

**Optimal Second Best Taxation of Addictive Goods
in Dynamic General Equilibrium:
A Revenue Raising Perspective***

Luca Bossi
Department of Economics
University of Pennsylvania
Philadelphia, PA 19104-6297
boluca@econ.upenn.edu

Pedro Gomis-Porqueras
School of Economics
HW Arndt Building 25a
Australian National University
ACT 0200 Australia
peregomis@gmail.com

and

David L. Kelly
Department of Economics
University of Miami
Box 248126
Coral Gables, FL 33124
dkelly@miami.edu

July 1, 2011

*We would like to thank Stephen Coate, Jang-Ting Guo, Narayana Kocherlakota, Adrian Peralta-Alva, Manuel Santos, Stephen E. Spear, Richard Suen, and seminar participants at the Society for the Advancement of Economic Theory meetings, the North American Summer Meeting of the Econometric Society, the Latin American Meetings of the Econometric Society, the University of Miami, the University of California at Berkeley ARE, the San Francisco FED, and UC Riverside for helpful comments and suggestions.

Abstract

In this paper we derive conditions under which optimal tax rates for addictive goods exceed tax rates for non-addictive consumption goods within a rational addiction framework where exogenous government spending cannot be financed with lump sum taxes. We reexamine classic results on optimal commodity taxation and find a rich set of new findings. Two dynamic effects exist. First, households anticipating higher future addictive tax rates reduce current addictive consumption, so they will be less addicted when the tax rate increases. Therefore, addictive tax revenue falls prior to the tax increase. Surprisingly, the optimal tax rate on addictive goods is generally decreasing in the strength of tolerance, since strong tolerance strengthens this tax anticipation effect. Second, high current tax rates on addictive goods make households less addicted in the future, affecting all future tax revenues in a way which depends on how elasticities are changing over time. Classic results on uniform commodity taxation emerge as special cases when elasticities are constant and the addiction function is homogeneous of degree one. Finally, we also study features of addictive goods such as complementarity to leisure that, while unrelated to addiction itself, are nonetheless common among many addictive goods.

Keywords: Ramsey model, dynamic optimal taxation, addictive goods, habit formation.

JEL Codes: E61, H21, H71.

“People should understand: Those who drink, those who smoke are doing more to help the state...If you smoke a pack of cigarettes, that means you are giving more to help solve social problems such as boosting demographics, developing other social services and upholding birth rates.”

Russian finance minister Alexei Kudrin on September 2, 2010 announcing the Russian government’s plan to raise excise duties on alcohol and cigarettes.

1 Introduction

A popular and increasingly common way for local, state, and federal governments to raise revenue is through taxation of addictive goods, including cigarettes, alcohol, and gambling. According to the Tax Policy Center, in 2008 the US cumulative state and local alcohol, tobacco and lottery tax revenues exceeded 43 billion dollars. What is the optimal excise tax for addictive goods, when the government must raise revenue to finance a stream of exogenous government expenditures? The goal of this paper is to characterize and analyze the conditions under which taxation of addictive goods might differ from taxes on labor and non-addictive consumption goods (hereafter *ordinary goods*).

This paper extends classic results of optimal commodity taxation (e.g. Atkinson and Stiglitz 1972) to the case of addictive goods and obtains a rich set of new dynamic findings not found in typical models which are either static or assume very specific utility functions or both. For common cases such as homothetic and separable utility, we show that the classic results obtained in the literature on optimal taxation of ordinary goods do not necessarily extend to addictive goods, when addiction is modeled as a rational dynamic process as in Becker and Murphy (1988), hereafter BM.

Two dynamic effects emerge. Both stem from the *tolerance* property of addictive goods: past consumption decreases current utility by raising the amount of consumption needed to sustain the addiction (BM). The first effect we label the *tax anticipation effect*. To gain insight, let us consider an announced increase in the addictive goods excise tax. Households anticipate prior to the tax increase that higher addictive taxes will reduce future addictive consumption. This provides an incentive to reduce addictive consumption in the period prior to the increase in taxes, since doing so makes the household less addicted when the tax increase occurs. Since households become less addicted, consumption in excess of that required to sustain addiction (hereafter *effective consumption*) does not fall as much: addictive consumption falls, but so does the level of addiction. The decrease in addictive consumption in the period prior to the increase in addictive taxes reduces addictive tax revenue and

therefore moderates the total revenue raised over time.

We show that, under the mild assumption of homogeneity of the function which maps addictive consumption into effective consumption, the tax anticipation effect is equivalent to a single current period revenue effect. In particular, the dynamic tax anticipation effect is equivalent to a smaller tax in a model without addiction. This allows us to evaluate the merits of taxing addictive goods in familiar terms, such as the static complementarity between effective consumption and leisure.

Second, high current tax rates on addictive goods causes households to be less addicted in the future. Future effective consumption rises, affecting future addictive, ordinary, and labor tax revenues. We call this effect the *addiction stock effect*. Suppose for example that effective consumption is becoming more complementary with leisure over time. Standard public finance theory (Corlett and Hague 1953) suggests that the tax rate on addictive goods should be relatively high, since reduced consumption of addictive goods will increase labor supply, thus raising labor income tax revenues. However, with tolerance, reduced current consumption of addictive goods raises future effective consumption (households are less addicted in the future, and therefore get more effective consumption in the future from a given quantity of addictive goods). But then future labor supply falls, and future labor income tax revenues fall, offsetting some of the revenue gains in the current period. Whether total tax revenues rise or fall depends on how elasticities are changing over time. For example, if the complementarity of addictive consumption with leisure is decreasing over time, labor tax revenues are less sensitive to changes in current addictive consumption than to changes in future addictive consumption. Taxing addictive goods becomes less attractive over time. Optimal addictive taxation should smooth distortions: taxing addictive goods today may make taxing addictive goods more or less distortionary in the future depending, for example, on their relationship with leisure.

The addiction stock effect smoothes intertemporal distortions caused by taxation for revenue raising. In this sense, our results are related to those on capital taxation (Chamley 1986, Chari and Kehoe 1998). A disadvantage of capital taxation is that it reduces the future capital stock and thus the future tax base. Similarly, taxation of addictive goods reduces the future stock of addiction, and through the elasticity, the future tax base. Nonetheless, our results differ because addictive consumption acts like both a finished good (current addictive consumption) and as an intermediate good (current addictive consumption affects future addictive consumption). For instance, optimal steady state tax rates on addictive goods equal tax rates on ordinary consumption in some cases where optimal capital tax rates are

zero, because addictive consumption acts like a finished good.

Classic results of uniform taxation for addictive and ordinary goods emerge as a special case when elasticities are constant and the effective consumption function is homogeneous of degree one. Constant elasticities ensures no distortion smoothing motivation exists, whereas homogeneity of degree one ensures the tax on addictive consumption is equivalent to a tax on effective consumption.

The literature typically models effective consumption in one of two ways: the subtractive specification (e.g. Campbell and Cochrane 1999, Chugh 2007) and the multiplicative specification (e.g. Abel 1990). These two models differ in terms of their homogeneity properties.¹ In this paper, we show that the optimal tax policy depends crucially on the degree of homogeneity of the addiction function.

In particular, we show that, given separable or homothetic utility with constant relative risk aversion, raising the degree of homogeneity makes addictive consumption more income inelastic in the equivalent static model without addiction. Thus, taxation of addictive goods is more attractive if the addiction model is homogeneous of degree one, as in the subtractive case, than if the addiction model is homogeneous of degree less than one, as in the multiplicative case, since it is optimal to tax necessities at a higher rate. Further, strong tolerance in the multiplicative model decreases the degree of homogeneity, making addictive goods more income elastic, which therefore *lowers* the optimal tax rate on addictive goods.

To compute optimal tax rates, we use the primal approach to Ramsey taxation. A large literature exists using the primal approach. Chari and Kehoe (1998) and Erosa and Gervais (2001) derive many results with both finitely and infinitely lived agents. More recent examples include Kyung (2009), Armenter (2008), Chugh (2007), Cunha (2008), Erosa and Gervais (2002), and Sleet and Yeltekin (2006). Chugh (2007) in particular undertakes a quantitative study of optimal inflation rates in a model where consumption goods are habit forming. The primal approach does restrict taxes to be linear, but this is not restrictive for the case of taxes on addictive goods.²

¹A good is habit forming if the marginal utility of the good is increasing in past consumption. We use the standard definition of addiction, which is when current consumption is increasing in past consumption, holding fixed the marginal utility of wealth and prices. Habit formation is often used in the macro literature, whereas addiction was introduced by BM. It is straightforward to show that the subtractive model of habit formation implies the good is addictive, and the multiplicative model of habit formation implies the good is addictive with an additional restriction.

²If households are heterogeneous, high addictive consumption may signal high stock of addiction, and therefore inelastic demand, which may imply higher optimal tax rates. While allowing for nonlinear taxes is possible using the Mirrless approach (Kocherlakota 2005), we chose not to adopt that framework for three reasons. First, nonlinear taxation of addictive goods is unlikely to be feasible in practice. Households can evade nonlinear taxes by buying addictive goods in small quantities or over the Internet (Goolsbee,

In the next section we describe the main alternative motives for taxing addictive goods found in the literature. In sections 3 and 4, we develop a dynamic, rational addiction model and determine conditions under which optimal tax rates for addictive goods exceed tax rates for ordinary consumption goods. Sections 5 give results for a general class of preferences and for a specific analytical example. The appendix contains the proofs of all the propositions and extends the analysis to two other preferences classes.

2 Taxing Addiction

Three classical motivations exist in the literature for taxing addictive goods differently than ordinary goods. The first is to lower the external costs often associated with consumption of addictive goods. The second is because some households fail to take into account some private costs and thus over-consume. The motivation which is the focus of this paper is to raise revenue.

