valuation of income. If a single consumption tax is used by the planner in conjunction with the uniform lump sum transfer, we show that the optimal consumption tax formula consists of two terms: the classical term trading off the redistributive objective and the efficiency cost of redistribution, and a term which increases with the ex post utility cost of myopia. In the case where two consumption taxes are used (one for each period of consumption), we show how cross price effects affect the optimal tax formulas obtained above. Finally, in the case of demand functions exhibiting multiplicative separability, we obtain that the tax on first period consumption is corrected by a Pigouvian term and the taxation of second period's consumption is only used for redistribution. This is in line with Sandmo (1975)'s principle of targeting which says that to correct for the consequences of externalities Pigouvian terms appear only in the expressions for the taxes of the externality generating goods.

We finally resort to simulations to deal with the case of a society composed of both farsighted and myopic individuals, and where the individual's type is not observable to the planner. We first consider a society where all individuals, myopic or farsighted, have the same productivity. We obtain that both types of individuals get better off at the second-best optimum when the share of myopics increases. The reason is twofold: myopics are better off because the second-best policy is more tailored to their needs, while farsighted agents benefit from the redistribution from the higher-consumption myopics. When productivity heterogeneity is introduced, our simulations show that myopia allows the planner to redistribute more, and to reach a larger average utility than with farsighted agents.

In most cases, the optimal tax rates are not the same in both periods of life. The idea that taxation should vary with age is not new. Banks and Diamond (2010) following what is called the new dynamic public finance argue in favor of an age-dependent taxation. Lozachmeur (2006) reaches the same conclusion, but showing that elderly workers should be subject to a lower tax than the others because they take both intensive (how many hours a week?) and extensive (when to retire?) labor supply decisions. In this paper, we also reach the conclusion that second period consumption should be taxed at a lower rate than first period's. The reason is that one has to correct for a myopic behavior which leads individuals to save too little and forces them to work longer than initially expected.

Appendix

A Impact of α on savings and labor supply

Differentiating the FOCs (1) and (2) with respect to α, and denoting the derivatives of s^e and ℓ^p with respect to α by s^e_α and ℓ^p_α, we obtain

$$u''(c)s^e_\alpha + (1+\alpha)u''(d-\alpha c)[w\ell^p_\alpha + (1+\alpha)s^e_\alpha] + u'(d-\alpha c) - (1+\alpha)u''(d-\alpha c)[w-s^e] = 0,$$

$$wu''(d-\alpha c)[w\ell^p_\alpha + (1+\alpha)s^e_\alpha] - wu''(d-\alpha c)[w-s^e] - h''(\ell^p)\ell^p_\alpha = 0,$$

which we express in matrix form as follows:

$$\begin{bmatrix} s^e_\alpha \\ \ell^p_\alpha \end{bmatrix} \begin{bmatrix} u''(c) + (1+\alpha)^2 u''(d-\alpha c) & (1+\alpha)wu''(d-\alpha c) \\ (1+\alpha)wu''(d-\alpha c) & w^2u''(d-\alpha c) - h''(\ell^p) \end{bmatrix} = \begin{bmatrix} -u'(d-\alpha c) + (1+\alpha)u''(d-\alpha c)[w-s^e] \\ wu''(d-\alpha c)[w-s^e] \end{bmatrix}$$

Using Cramer's rule we get the expressions

$$\begin{aligned} s^e_\alpha &= \frac{\underset{-}{R}[w^2\underset{-}{u''(d-\alpha c)} - h''(\ell^p)] - (1+\alpha)w^2(u''(d-\alpha c))^2\underset{+}{[w-s^e]}}{D}, \\ \ell_\alpha &= \frac{[u''(c) + (1+\alpha)^2 \underset{-}{u''(d-\alpha c)}]w\underset{-}{u''(d-\alpha c)}[w-s^e]}{D} \\ &\quad - \frac{\underset{+}{(1+\alpha)wu''(d-\alpha c)R}}{D}. \end{aligned}$$

where

$$\begin{aligned} D &= [u''(c) + (1+\alpha)^2u''(d-\alpha c)][w^2u''(d-\alpha c) - h''(\ell^p)] \\ &\quad -[(1+\alpha)wu''(d-\alpha c)][(1+\alpha)wu''(d-\alpha c)] > 0 \end{aligned}$$

and

$$R = -u'(d-\alpha c) + (1+\alpha)u''(d-\alpha c)[w-s^e] < 0.$$

From there, it is easy to show that $s^e_\alpha > 0$, since its numerator is equal to