2.1 Addictive Goods and Fiscal Concerns

Taxation of many addictive goods, such as lotteries, has an obvious revenue raising component. Taxes on many other addictive goods are dictated by fiscal concerns. For example, the repeal of alcohol Prohibition in the 1930's was clearly approved for tax revenue considerations as the onset of the Great Depression negatively impacted all other government revenues (Boudreaux and Pritchard 1994). Parry, Laxminarayan, and West (2009) also note that the last two increases in federal alcohol taxes were part of deficit reduction packages.³

In November 2010 in California a ballot measure that would have allowed local government to legally sell marijuana was put to vote and rejected with 54% of the voters against it. The "Regulate, Control and Tax Cannabis Act" would have permitted local governments

Lovenheim, and Slemrod 2009), or by bootlegging across states, and the government cannot easily aggregate all addictive good purchases made by the household. Second, the majority of addictive taxes in the US are administered locally and states and municipalities tax addictive goods mostly because they are concerned with raising revenue (see the next paragraph for evidence) rather than optimal redistribution. Indeed, the vast majority of redistribution efforts are carried out at the federal level, since tax competition across states limits the ability of state's to maintain discriminatory tax rates and since incomes within states and municipalities are more homogeneous than across the US. Lastly, in practice (probably because of the first two reasons), in the US tax system we see almost exclusively linear excise taxes.

³For lotteries, external costs are presumably small, but the nationwide average lottery tax ranges from 40% in 1989 (Clotfelter and Cook 1990) to 31% in 2003 (Hansen 2004), accounting for 2% of state tax revenues. States spent about \$272 million on lottery advertising in 1989, which is at least a strong indication that states are motivated by revenue concerns, rather than the external costs of lotteries and other forms of gambling. Finally, proposals exist to use higher cigarette taxes to close budget deficits in Florida, Illinois, West Virginia, and elsewhere.

to impose and collect cannabis-related fees and taxes similarly to what is done with alcohol. The revenue potential of taxing marijuana was the major argument in favor of this ballot measure. In an interesting and thought provoking paper, Miron (2010) analyzes the budgetary implications of legalizing drugs. He finds that the savings in government expenditure and the gains in tax revenue (at the federal and local level) that would result from replacing drug prohibition with a regime in which drugs are legal, and taxed like alcohol and tobacco to be \$83 billion per year.

Since the poor presumably spend a higher fraction of income on addictive goods, taxing addictive goods for revenue raising must be justified on efficiency grounds, rather than redistribution. A few papers consider the efficient revenue raising motivation by treating addictive goods in a static way as simply goods with external costs, which are possibly complementary with leisure. If so, one can apply the ideas from the “double dividend” literature (e.g. Bovenberg and Goulder 1996). Taxing a good with external costs raises revenues which can be used to reduce taxes on labor income (the “revenue recycling effect”). If taxing addictive goods results in lower dead-weight losses than taxing labor, then the revenue recycling effect is positive and it is optimal to tax addictive goods at a relatively high rate. Moreover, a good with external costs may also be taxed above its Pigouvian rate for revenue raising if it is complementary with leisure, since the tax increases labor supply and labor income tax revenues (the “tax interaction effect”).⁴

However, this literature models addiction in a static way as simply a good with external costs; the dynamic nature of addiction is ignored. It remains unclear how dynamic addictive properties such as tolerance affect optimal revenue raising. This paper fills this gap in the literature by considering a dynamic model of rational addiction while explicitly considering a revenue raising motive. We model addiction using BM’s rational addiction framework. In this approach, addictive consumption is linked to past consumption, but the first welfare theorem holds. If no externalities exist, then the only rationale in the BM setting for taxing addictive goods is a revenue raising one. Unlike static models, in our dynamic framework changes in tax rates on addictive goods affects future revenues, by changing future elasticities.

⁴Sgontz (1993) finds the revenue recycling effect to be positive, and Parry, Laxminarayan, and West (2009) finds both the revenue recycling effect and the tax interaction effect to be positive: alcohol is complementary to leisure and also reduces labor productivity. Therefore, they find it is optimal to tax alcohol above its Pigouvian rate as part of the optimal revenue raising package.

2.2 Addictive Goods and Externalities

The standard economic rationale for taxation of addictive goods is that their consumption is often associated with external costs, such as second-hand smoke, drunk driving, and crime. The literature often finds addictive goods are taxed at a rate less than the rate which is second best in the sense that the government cannot discriminate between consumers who generate external costs and responsible consumers.⁵

If we were to include a negative consumption externality for the addictive good in our model, then the results go through under slightly different conditions. In particular, the conditions for which the optimal tax rate of an addictive good is above the rate which corrects the externality is similar to the conditions derived here for which the optimal tax rate of an addictive good is above the tax rate for ordinary consumption goods. Hence, our results can be simply interpreted as relative to the tax ratio which corrects the externality.

However, we decided not to include an externality in our problem for three reasons: first, it is well known (Kenkel 1996, Pogue and Sgontz 1989) that taxing an addictive good whose consumption is imperfectly correlated with an externality is a second-best solution. Taxing the actual behavior causing the externality (e.g. make the punishment for drunk driving more severe or banning smoking in public places) is more efficient. Indeed, Parry, Laxminarayan, and West (2009) show that welfare gains from increasing drunk driving penalties exceed those from raising taxes on alcohol, even when implementation costs and dead-weight losses associated with incarceration are included. Second, as noted above the results are similar when an externality is included. Finally, our focus is on the effect of addiction on revenue raising, rather than externalities.

⁵For example Kenkel (1996) finds that a tax rate on alcohol of about 42% is optimal for the drunk driving externality, while the actual average tax rate ranges from over 50% in 1954 to 20% in the 1980s. Moreover, Grinols and Mustard (2006) estimate external costs of casino gambling are 47% of revenues, thus the optimal tax would be higher than 47% if demand for casino gambling is inelastic, or less than 47% if a significant fraction of casino gamblers do not impose external costs. Anderson (2005) reports that casinos pay 16% of gross revenues in taxes. The empirical evidence is, however, mixed for cigarettes taxation: Manning, Keeler, Newhouse, Sloss, and Wasserman (1989) estimated the gross external cost of smoking in the U.S. of approximately \$0.43 per pack, but only \$0.16 per pack once reductions in health care expenditures stemming from premature deaths were included. Viscusi (1995) finds that after accounting also for the lower nursing home cost and retirement pension savings the net external costs of smoking are negligible for the U.S. Conversely, Gruber and Koszegi (2001) estimate external costs of smoking at \$0.94 to \$1.75 per pack, versus an average excise tax of about \$0.65.

2.3 Addictive Goods and Non-market Internal Costs

Another source of non-market costs occurs if addiction is modeled as non-fully rational excess consumption. Suppose households fail to take into account the self-adverse health effects caused by consumption of addictive goods, either because they are unaware that addictive goods consumption has adverse health effects (e.g. Kenkel 1996) or because some households are exogenously assumed to be unable to take into account the health gains from reducing addictive goods consumption (e.g. Pogue and Sgontz 1989). When some households are exogenously assumed not to consider some private costs, they over-consume. The resulting “internality” causes the optimal second best (again, in the sense that the government cannot distinguish between naive and rational households) tax rate to rise considerably.⁶

A related, subsequent literature makes excess consumption endogenous and rational by defining “sin goods” as goods for which preferences are time inconsistent (Gruber and Koszegi 2001, Gruber and Koszegi 2004, O’Donoghue and Rabin 2003, O’Donoghue and Rabin 2006). In this approach, households optimally choose to consume more now and less in the future. However, next period households also optimally choose to consume more now and less in the future. Hence households are rational, but over-consume in the sense that welfare increases with a tax that reduces consumption to a level which households would choose if they could pre-commit to consume less in the future.⁷

Given the stated concern for revenue raising in many of the state and municipal addictive tax programs, in this paper we decided to focus on how the dynamic properties of the utility function in a rational addiction framework affect optimal revenue raising.

3 Model

We consider an infinite horizon closed economy in discrete time. The economy is populated by a continuum of identical households of measure one who maximize the discounted sum of instantaneous utilities. A large number of identical firms produce both addictive and ordinary goods using a constant return to scale technology. Finally, there is a government that needs to finance a constant stream of government expenditures through fiscal policy.

⁶Kenkel (1996) finds the optimal tax rate on alcohol rises to about 106% while Pogue and Sgontz (1989) find the optimal tax rate on alcohol rises to 306%.

⁷O’Donoghue and Rabin (2006) compute numerical examples where the optimal tax on unhealthy foods ranges from 1-72%. Gruber and Koszegi (2001) show that the optimal tax on cigarettes rises to at least \$1 per pack when the time inconsistency problem is included.

3.1 Firms

A large number of identical firms at time t rent capital k_t and labor h_t from households to produce a composite good using a technology $F(k_t, h_t)$. We assume throughout the paper that:

Assumption A1 $F(k_t, h_t)$ is constant returns to scale and increasing, concave, and satisfies Inada conditions in each input.

Let w_t denote the wage rate and r_t the rental rate of capital, then the objective of the firm is to maximize profits, which equal:

$$\max_{k_t, h_t} \{F(k_t, h_t) - r_t k_t - w_t h_t\}. \quad (3.1)$$

Let subscripts on functions denote corresponding partial derivatives. The equilibrium rental rate and wage rate are:

$$r_t = F_k(k_t, h_t), \quad (3.2)$$

$$w_t = F_h(k_t, h_t). \quad (3.3)$$

For simplicity we assume that the composite good can be used for either addictive or consumption or investment.⁸

3.2 Households

A representative household derives utility from consumption of an ordinary (non-addictive) good, c_t , the fraction of time allocated to leisure, $1 - h_t \equiv l_t \in [0, 1]$,⁹ and effective consumption (i.e. consumption in excess of that required to sustain the addiction), s_t , of an addictive good, d_t .

We model addiction using the rational addiction framework of Becker and Murphy (1988). In this approach, consumption of the addictive good is specifically related to past consumption. Evidence for rational addiction exists for cigarettes,¹⁰ alcohol,¹¹ drugs (Grossman and

⁸Note that it is possible (but cumbersome) to extend the analysis to allow the production technology to differ by consumption goods.

⁹Endogenous labor supply allows the model to determine the effect of complementarities with leisure on taxation of addictive goods (see section 7.6).

¹⁰See for example Arcidiacono, Sieg, and Sloan (2007), Gruber and Koszegi (2001), Becker, Grossman, and Murphy (1994), Chaloupka (1991) and Sung, Hu, and Keeler (1994).

¹¹See for example, Grossman, Chaloupka, and Sirtalan (1998) Baltagi and Griffin (2002), Bentzen, Eriks-

Chaloupka 1998), food consumption (Carrasco, Labeaga, and Lopez-Salido 2005), carbonated drinks (Liu and Lopez 2009), and caffeine (Olekalns and Bardsley 1996). Dragone (2009) and Levy (2002) give a positive theory for the worldwide diffusion of obesity by proposing a rational addiction model for (junk) food intake behavior that is capable of displaying binges, diets and cycles for food consumption.

The main alternative, modeling addiction as either rational or irrational excess consumption, has intuitive appeal but also some drawbacks. First, in that framework the degree of excess consumption must be heterogeneous across the population. The optimal tax is sensitive to both the degree of excess consumption and the fraction of the population that suffers from excess consumption. Furthermore, time inconsistent preferences require separability in addictive and ordinary goods, no savings, and often *ad hoc* utility functions for tractability purposes. Since we show here that properties of the utility function such as homotheticity, tolerance, and homogeneity of the addiction function are crucial in establishing optimal taxation results, our modeling choice is ex-ante less restrictive than alternatives, especially if the government has a revenue raising motive.

Let $s_t = s(d_t, d_{t-1})$ map addictive consumption into effective consumption.¹² We assume throughout the paper that:

Assumption A2 *Per-period utility, $u(c_t, s_t, l_t)$, is strictly increasing, concave, and satisfies the Inada conditions in each argument.*

Lifetime utility is:

$$U = \sum_{t=0}^{\infty} \beta^t u(c_t, s_t, l_t); \quad (3.4)$$

where β is the discount factor with rate of time preference $\rho = \frac{1-\beta}{\beta}$.

For effective consumption, we assume throughout the paper that:

Assumption A3 *$s(d_t, d_{t-1})$ is homogeneous of degree α in $[d_t, d_{t-1}]$ ($HD-\alpha$) and satisfies $s_1 > 0$, $s_2 < 0$, $s_{11} \leq 0$, and $\sigma_s(c_t, s_t, l_t) \equiv \frac{-u_{ss}(c_t, s_t, l_t)s_t}{u_s(c_t, s_t, l_t)} \geq \frac{s_{22} \cdot s}{s_2^2}$.*

son, and Smith (1999), Baltagi and Geishecker (2006), and Waters and Sloan (1995).

¹²Here we follow the literature (e.g. Campbell and Cochrane 1999) and assume s has only one lag. This specification is clearly equivalent to Becker, Grossman, and Murphy (1994), who assume a utility function of the form $u(c_t, d_t, d_{t-1})$, except they assume no preferences for leisure. Our assumption below that s is homogeneous is the main restriction we impose on their utility specification. However, our specification is considerably more general than the addiction literature which typically assumes one of the two special cases for s given below.

The first inequality states that households get positive marginal utility from consumption of the addictive good. The second inequality states that the addictive good has the *tolerance* property, meaning past consumption lowers current utility, which is also known as a harmful addiction.¹³ The third and fourth inequalities are sufficient conditions which ensure that U is concave in the choice set $[c_t, l_t, d_t]$ if u is concave when $s_t = d_t$ (i.e. the standard problem with no addiction is concave, see assumption A2). The role of homogeneity is discussed below.

The two most commonly used specifications of effective consumption, s , in the literature are the *subtractive model* (see for example Campbell and Cochrane 1999, Chugh 2007), where effective consumption is:

$$s_t = d_t - \gamma d_{t-1}, \tag{3.9}$$

and the *multiplicative model* (see for example Abel 1990), which specifies effective consumption

¹³The term habit formation is often used in the macro literature, whereas addiction was introduced by BM. These two notions are equivalent under certain conditions. Gruber and Koszegi (2004) and others define habit formation as past consumption increasing the taste for current consumption. Therefore, a good is *habit forming* if and only if:

$$\frac{\partial^2 u}{\partial d_t \partial d_{t-1}} > 0. \tag{3.5}$$

From the assumptions on s , a good is habit forming if and only if:

$$\sigma_s(c_t, s_t, l_t) > \frac{s_t s_{12}(d_t, d_{t-1})}{s_1(d_t, d_{t-1}) s_2(d_t, d_{t-1})}. \tag{3.6}$$

BM and others define *addiction* as when past consumption increases current consumption, holding fixed prices and the marginal utility of ordinary consumption. Let $c_t = y_t - p_t d_t$, where y_t represents income in period t and p_t is the price of d in period t , then d is addictive if and only if:

$$\frac{\partial d_t}{\partial d_{t-1}} = \frac{\frac{\partial^2 U}{\partial d_t \partial d_{t-1}}}{-\frac{\partial^2 U}{\partial d_t^2}} > 0, \tag{3.7}$$

holding fixed the marginal utility of consumption. Using the concavity assumptions, equation (3.7) simplifies to:

$$\frac{\partial^2 U}{\partial d_t \partial d_{t-1}} = \frac{\partial^2 u}{\partial d_t \partial d_{t-1}} > 0. \tag{3.8}$$

Thus d is addictive if and only if d is habit forming given the one-lag specification of effective consumption, and the concavity assumptions. In general, if s has more than one lag, addiction is more restrictive than habit formation. Thus, for example, habit formation and addiction are not equivalent in BM, but are equivalent in Becker, Grossman, and Murphy (1994).

tion as:

$$s_t = \frac{d_t}{d_{t-1}^\gamma}. \quad (3.10)$$

In both models $\gamma \geq 0$ denotes the strength of tolerance. If $\gamma = 0$, then past consumption has no weight at all, in which case the model reduces to the standard time separable model, and utility is fully determined by consumption levels and not by the changes in consumption.

The subtractive specification satisfies assumption A3, and the multiplicative specification satisfies A3 for $\sigma_s(c_t, s_t, l_t) > (1 + \gamma)/\gamma \forall t$. However, two key differences exist. In the subtractive model, effective consumption is HD-1. In the multiplicative model, effective consumption is HD- $(1 - \gamma)$, and the degree of homogeneity depends on the degree of tolerance. Moreover, equation (3.8) implies that if s is subtractive, then d is addictive for all $\gamma > 0$. However, if s is multiplicative, then d is addictive if and only if $\sigma_s(c_t, s_t, l_t) > 1$ for all $[c_t, s_t, l_t]$.

Household Resources and Optimal Decisions

The household budget constraint sets after tax wage and rental income and government bond redemptions (equal to $R_t^b b_t$, where b_t are bonds issued in $t - 1$ and redeemed in t) equal to after tax expenditures on government bond issues and consumption of addictive, ordinary, and investment goods given by $i_t = k_{t+1} - (1 - \delta)k_t$, where δ is the depreciation rate. Since consumption of ordinary, addictive, and investment goods all have the same production technology, they have the same pre-tax price, which is normalized to one. Let τ_c and τ_d be the tax rates on consumption of ordinary and addictive goods, respectively and let τ_h be the tax rate on labor income. The household budget constraint is then:

$$R_t^b b_t + r_t k_t + (1 - \tau_{h,t}) w_t h_t = (1 + \tau_{c,t}) c_t + (1 + \tau_{d,t}) d_t + i_t + b_{t+1}. \quad (3.11)$$

In the next subsection we show that introducing a tax on capital would not change the results. Also, we will be more specific on the tax smoothing role played by government bonds. Let λ_t denote the Lagrange multiplier on the budget constraint. The resulting household first order conditions are:

$$(1 + \tau_{c,t}) \lambda_t = \beta^t u_c(c_t, s_t, l_t), \quad (3.12)$$

$$(1 - \tau_{h,t}) w_t \lambda_t = \beta^t u_l(c_t, s_t, l_t), \quad (3.13)$$

$$(1 + \tau_{d,t}) \lambda_t = \beta^t u_s(c_t, s_t, l_t) s_1(d_t, d_{t-1}) + \beta^{t+1} u_s(c_{t+1}, s_{t+1}, l_{t+1}) s_2(d_{t+1}, d_t), \quad (3.14)$$

$$\lambda_t R_t = \lambda_{t-1}, \quad t \geq 1, \quad (3.15)$$

$$\lambda_t R_t^b = \lambda_{t-1}, \quad t \geq 1, \quad (3.16)$$

$$R_t = r_t + 1 - \delta. \quad (3.17)$$

Equations (3.12)-(3.16), the budget constraint (3.11), initial conditions k_0 and d_{-1} , and the appropriate transversality conditions characterize the optimal household decisions k_t , b_t , h_t , c_t , d_t , and λ_t . In equation (3.14), the household increases effective consumption by increasing d_t (first term on the right hand side), but also increases tolerance and therefore reduces future effective consumption (second term on the right hand side). From equations (3.12) and (3.14) we have:

$$\frac{1 + \tau_{d,t}}{1 + \tau_{c,t}} = \frac{u_s(c_t, s_t, l_t) s_1(d_t, d_{t-1}) + \beta u_s(c_{t+1}, s_{t+1}, l_{t+1}) s_2(d_{t+1}, d_t)}{u_c(c_t, s_t, l_t)} \equiv \frac{MU_{d,t}}{MU_{c,t}}; \quad (3.18)$$

where $MU_{i,t}$ represents the marginal utility of good i at time t .