$$
\begin{aligned}
&[-u'(d-\alpha c)+(1+\alpha)u''(d-\alpha c)[w-s^e]][w^2u''(d-\alpha c)-h''(\ell^p)] \\
&\quad -(1+\alpha)w^2(u''(d-\alpha c))^2[w-s^e]= \\
&\quad -u'(d-\alpha c)w^2u''(d-\alpha c)+(1+\alpha)w^2(u''(d-\alpha c))^2[w-s^e] \\
&\quad -Rh''(\ell^p)-(1+\alpha)w^2(u''(d-\alpha c))^2[w-s^e]= \\
&-u'(d-\alpha c)w^2u''(d-\alpha c)-[-u'(d-\alpha c)+(1+\alpha)u''(d-\alpha c)[w-s^e]]h''(\ell^p)= \\
&\quad -u'(d-\alpha c)w^2u''(d-\alpha c)-Rh''(\ell^p)>0.
\end{aligned}
$$

References

Atkinson, A. and J. Stiglitz, 1980, *Lectures on Public Economics*, Mc Graw Hill, New York.

Banks, J. and P. Diamond, 2010, The base for direct taxation, in S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Gammie, P. Johnson, G. Myles and J. Poterba, eds., *Dimensions of Tax Design: The Mirrlees Review*, Oxford Univesity Press, Oxford, 548–648.

Bernheim B. D. and A. Rangel, 2009, Beyond revealed preference: choice-theoretic foundations for behavioral welfare economics. *Quarterly Journal of Economics*, 124, 1, 51–104.

Brown, T.N., 1952, Habit persistence and lags in consumer behavior, *Econometrica*, 20, 355–371.

Camerer, C., S. Issacharoff, G. Loewenstein, T. O'Donoghue and M. Rabin, 2003, Regulation for conservatives: behavioral economics and the case for "asymmetric paternalism", *University of Pennsylvania Law Review 151*, 1211–1254.

Cremer, H., P. De Donder, D. Maldonado, P. and Pestieau, 2009, Forced Saving, Redistribution and Non Linear Social Security Scheme, *Southern Economic Journal*, 76, 86–98.

Diamond, P. and B. Koszegi, 2003, Quasi-hyperbolic discounting and retirement. *Journal of Public Economics*, 87, 1839–1872.

Diamond, P. and J. Mirrlees, 2000, Adjusting one's standard of living: Two period models, in P. J. Hammond and G. D. Myles (eds.), *Incentives, Organization, and Public Economics papers in Honour of Sir James Mirrlees.* Editors, Oxford University Press, Oxford.

Fehr, E. and P. K. Zych, 2008, Intertemporal choice under habit formation, in: C. Plott and V.L. Smith, eds., *Handbook of Experimental Economics*, Elsevier, Amsterdam, The Netherlands, 923–928.

Gruber, J. and B. Koszegi, 2001, Is addiction 'rational'? theory and evidence, *Quarterly Journal of Economics,* 116, 1261–1303.

Gruber, J. and B. Koszegi, 2004, Tax incidence when individuals are time-inconsistent: the case if cigarette excise taxes, *Journal of Public Economics*, 88, 1959–1987.

Layard, R., 2005, Rethinking public economics: The implications of rivalry and habit, in L. Bruni and L. Porta, eds, *Economics and Happiness: Framing the Analysis*, Oxford University Press, Oxford, UK.

Lozachmeur, J-M., 2006, Optimal age specific income taxation, *Journal of Public Economic Theory*, 8, 697–711.

Maestas, N., 2007, Back to work: Expectations and Realizations of work after retirement, Rand Labor and Population WR 196–2.

Muellbauer, J., 1988, Habits, rationality and myopia in the life cycle consumption function, *Annales d'Economie et de Statistique* 9, 47–70.

O'Donoghue, T. and M. Rabin, 1999, Procrastination in preparing for retirement, in H. J. Aaron, ed., *Behavioral Dimensions of Retirement Economics*, Brookings Institution, Washington, D.C.

O'Donoghue, T. and M. Rabin, 2003, Studying optimal paternalism illlustrated with a model of sin taxes, *American Economic Review Papers and Proceedings*, 93, 186–191.

O'Donoghue, T. and M. Rabin, 2006, Optimal sin taxes, *Journal of Public Economics*, 90, 1825–1849.

Sandmo, A., 1975, Optimal Taxation in the Presence of Externalities, *The Swedish Journal of Economics*, 77, 86–98.

Woittiez, I. and A. Kapteyn, 1998, Social interactions and habit formation in a model of female labour supply, *Journal of Public Economics*, 70, 185–205.