Any difference in tax rates drives a wedge between the marginal utilities of the consumption of ordinary and addictive goods. Thus the optimal tax rate of addictive goods exceeds the tax rate of ordinary consumption goods ($\tau_{d,t} > \tau_{c,t}$) if and only if $MU_{d,t} > MU_{c,t}$. The goal of this paper is to find conditions under which the marginal utility of addictive goods differs from that of ordinary goods.¹⁴

3.3 Government

The government finances an exogenous sequence of expenditures, g_t , with bond issues and consumption and labor income tax revenues. The government budget constraint is:

$$g_t = \tau_{h,t} w_t h_t + \tau_{c,t} c_t + \tau_{d,t} d_t + b_{t+1} - R_t^b b_t. \quad (3.19)$$

As will be clear below, three wedges exist in our model: one between the marginal utility

¹⁴Note that τ_c is defined as a tax on ordinary goods only, not all consumption goods. Thus, the ordinary consumption tax is not functionally equivalent to a wage tax, as would be the case with a single consumption good. Any change in the wage tax would equally raise the leisure cost of both types of consumption goods, and so does not affect the wedge.

of addictive and ordinary consumption, a second between the wage and the marginal rate of substitution between consumption and leisure, and a third between the intertemporal marginal rate of substitution and the rate of interest. Thus we need only three tax instruments for a complete tax system. We therefore set interest taxes equal to zero, noting that the government can affect all three margins by setting a time-varying consumption tax, a wage tax, and an addictive goods tax.¹⁵ To see this, use equation (3.12) to eliminate λ_t from the household first order conditions so that equation (3.15) becomes:

$$\left(\frac{1 + \tau_{c,t-1}}{1 + \tau_{c,t}}\right) R_t = \frac{u_{c,t-1}}{\beta u_{c,t}}. \quad (3.20)$$

Given a tax rate $\tau_{c,t-1}$, the planner can create an effective interest tax by varying $\tau_{c,t}$, while using $\tau_{d,t}$ to alter the wedge between the marginal utility of addictive and ordinary consumption, and $\tau_{h,t}$ to alter the wedge between the wage and the marginal rate of substitution between consumption and leisure. Alternatively, suppose we remove the tax on consumption and add a tax τ_k on interest so that (3.20) becomes:

$$(1 - \tau_{k,t}) R_t = \frac{u_{c,t-1}}{\beta u_{c,t}}. \quad (3.21)$$

Now the planner can equivalently use $\tau_{k,t}$ to vary the wedge between the intertemporal marginal rate of substitution and the interest rate, while again using $\tau_{d,t}$ and $\tau_{h,t}$ to vary the other wedges. Thus, the optimal allocations derived below can be decentralized using either tax system.¹⁶

As is common in the literature (see for example Chamley 1986), the government optimally uses bonds to smooth tax burdens over time. In the absence of bonds, the government may favor the tax with better smoothing properties.¹⁷ Changes in current addictive goods tax rates affects both current and future tax revenue. The existence of government bonds enables us to conveniently summarize the effect of a change in current addictive tax rates on all

¹⁵We also do not allow a tax on effective consumption, since informational asymmetries rule out taxes on effective consumption in practice.

¹⁶Although labor taxes do not affect the wedge between the marginal utilities of ordinary and addictive consumption, they play an important role in the analysis in that without the tax the tax system would be incomplete and the government would be forced to have the same wedge for consumption/savings as labor/leisure. The government would then have to consider the effect of addictive taxes on this combined wedge, considerably complicating the analysis in a way that is unlikely to matter in practice (state and local governments can equivalently affect all three wedges using property (capital), addictive, and ordinary consumption taxes).

¹⁷However, addictive taxes are common at the state and local level, which frequently have constitutional borrowing restrictions. We leave this interesting case to future research.

periods as the effect on the infinite horizon version of the government's budget constraint.

Let $\pi = [(\tau_{c,t})_{t=0}^{\infty}, (\tau_{d,t})_{t=0}^{\infty}, (\tau_{h,t})_{t=0}^{\infty}, (g_t)_{t=0}^{\infty}]$ denote an infinite sequence of government policies. As is standard in the literature (e.g. Gruber and Koszegi 2001), we assume the existence of a commitment technology, so that the government commits to all future policies at time zero.¹⁸

4 Equilibrium and Ramsey Problem

Equations (3.2), (3.3), (3.11), (3.12) - (3.15), and (3.19) form a system of nine nonlinear equations that characterize the competitive equilibrium. Hence:

Definition 1 *Given initial values k_0 and d_{-1} , a competitive equilibrium is a set of allocations $\{c_t, d_t, h_t, k_t\}$, prices $\{w_t, r_t, R_t^b\}$ and a sequence of policies π that satisfy the household budget constraint (3.11), firm profit maximization (3.1), the government budget constraint (3.19), and household maximization of (3.4) for all t .*

We use the primal approach to determine optimal taxation. The primal approach uses household and firm first order conditions to eliminate prices and policies from the equations that define the competitive equilibrium. The planner then chooses allocations which maximize welfare subject to the remaining equations from the competitive equilibrium. These equations are the resource constraint:

$$F(k_t, h_t) = c_t + d_t + k_{t+1} - (1 - \delta)k_t + g_t, \quad (4.1)$$

and the *implementability constraint* (IMC):

$$\sum_{t=0}^{\infty} \beta^t \left\{ u_c(c_t, s_t, l_t) c_t + \left[u_s(c_t, s_t, l_t) s_1(d_t, d_{t-1}) + \beta u_s(c_{t+1}, s_{t+1}, l_{t+1}) \cdot s_2(d_{t+1}, d_t) \right] d_t - u_l(c_t, s_t, l_t) h_t \right\} - \frac{u_c(c_0, s_0, l_0) (R_0 k_0 + R_0^b b_0)}{1 + \tau_{c,0}} = 0. \quad (4.2)$$

The IMC uses the household first order conditions to substitute out for all prices and policies in the budget constraint and then recursively eliminates λ_t . Thus, the IMC is the infinite horizon version of the household budget constraint where all prices and policies have been written in terms of their corresponding marginal utilities. It is immediate from Walras' Law and the resource constraint that the IMC can also be thought of as the infinite horizon version

¹⁸In principle the government could promise low future taxes on addictive goods, and then find it optimal to renege on the promise once households become addicted.

of the government budget constraint. The Ramsey approach is therefore very convenient in that the planner can, through the IMC, determine the effect of a change in d_t on government revenues over the infinite horizon.

Let IMC with a variable subscript denote the partial derivative of IMC with respect to that variable, and think of IMC as the infinite horizon version of the government budget constraint. The marginal tax revenue of ordinary consumption depends on how a small change in ordinary consumption affects ordinary consumption tax revenue, addictive tax revenue, and labor income tax revenue:

$$\frac{IMC_{c_t}}{\beta^t} = u_{c,t} + u_{cc,t}c_t + \alpha u_{cs,t}s_t - u_{cl,t}h_t. \quad (4.3)$$

An increase in c_t directly increases ordinary consumption tax revenue (first term), but decreases the marginal utility of consumption and thus requires the planner to lower the ordinary consumption tax rate in order to maintain equilibrium, which lowers tax revenues (second term). The third term contains two offsetting effects. Suppose for example that $u_{cs} > 0$. Then an increase in c_t raises $MU_{d,t}$, so the planner must raise the tax on d_t to maintain equilibrium, which increases addictive tax revenues. However, an increase in c_t with $u_{cs} > 0$ lowers $MU_{d,t-1}$: consuming d_{t-1} is less attractive because it causes effective consumption to fall in t (tolerance), when the marginal utility of effective consumption is relatively high due to the increase in c_t . Thus, the planner must also lower $\tau_{d,t-1}$, reducing revenues. Thus an increase in c_t has offsetting addictive tax revenue effects, but both work through the u_{cs} term. Given the homogeneity assumption, these two dynamic effects can be combined into a single effect, as if a smaller tax on s_t , rather than d_t , existed. Finally, the fourth term implies that an increase in c_t increases preferences for leisure, and thus causes the planner to decrease the labor income tax rate to maintain equilibrium, if and only if $u_{cl} > 0$.

The marginal tax revenue of addictive goods depends on revenue effects in periods $t - 1$, t , and $t + 1$:

$$\begin{aligned} \frac{IMC_{d_t}}{\beta^t} = & (u_{s,t}s_{12,t} + u_{ss,t}s_{1,t}s_{2,t})d_{t-1} + MU_{d,t} + (u_{s,t}s_{11,t} + u_{ss,t}s_{1,t}^2 + \\ & \beta u_{s,t+1}s_{22,t+1} + \beta u_{ss,t+1}s_{2,t+1}^2) d_t + (\beta u_{s,t+1}s_{12,t+1} + \beta u_{ss,t+1}s_{1,t+1}s_{2,t+1}) d_{t+1} \\ & + u_{cs,t}s_{1t}c_t + \beta u_{cs,t+1}s_{2,t+1}c_{t+1} - u_{sl,t}s_{1,t}h_t - \beta u_{sl,t+1}s_{2,t+1}h_{t+1}. \end{aligned} \quad (4.4)$$

Equation (4.4) shows an increase in d_t affects addictive tax revenue in periods $t - 1$ (first term on the right hand side), t (second and third terms), and $t + 1$ (fourth term). Ordinary consumption and labor tax revenues are also affected in periods t and $t + 1$ (remaining terms).

Using the homogeneity of s and its first derivatives, equation (4.4) simplifies to:

$$\frac{IMC_{d_t}}{\beta^t} = \alpha MU_{d,t} + \alpha u_{ss,t} s_{1,t} s_t + \alpha \beta u_{ss,t+1} s_{2,t+1} s_{t+1} + u_{cs,t} s_{1,t} c_t + \beta u_{cs,t+1} s_{2,t+1} c_{t+1} - u_{sl,t} s_{1,t} h_t - \beta u_{sl,t+1} s_{2,t+1} h_{t+1}. \quad (4.5)$$

The first term on the right hand side shows that many of the dynamic effects on addictive taxation in periods $t - 1$, t , and $t + 1$ are equivalent to smaller revenue effects in period t . The dynamic revenue effects in $t - 1$ and $t + 1$ reduce the revenue raised from a small increase in d_t from $MU_{d,t}$ in equation (4.4) to $\alpha MU_{d,t}$. The revenue effects on addictive taxes in $t - 1$ and $t + 1$ tend to offset the revenue gains in t .

The remaining terms in equation (4.5) show that the planner generally must adjust current and future tax rates differently to maintain equilibrium. A small increase in d_t reduces the marginal utility of effective consumption, which requires the planner to decrease $\tau_{d,t}$ to maintain equilibrium, reducing revenues (second term on the right hand side of 4.5). An increase in d_t increases the marginal utility of ordinary consumption and thus increases $\tau_{c,t}$ if and only if $u_{cs} > 0$. An increase in d_t increases the marginal utility of leisure and thus decreases $\tau_{h,t}$ if and only if $u_{sl} > 0$. An increase in d_t also increases tolerance in period $t + 1$, thus reducing s_{t+1} . Thus an increase in d_t affects all three types of tax revenue in period $t + 1$ as well but *in the opposite direction*.

The first proposition gives the relationship between the competitive equilibrium and the IMC and resource constraint.

PROPOSITION 1 *Let assumptions (A1)-(A3) hold. Given k_0 , d_{-1} , $\tau_{h,0}$, and $\tau_{c,0}$, the allocations of a competitive equilibrium satisfy (4.1) and (4.2). In addition, given k_0 , d_{-1} , $\tau_{h,0}$, and $\tau_{c,0}$, and allocations which satisfy (4.1) and (4.2), prices and policies exist which, together with the allocations, are a competitive equilibrium.*

All proofs are in the appendix.

The Ramsey Problem (RAM) determines the optimal tax package that maximizes welfare subject to the IMC and resource constraint:

$$\begin{aligned} \text{RAM} = \max_{c_t, d_t, h_t, k_t} & \left\{ \sum_{t=0}^{\infty} \beta^t \left(u(c_t, s_t, 1 - h_t) + \mu \left[u_c(c_t, s_t, 1 - h_t) c_t + \right. \right. \right. \\ & (u_s(c_t, s_t, 1 - h_t) s_1(d_t, d_{t-1}) + \beta u_s(c_{t+1}, s_{t+1}, l_{t+1}) s_2(d_{t+1}, d_t)) d_t - \\ & \left. \left. \left. u_l(c_t, s_t, 1 - h_t) h_t \right] \right) - \mu \frac{u_c(c_0, s_0, l_0) (R_0 k_0 + R_0^b b_0)}{1 + \tau_{c,0}} + \right. \end{aligned}$$

$$\phi_t \left[F(k_t, h_t) - c_t - d_t - k_{t+1} + (1 - \delta)k_t - g_t \right] \}. \quad (4.6)$$

The Ramsey problem consists in maximizing welfare subject to the IMC and resource constraints. In (4.6), μ and ϕ_t are the Lagrange multipliers on the IMC and resource constraints, respectively.

The first order conditions that characterize optimal taxation are:

$$\phi_t = \beta^t MU_{c,t} + \mu IMC_{c_t}, \quad (4.7)$$

$$\phi_t = \beta^t MU_{d,t} + \mu IMC_{d_t}, \quad (4.8)$$

$$\phi_t F_h(k_t, h_t) = \beta^t u_l(c_t, s_t, 1 - h_t) - \mu IMC_{h_t}, \quad (4.9)$$

$$\beta \phi_t (F_k(k_t, h_t) + 1 - \delta) = \phi_{t-1}. \quad (4.10)$$

Conditions (4.7)-(4.9) equate the marginal social welfare of c , d , and l with the marginal resource cost ϕ . Equation (4.10) equates the return on capital with the intertemporal marginal rate of substitution.

From equations (4.7) and (4.8):

$$\beta^t (MU_{c,t} - MU_{d,t}) = \mu (IMC_{d_t} - IMC_{c_t}). \quad (4.11)$$

Hence using equation (3.18), we find that addictive goods are taxed at a higher rate than ordinary goods if and only if the derivative of the IMC with respect to d_t is smaller than the derivative with respect to c_t :

$$\tau_{d,t} > \tau_{c,t} \text{ iff } IMC_{d_t} < IMC_{c_t}. \quad (4.12)$$

Since the marginal rate of transformation between c and d is one, the marginal rate of substitution minus one is the tax wedge. The Ramsey problem computes the optimal wedges between marginal utilities as:

$$\frac{MU_{d,t}}{MU_{c,t}} = 1 + \text{Wedge} = 1 + \frac{\mu}{\beta^t MU_{c,t}} (IMC_{c_t} - IMC_{d_t}). \quad (4.13)$$

From equation (4.11), if $MU_{d,t} > MU_{c,t}$, reallocating a marginal resource from ordinary to addictive consumption raises welfare by $MU_{d,t} - MU_{c,t}$. Thus, tax revenue must fall by $IMC_{c_t} - IMC_{d_t}$, resulting in a loss of welfare of μ times the loss of tax revenue. Hence addictive goods are taxed at a higher rate than ordinary goods if and only if moving a resource unit from addictive to ordinary consumption raises revenue, that is, if the marginal tax revenue of ordinary goods exceeds that of addictive goods. Equivalently, addictive goods are taxed at a higher rate if and only if the welfare distortion induced by addictive taxation is less than that of ordinary taxation.

In summary then, simple static results and intuition might indicate that taxing addictive goods is a good revenue raiser because addictive goods tend to be complementary to leisure. However, dynamic considerations are likely to yield more moderate results. To see this consider, for example, the last two terms of (4.5). If $u_{sl} > 0$, then a decrease in d_t raises labor income tax revenues in period t (second to last term), but increases s_{t+1} , reducing labor income tax revenues in period $t + 1$ (last term). Thus the dynamic and static terms offset (a similar intuition holds for the other terms). In addition, the stronger the tolerance, the stronger is the dynamic effect. To obtain further results requires more specific preference assumptions.

5 Results for General Preference Classes

As in the literature on optimal commodity taxation, characteristics of the utility function play an important role in determining any deviations from uniform taxation. In this section we first establish theoretical results for the well-known case of homothetic utility; we then proceed to numerically investigate the linear quadratic utility case. In the appendix we study the separable utility function and a utility function which is not weakly separable in leisure.

5.1 Homothetic Utility with weak separability of leisure

In this section we assume the utility function takes the form:

$$u(c_t, s_t, l_t) = q(v(c_t, s_t), l_t), \quad (5.1)$$

where $v(\cdot)$ is homothetic and $q(\cdot)$ is an increasing function. Weak separability in leisure implies that the more income inelastic good is also more complementary with leisure. Further, homotheticity implies c and s are equally income elastic, and thus equally substitutable with leisure. Thus, homotheticity and weak separability imply uniform taxation of c and s . To

what extent does this result change for the more realistic case where d is taxed rather than s ?

To find out, we combine equations (4.12), (4.3), and (4.5), assuming (5.1). Let us define the following elasticities:

$$\sigma_{cs,t} \equiv \frac{u_{cs}(c_t, s_t, l_t) c_t}{u_s(c_t, s_t, l_t)}, \quad \sigma_{sc,t} \equiv \frac{u_{cs}(c_t, s_t, l_t) s_t}{u_c(c_t, s_t, l_t)}, \quad \sigma_{hs,t} \equiv \frac{u_{sl}(c_t, s_t, l_t) h_t}{u_s(c_t, s_t, l_t)}, \quad (5.2)$$

and let $\sigma_{hc,t}$ be defined analogously. Then we have the following result.

PROPOSITION 2 *Let assumptions (A1)-(A3) hold. In addition, let $u(\cdot)$ be of the form given in equation (5.1). Then $\tau_{d,t} > \tau_{c,t}$ if and only if:*

$$(1 - \alpha)(1 - \sigma_{s,t} - \sigma_{sc,t}) > J(\alpha \Delta \sigma_s - \Delta \sigma_{cs} + \Delta \sigma_{hs}), \quad (5.3)$$

where

$$J \equiv \frac{-\beta u_{s,t+1} s_{2,t+1}}{MU_{d,t}}, \quad \Delta \sigma_i \equiv \sigma_{i,t+1} - \sigma_{i,t}, \quad i = s, cs, hs.$$

Homotheticity of c and s does not result in uniform taxation of addictive and ordinary goods. Three effects determine the intuition behind condition (5.3). First households anticipate higher taxes in period t and respond by changing behavior in period $t - 1$, changing tax revenue in period $t - 1$. We call this the *tax anticipation effect*.¹⁹ Second, households are less addicted in period $t + 1$ which changes tax revenue in period $t + 1$. We call this the *addiction stock effect*. Third, homotheticity in c and s does not necessarily imply homotheticity in c and d , the goods which are actually taxed. To see the dynamic effects at play in isolation, first assume $s_{1,t} d_t = s_t$ and $s_{11,t} = 0$ so that homotheticity of c and s is equivalent to homotheticity of c and d . Then only the *tax anticipation effect* and the *addiction stock effect* cause deviations from uniform taxation. Note that these conditions imply that (3.6) simplifies to $\sigma_{s,t} > 1$.

Next, suppose households do not consider in period t that their addictive consumption affects their future addiction stocks. However, continue to assume households anticipate tax increases in $t - 1$. That is, assume all terms multiplied by β in equation (4.5) equal

¹⁹Strictly speaking, given no uncertainty, all addictive consumption levels over time are endogenously determined by the planner through the tax rates at time 0, and the first order condition for d_t is written with all other values of d_t over time held at their optimal levels. Therefore, the terms multiplied by d_{t-1} and d_{t+1} in equation (4.5) indicate that a change in the tax rate in t affects incentives in $t - 1$ and $t + 1$, forcing the planner to alter the tax rates in $t - 1$ and $t + 1$ to maintain equilibrium with the same optimal values of d_{t-1} and d_{t+1} . The change in tax rates is what affects revenues in $t - 1$ and $t + 1$. Rates in other periods are unaffected since the economy remains on the same equilibrium path.

zero and $J = 0$ in condition (5.3). Thus, the left hand side represents the change in the optimal tax rate caused by households being able to anticipate higher tax rates in period t . By condition (3.6), the left hand side is negative, which favors taxing c_t at a higher rate, unless σ_{sc} is sufficiently negative. Households anticipate that an increase in taxes in period t will reduce addictive consumption in period t . This reduces the marginal utility of addictive consumption in $t - 1$ as addictive consumption in $t - 1$ will raise the stock of addiction in t , reducing effective consumption when it is already low due to the tax. Therefore households desire to reduce addictive consumption in $t - 1$, so the planner must decrease the addictive tax rate in $t - 1$ to maintain equilibrium, reducing revenues gained from raising the tax in t . Thus, the *tax anticipation effect* tends to reduce optimal addictive tax rates. Considering only the tax anticipation effect, the optimal addictive tax is less than the ordinary consumption tax unless raising addictive taxes generates enough consumption tax revenue (σ_{sc} is sufficiently negative).

The left hand side of condition (5.3) also shows that the dynamic effects in $t - 1$ and t reduce to a single current period term which depends only on the elasticities and homogeneity, and should therefore be straightforward to check in empirical applications.

To see the addiction stock effect, continue to assume $s_{1,t}d_t = s_t$ and $s_{11,t} = 0$, but now assume that the terms multiplied by d_{t-1} in equation (4.5) and (4.3) are zero. Condition (5.3) then reduces to:

$$0 > J(1 - \alpha + \alpha\Delta\sigma_s - \Delta\sigma_{cs} + \Delta\sigma_{hs}). \quad (5.4)$$

The difference between the right hand side of (5.3) and (5.4) reflects that condition (5.3) assumes homotheticity in c and s whereas (5.4) assumes homotheticity in c and d . The right hand side of condition (5.4) consists of terms which reflect the change in all three types of tax revenues in periods t and $t + 1$ due to a change in the addiction stock. The period t and $t + 1$ effects tend to offset. For example, consider the impact of an increase in the addictive tax in period t , which decreases d_t and thus s_t . If $\sigma_{hs,t} > 0$, the decrease in s_t causes households to work more, which means the planner must raise the labor tax rate to maintain equilibrium, increasing labor tax revenues. This makes taxing d_t more attractive. However, because the household will be less addicted in $t + 1$, s_{t+1} rises, inducing the opposite effect in $t + 1$ assuming $\sigma_{hs,t+1} > 0$. A similar intuition holds for the other taxes.

In general the addiction stock effect implies addictive goods should be taxed at a higher rate when responses to tax changes are becoming more elastic over time. Suppose that σ_{hs} is falling, so that $J \cdot \Delta\sigma_{hs}$ is becoming more negative. Then taxing d_t raises labor tax revenues

today, but decreases labor tax revenues less in $t + 1$, when σ_{hs} is lower. Because the hours response is more elastic in period $t + 1$, the planner must raise the tax rate by a relatively small amount in $t + 1$ to maintain equilibrium, whereas in period t the planner must lower the tax rate by a relatively large amount. In this case, taxing addictive goods is more attractive than taxing ordinary consumption goods. Optimal addictive taxation smoothes distortions, because taxing addictive goods now makes taxing addictive goods more distortionary later. As shown in the following propositions, however, some common specifications for v and s induce constant elasticities which imply uniform taxation.

PROPOSITION 3 *Let the conditions of Proposition 2 hold, and let $u(.) = z(l) + \frac{(c^\xi s^{1-\xi})^{1-\sigma} - 1}{1-\sigma}$, and $z(.)$ be concave, then $\tau_d = \tau_c$ for all t .*

Although we have assumed here that $v(.)$ is constant relative risk aversion (CRR), this proposition is considerably more realistic than the existing literature which assumes a static utility function and/or separable quadratic utility for tractability. If utility is CRR in c and s and separable in l , then we obtain the classic result of uniform commodity taxation as in Atkinson and Stiglitz (1972).

Elasticities are constant in the steady state, which eliminates the addiction stock effect. Let \bar{y} denote the steady state value of any variable y , then:

PROPOSITION 4 *Let the conditions of Proposition 2 hold. Then $\bar{\tau}_d > \bar{\tau}_c$ if and only if:*

$$(1 - \alpha)(1 - \bar{\sigma}_s - \bar{\sigma}_{sc}) > 0. \quad (5.5)$$

In the steady state with homotheticity in c and s , an increase in addictive tax revenues in t has an equal and opposite decrease in addictive tax revenues in $t + 1$, eliminating the addiction stock effect. In contrast, the tax anticipation effect depends on the homogeneity of the addiction function. For HD-1 addiction functions, including the subtractive model, proposition 4 indicates that the steady state tax rates are uniform. For the multiplicative case, the degree of homogeneity is decreasing in the strength of tolerance. Therefore, an increase in tolerance can cause a decrease in the steady state addictive tax rate if:

$$\bar{\sigma}_s - 1 < -\bar{\sigma}_{sc}. \quad (5.6)$$

If condition (5.6) holds the steady state tax anticipation effect is negative: high tolerance given the multiplicative model means that households anticipate higher tax rates by strongly decreasing $t - 1$ addictive consumption, reducing addictive tax revenue enough to offset the

gain in period t revenue, as well as any gains arising from increased ordinary consumption in t .

Proposition 4 and Proposition 2 indicate that the choice of addiction function is not innocuous when designing optimal tax policies. In particular, condition (5.5), implies uniform steady state taxation for the subtractive model, but not necessarily for the multiplicative model.

5.2 An Analytical Example: The Quadratic Case

In this section, we consider a linear-quadratic utility function. The linear-quadratic utility, a common specification in the literature (e.g., Becker, Grossman, and Murphy 1994, Gruber and Koszegi 2001), offers several advantages. First, we obtain an analytic solution, which allows us to study how the dynamics of addictive consumption change optimal taxation. Second, since this specification has no income effects to complicate the dynamics, we can derive a more precise relationship between tolerance and addictive taxation. In the appendix we consider classes of utility functions with non-trivial income and labor supply effects.

Suppose the subtractive specification (3.9) for effective consumption and that the utility and production function are:

$$u(c_t, s_t, l_t) = \omega c_t + \nu s_t - \frac{s_t^2}{2} + e l_t - \frac{l_t^2}{2}, \quad e < 1, \nu > \frac{\omega}{1 - \beta\gamma}, \quad (5.1)$$

$$F(k_t, h_t) = k_t^\theta h_t^{1-\theta}. \quad (5.2)$$

Here, the assumption on e ensures positive steady state hours and the assumption on ν is required for positive steady state addictive consumption (see Proposition 6).

Inspection of equations (4.7) and (4.8), given the utility function (5.1), reveals that the marginal utility of d_t is constant in the optimal second best allocation. In particular:

$$MU_{d,t} = \frac{\omega + \mu(\omega + \nu(1 - \beta\gamma))}{1 + 2\mu}. \quad (5.3)$$

The marginal utility of d_t divided by $MU_{c,t} = \omega$ equals the tax ratio given by equation (3.18). Hence the tax ratio is constant over time. Furthermore, inspection of equations (3.12), (3.15), (4.7), and (4.10) indicates that τ_c is constant over time. Therefore, τ_d and τ_h are also constant over time. Thus the implicit interest tax rate is zero for all t .

Equation (5.3) implies $MU_{d,t} > MU_{c,t}$ and thus $\tau_{d,t} > \tau_{c,t}$ for all t . Hence we have shown:

PROPOSITION 5 *Let $u(.,.,.)$ and $F(.,.)$ be given by equations (5.1) and (5.2) and let effective consumption be given by the subtractive model. Then $\tau_{d,t} > \tau_{c,t}$ for all t and the ratio of tax rates $\frac{1+\tau_{d,t}}{1+\tau_{c,t}}$ is constant over time.*

In the static version of the model without addiction, $d_t = s_t$ has an income elasticity equal to zero whereas the income elasticity of c_t is positive. Further, ordinary consumption and leisure are substitutes, whereas $\frac{\partial s}{\partial w} \frac{w}{s} = 0$. Thus, it is optimal to tax d_t at a higher rate because, regardless of k_t or d_{t-1} , c is more substitutable with leisure.

It is also clear from equation (5.3) that the second best optimal d_t is the solution to a linear second order difference equation,

$$-\beta\gamma d_{t+1} + (1 + \beta\gamma^2) d_t - \gamma d_{t-1} = \frac{1 + \mu}{1 + 2\mu} (\nu(1 - \beta\gamma) - \omega), \quad (5.4)$$

and that the second best optimal s_t is the solution to a linear first order difference equation. However, before computing the solution to d_t , we must verify that a solution exists for μ . In appendix 7.4, we prove that a unique, positive solution exists if government spending is not so large as to exhaust the maximum feasible revenue in the economy, and not so small that given initial tax rates are sufficient to pay for all current and future government expenditures.

Given a unique solution for μ , d_t is the solution to the second order difference equation (5.3).

PROPOSITION 6 *Let the conditions for Proposition 5 hold. Then the explicit solution for d_t is:*

$$d_t = \frac{\nu(1 - \beta\gamma) - \omega}{(1 - \gamma)(1 - \beta\gamma)} \left(\frac{1 + \mu}{1 + 2\mu} \right) (1 - \gamma^{t+1}) + d_{-1}\gamma^{t+1}. \quad (5.5)$$

The solution for d_t , given by equation (5.5), allows us to derive some interesting properties of the second best solution, both over time and as compared to the first best solution ($\mu = 0$). First, optimal consumption of d_t increases over time, assuming d_{-1} is less than the steady state. The planner decreases d_t relative to the first best solution through the tax. However, since $0 < \mu < \infty$, equation (5.5) implies d_t in the second best optimum is at least half of the first best level in the steady state. The planner also decreases the growth rate of d_t since:

$$gr_t = \frac{d_t - d_{t-1}}{d_{t-1}} = \frac{\gamma^t(1 - \gamma)(\hat{d}x - d_{-1})}{\hat{d}x(1 - \gamma^t) + \gamma^t d_{-1}}, \quad \hat{d} \equiv \frac{\nu(1 - \beta\gamma) - \omega}{(1 - \gamma)(1 - \beta\gamma)}, \quad x \equiv \frac{1 + \mu}{1 + 2\mu}, \quad (5.6)$$

which is decreasing in μ because:

$$\frac{\partial gr_t}{\partial \mu} = -\frac{\gamma^t (1 - \gamma) d_{-1}}{d_{t-1}^2 (1 + 2\mu)^2} < 0. \quad (5.7)$$

We can also explore how the strength of tolerance affects second best addictive consumption. Since the solution for μ is unique, we can use the implicit function theorem to derive comparative statics using equation (5.3). Our intuition is that strong tolerance should moderate the optimal tax ratio, as gains in current tax revenue from taxation of addictive goods are offset by losses in future tax revenues. If d_{-1} is sufficiently large, it is indeed true that the optimal tax ratio is inversely related to the degree of tolerance. In particular, we have:

$$d_{-1} > \frac{\beta\omega}{(1 - \beta\gamma)(1 - \beta)} \Rightarrow \frac{\partial \frac{1+\tau_d}{1+\tau_c}}{\partial \gamma} < 0. \quad (5.8)$$

Condition (5.8), derived in appendix 7.4, is a sufficient condition calculated assuming $\mu = \infty$. In practice, for μ small, the optimal tax ratio is decreasing in the degree of tolerance under much less restrictive conditions.

Table 1 gives parameter values for a numerical example. Table 1 indicates that the optimal tax ratio is decreasing in the degree of tolerance, even though condition (5.8) is violated, since the parameter G , set to 30% of GDP for all t , generates at most a value of only $\mu = 4.74$. The planner relies increasingly on labor taxes and less on addictive taxes as the degree of tolerance increases. For $\gamma = 0.55$, taxation is nearly uniform. Figure 1 shows the time path of the first and second best levels of d for various values of γ . Increasing the level of tolerance severely reduces addictive consumption since the future costs of current consumption are higher. As expected, the difference between first and second best addictive consumption is widest at the steady state.

6 Conclusions

This paper is the first attempt in the literature to characterize and analyze the conditions under which taxation of addictive goods might differ from taxes on labor and ordinary consumption goods in a dynamic rational addiction setting. We derive a rich set of results. In particular, we derive conditions for a variety of classes of utility functions for which tax rates for addictive goods exceeds tax rates for ordinary goods in an environment where exogenous government spending cannot be financed with lump sum taxes.

We find two effects which differentiate taxation of addictive goods from ordinary goods.

First a tax anticipation effect exists in that households anticipate future increases in addictive tax rates and reduce addictive consumption, decreasing addictive tax revenues.²⁰ The tax anticipation effect tends to moderate the optimal tax. Second, an addiction stock effect exists. Current addictive taxes change the future tax base. Households will be less addicted in the future following an addictive tax increase. Thus future addictive, ordinary, and labor tax revenues are affected by a change in the current addictive tax rate. Whether the addiction stock effect increases revenues or not depends on whether or not elasticities are rising or falling. In the steady state, or with CRR utility, elasticities are constant and thus the addiction stock effect vanishes. Surprisingly, the tax anticipation effect makes taxing addictive goods in a dynamic setting equivalent to taxing effective consumption, in which the addictive good is less complementary with leisure to a degree which depends on the homogeneity of the addiction function. Therefore, classic results on uniform taxation re-emerge in some special cases such as CRR utility.

We consider homogeneous addiction functions, an improvement over the literature which typically assumes either subtractive or multiplicative addiction functions. Higher tolerance strengthens the tax anticipation effect and therefore reduces the optimal addictive tax. This result is quite surprising as the textbook intuition suggests that stronger addiction would make taxing addictive goods more attractive.

We also consider features such as complementarity to leisure that, while unrelated to addiction itself, are nonetheless common among addictive goods. In general, such effects are weaker in our dynamic setting. Examples are constructed where failing to account for the dynamic effects results in taxing addictive goods at a higher rate, when the optimal addictive tax is actually less than that of ordinary goods, due to the tax anticipation effect.

Our results come with a few caveats. First, one common feature of addictive goods, the presence of externalities, has not been considered in this paper. However, it is unclear whether externalities are not better dealt with by regulating the exact behavior that causes the externality, such as banning smoking in public places, rather than the second best solution of taxing consumption of the addictive good. In addition, if we include an externality, then our results go through under slightly different conditions. That is, our results can be interpreted as relative to the tax ratio which corrects the externality.

Second, we consider only the optimal tax package, not the optimal addictive goods tax taking as given other taxes. Certain features of the tax code such as balanced budget rules and positive capital tax rates, if added as extra constraints, may change our results. Third,

²⁰Some evidence exists that households do anticipate changes in addictive tax rates, see footnote 11.

we have no heterogeneity in addictive consumption or wealth. However, if we assume that the poor are more likely to consume addictive goods, then our results would likely strengthen, because consumers of addictive goods would have a higher marginal utility of income. One could also introduce heterogeneity via an overlapping generations model. Although beyond the scope of the present paper, it would seem plausible that the tax anticipation effect would be weaker for the young since they are not initially addicted. Conversely, the addiction stock effect will be weaker for the old.

Finally, by choosing to adopt the Ramsey approach, we are subject to the typical criticisms made to this framework: it takes the set of possible fiscal instruments as given and it requires linear taxes. Our results might change if we allowed for heterogeneous households and nonlinear tax rates within a Mirrless framework. However, while capital and income taxes are nonlinear in the data, addictive taxes are typically linear. In addition, it would be very hard, in particular at the state level, to monitor addictive consumption for each household. Most importantly, the majority of redistributive fiscal policy is carried out at the federal level in the US. Excise taxes at the state level are not typically set with a redistributive intent, as in a Mirrless framework, but rather to generate fiscal revenues and balance the state budgets. We leave the interesting theoretical issues concerning taxation of addictive goods in the Mirrless framework for further research as well.

References

- Abel, A., 1990, "Asset Prices Under Habit Formation and Catching up with the Joneses," *American Economic Review*, 80, 38–42.
- Anderson, J., 2005, "Casino Taxation in the United States," *National Tax Journal*, 58, 303–24.
- Arcidiacono, P., H. Sieg, and F. Sloan, 2007, "Living Rationally Under the Volcano? An Empirical Analysis of Heavy Drinking and Smoking," *International Economic Review*, 48, 37–65.
- Armenter, R., 2008, "A Note on Incomplete Factor Taxation," *Journal of Public Economics*, 92, 2275–81.
- Atkinson, A. B., and J. E. Stiglitz, 1972, "The Structure of Indirect Taxation and Economic Efficiency," *Journal of Public Economics*, 1, 97–119.
- Baltagi, B., and I. Geishecker, 2006, "Rational Alcohol Addiction: Evidence from the Russian Longitudinal Monitoring Survey," *Health Economics*, 15, 893–914.
- Baltagi, B., and J. Griffin, 2002, "Rational Addiction to Alcohol: Panel Data Analysis of Liquor Consumption," *Health Economics*, 11, 485–91.
- Becker, G., M. Grossman, and K. Murphy, 1994, "An Empirical Analysis of Cigarette Addiction," *American Economic Review*, 84, 396–418.
- Becker, G., and K. Murphy, 1988, "A Theory of Rational Addiction," *Journal of Political Economy*, 96, 675–700.
- Bentzen, J., T. Eriksson, and V. Smith, 1999, "Rational Addiction and Alcohol Consumption: Evidence from Nordic Countries," *Journal of Consumer Policy*, 22, 257–79.
- Boudreaux, D., and A. Pritchard, 1994, "Price of Prohibition, The," *Arizona Law Review*, 36, 1–10.
- Bovenberg, A., and L. H. Goulder, 1996, "Optimal Environmental Taxation in the Presence of Other Taxes: General-Equilibrium Analyses," *American Economic Review*, 86, 985–1000.
- Campbell, J., and J. Cochrane, 1999, "By Force of Habit: a Consumption Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107, 205–51.
- Carrasco, R., J. Labeaga, and D. Lopez-Salido, 2005, "Consumption and Habits: Evidence from Panel Data," *Economic Journal*, 115, 144–65.
- Chaloupka, F., 1991, "Rational Addictive Behavior and Cigarette Smoking," *Journal of Political Economy*, 99, 722–42.

- Chamley, C., 1986, "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives," *Econometrica*, 54, 607–622.
- Chari, V. V., and P. J. Kehoe, 1998, "Optimal Fiscal and Monetary Policy," Discussion Paper 251, Federal Reserve Bank of Minneapolis Research Department Staff Report.
- Chugh, S., 2007, "Optimal Inflation Persistence: Ramsey Taxation with Capital and Habits," *Journal of Monetary Economics*, 54, 1809–36.
- Clotfelter, C., and P. Cook, 1990, "On the Economics of State Lotteries," *Journal of Economic Perspectives*, 4, 105–119.
- Corlett, W., and D. Hague, 1953, "Complementarity and the Excess Burden of Taxation," *Review of Economic Studies*, 21, 21–30.
- Cunha, A., 2008, "The Optimality of the Friedman Rule When Some Distorting Taxes are Exogenous," *Economic Theory*, 35, 267–291.
- Dragone, D., 2009, "A rational eating model of binges, diets and obesity," *Journal of Health Economics*, 28, 799–804.
- Erosa, A., and M. Gervais, 2001, "Optimal Taxation in Infinitely Lived Agent and Overlapping Generations Models: A Review," *Economic Quarterly*, 87, 23–44, Federal Reserve Bank of Richmond.
- Erosa, A., and M. Gervais, 2002, "Optimal Taxation in Life Cycle Economies," *Journal of Economic Theory*, 105, 338–69.
- Goolsbee, A., M. Lovenheim, and J. Slemrod, 2009, "Playing With Fire: Cigarettes, Taxes, and Competition From the Internet," *American Economics Journal: Economic Policy*, forthcoming.
- Grinols, E., and D. Mustard, 2006, "Casinos, Crime, and Community Costs," *Review of Economics and Statistics*, 88, 28–45.
- Grossman, M., and F. Chaloupka, 1998, "The demand for cocaine by young adults: a rational addiction approach," *Journal of Health Economics*, 17, 427–474.
- Grossman, M., F. Chaloupka, and I. Sirtalan, 1998, "An Empirical Analysis of Alcohol Addiction: Results From Monitoring the Future Panels," *Economic Inquiry*, 36, 39–48.
- 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 of Cigarette Excise Taxes," *Journal of Public Economics*, 88, 1959–88.
- Hansen, A., 2004, "Lotteries and State Fiscal Policy," Discussion Paper 46, The Tax Foundation Background Paper.
- Kenkel, D., 1996, "New Estimates of the Optimal Tax on Alcohol," *Economic Inquiry*, 34, 296–319.
- Kocherlakota, N., 2005, "Zero Expected Wealth Taxes: A Mirrless Approach To Dynamic Optimal Taxation," *Econometrica*, 73, 1587–1621.
- Kyung, H., 2009, "Ramsey Fiscal Policy and Endogenous Growth," *Economic Theory*, 39, 377–98.
- Levy, A., 2002, "Rational Eating: Can it Lead to overweightness or underweightness?," *Journal of Health Economics*, 21, 887–899.
- Liu, X., and R. Lopez, 2009, "Rational Addiction Evidence from Carbonated Soft Drinks," University of Connecticut.
- Manning, W., E. Keeler, J. Newhouse, E. Sloss, and J. Wasserman, 1989, "The Taxes of Sin: Do Smokers and Drinkers Pay Their Way?," *Journal of the American Medical Association*, 261, 1604–9.
- Miron, J., 2010, "The Budgetary Implications of Drug Prohibition," Harvard University.
- O'Donoghue, T., and M. Rabin, 2003, "Studying Optimal Paternalism, Illustrated by a Model of Sin Taxes," *American Economic Review*, 83, 186–191.
- O'Donoghue, T., and M. Rabin, 2006, "Optimal Sin Taxes," *Journal of Public Economics*, forthcoming.
- Olekalns, N., and P. Bardsley, 1996, "Rational Addiction to Caffeine: An Analysis of Coffee Consumption," *Journal of Political Economy*, 104, 1100–04.
- Parry, I., R. Laxminarayan, and S. West, 2009, "Fiscal and Externality Rationales for Alcohol Taxes," *B.E. Journal of Economic Analysis and Policy*, 9, article 29.
- Pogue, T., and L. Sgontz, 1989, "Taxing to Control Social Costs: The Case of Alcohol," *American Economic Review*, 79, 235–43.
- Sgontz, L., 1993, "Optimal Taxation: The Mix of Alcohol and Other Taxes," *Public Finance Quarterly*, 21, 260–75.
- Sleet, C., and S. Yeltekin, 2006, "Optimal Taxation with Endogenously Incomplete Debt Markets," *Journal of Economic Theory*, 127, 36–73.

Sung, H., T. Hu, and T. Keeler, 1994, “Cigarette Taxation and Demand: An Empirical Model,” *Contemporary Economic Policy*, 12, 91–100.

Viscusi, W., 1995, “Cigarette Taxes and the Social Consequences of Smoking,” in J. Poterba (ed.), *Tax Policy and the Economy*, vol. 9, . pp. 51–101, MIT Press, Cambridge, MA.

Waters, T., and F. Sloan, 1995, “Why Do People Drink? Tests of the Rational Addiction Model,” *Applied Economics*, 27, 727–36.

7 Appendix: Proofs and Extras

7.1 Proof of Proposition 1

To see that a competitive equilibrium satisfies the IMC and resource constraint, we substitute the factor prices (3.2) and (3.3) into the budget constraint (3.11). Using constant returns to scale, we then have:

$$R_t^b b_t + F(k_t, h_t) - \tau_{h,t} F_h(k_t, h_t) h_t = (1 + \tau_{c,t}) c_t + (1 + \tau_{d,t}) d_t + i_t + b_{t+1}. \quad (7.1)$$

Combining the above equation with the government budget constraint (3.19) gives the resource constraint (4.1).

To derive the IMC from the budget constraint, we substitute the household first order conditions (3.12)-(3.14) into the budget constraint (3.11), eliminating the tax rates, so that:

$$\lambda_t R_t k_t + \lambda_t R_t^b b_t - \lambda_t k_{t+1} - \lambda_t b_{t+1} = \beta^t (u_c(c_t, s_t, l_t) c_t + MU_{d,t} d_t + u_l(c_t, s_t, l_t) h_t). \quad (7.2)$$

Next using the first order conditions (3.15) and (3.16), we have:

$$\lambda_t R_t (k_t + b_t) - \lambda_{t+1} R_{t+1} (k_{t+1} + b_{t+1}) = \beta^t (u_c(c_t, s_t, l_t) c_t + MU_{d,t} d_t + u_l(c_t, s_t, l_t) h_t). \quad (7.3)$$

The above equation characterizes a sequence of budget constraints that can be used to recursively eliminate $\lambda_t R_t (k_t + b_t)$, yielding:

$$\lambda_0 (R_0 k_0 + R_0^b b_0) - \lim_{t \rightarrow \infty} \lambda_{t+1} R_{t+1} (k_{t+1} + b_{t+1}) = \sum_{t=0}^{\infty} \beta^t (u_c(c_t, s_t, l_t) c_t + MU_{d,t} d_t - u_l(c_t, s_t, l_t) h_t). \quad (7.4)$$

The transversality conditions imply the second term on the left hand side equals zero. Again using the household first order conditions at period zero gives:

$$\frac{u_{c,0} (R_0 k_0 + R_0^b b_0)}{1 + \tau_{c,0}} = \sum_{t=0}^{\infty} \beta^t (u_c(c_t, s_t, l_t) c_t + MU_{d,t} d_t - u_l(c_t, s_t, l_t) h_t), \quad (7.5)$$

which is the IMC.

We next show that, given allocations which satisfy the IMC and resource constraint, prices and policies exist which, along with the allocations, are a competitive equilibrium. Let $\{c_t, k_t, h_t, d_t\}$ be a sequence which satisfies the IMC and resource constraint. Then r_t and w_t are defined using equations (3.2) and (3.3). Since $\tau_{c,0}$ is given, we can define λ_0 using equation (3.12). Then λ_t can be defined recursively using equation (3.15). Then R_t^b is defined using equation (3.16). Next, we define the government policies:

$$(1 + \tau_{c,t}) = \frac{\beta^t u_c(c_t, s_t, l_t)}{\lambda_t}, \quad (7.6)$$

$$(1 - \tau_{h,t}) = \frac{\beta^t u_l(c_t, s_t, l_t)}{\lambda_t F_h(k_t, h_t)}, \quad (7.7)$$

$$(1 + \tau_{d,t}) = \frac{\beta^t MU_{d,t}}{\lambda_t}, \quad (7.8)$$

Given the above prices and policies, all equations which define a competitive equilibrium are satisfied except the household and government budget constraints. We use b_t to satisfy the household budget constraint:

$$b_t = \frac{1}{R_t^b} (-r_t k_t - (1 - \tau_{h,t}) w_t h_t + (1 + \tau_{c,t}) c_t + (1 + \tau_{d,t}) d_t + i_t + b_{t+1}). \quad (7.9)$$

We can multiply the above equation by λ_t and recursively eliminate b_{t+1} from the above equation. After eliminating prices and policies using the household first order conditions (3.12)-(3.14), b_t is a function of the allocations:

$$b_t = \left(\prod_{i=0}^{t-1} (F_k(k_i, h_i) + 1 - \delta) \right) \frac{1 + \tau_{c,0}}{\tau_{c,0}} \sum_{i=t}^{\infty} \beta^i (u_{c,i} c_i + MU_{d,i} d_i - u_{l,i} h_i) - k_t \quad (7.10)$$

The above equation is the debt allocation which implies the household budget constraint is

satisfied.

Since the budget constraint is satisfied, we simply substitute the resource constraint into the budget constraint to see that the government budget constraint is satisfied. Finally, by substituting the prices and policies into the IMC and reversing the derivation of the IMC, we see that the transversality conditions are satisfied.

7.2 Proof of Proposition 2

First, we rewrite equations (4.3) and (4.5), using the σ definitions, so that:

$$IMC_{c_t} = \beta^t u_{c,t} (1 - \sigma_c + \alpha \sigma_{sc,t} - \sigma_{hc,t}), \quad (7.11)$$

$$IMC_{d_t} = \beta^t MU_{d,t} \left(\alpha - \alpha \sigma_{s,t} + \sigma_{cs,t} - \sigma_{hs,t} + \frac{\beta u_{s,t+1} s_{2,t+1}}{MU_{d,t}} (\alpha \sigma_{s,t} - \alpha \sigma_{s,t+1} - \sigma_{cs,t} + \sigma_{cs,t+1} + \sigma_{hs,t} - \sigma_{hs,t+1}) \right). \quad (7.12)$$

Now since v is homothetic, we know that:

$$\frac{v_c(\psi c, \psi s)}{v_s(\psi c, \psi s)} \equiv \frac{v_c(c, s)}{v_s(c, s)}, \quad (7.13)$$

which implies:

$$\frac{v_{cc}(c, s) c}{v_c(c, s)} + \frac{v_{cs}(c, s) s}{v_c(c, s)} = \frac{v_{ss}(c, s) s}{v_s(c, s)} + \frac{v_{cs}(c, s) c}{v_s(c, s)}, \quad (7.14)$$

which, using the definition of $u(\cdot)$ in equation (5.1), implies:

$$\sigma_{sc} - \sigma_c = \sigma_{cs} - \sigma_s. \quad (7.15)$$

It is also immediate from the definition of $u(\cdot)$ that $\sigma_{hc} = \sigma_{hs}$. These facts and equations (4.7), (4.8), (7.11), and (7.12) together imply:

$$\frac{MU_{d,t}}{u_{c,t}} = \frac{1 + \mu (1 - \sigma_{s,t} + \sigma_{cs,t} - (1 - \alpha) \sigma_{sc,t} - \sigma_{hs,t})}{1 + \mu (\alpha - \alpha \sigma_{s,t} + \sigma_{cs,t} - \sigma_{hs,t} + J(\alpha \Delta \sigma_s - \Delta \sigma_{cs} + \Delta \sigma_{hs}))}. \quad (7.16)$$

Hence, $\tau_{d,t} > \tau_{c,t}$ if and only if the right hand side is greater than one, or:

$$1 - (1 - \alpha) \sigma_{sc,t} - \sigma_{s,t} > \alpha - \alpha \sigma_{s,t} + J(\alpha \Delta \sigma_s - \Delta \sigma_{cs} + \Delta \sigma_{hs}), \quad (7.17)$$

which simplifies to the desired result.

7.3 Proof of Propositions 3-4

For the CRR case, note that $\sigma_{sl} = 0$, $\sigma_s = 1 - (1 - \xi)(1 - \sigma)$, and $\sigma_{sc} = (1 - \xi)(1 - \sigma)$, which implies the left hand side of condition (5.3) is zero. The right hand side of (5.3) is also zero since σ_{cs} and σ_s are constant.

For the steady state case, $\sigma_{i,t} = \sigma_{i,t+1}$ for all $i \in \{s, sc, cs, hs\}$, so the result follows immediately from condition (5.3).

7.4 Details of the linear quadratic case: Proof of existence and of Propositions 5-6 and of condition (5.8)

Proposition 5 was proved in the text. Now we prove that a unique positive solution for μ exists.

PROPOSITION 7 *Let the conditions for Proposition 5 hold. Let g_t be a stationary sequence with limiting value \bar{g} . Then there exists an interval $[\zeta_l, \zeta_h]$, with $0 < \zeta_l < \zeta_h < \infty$ such that if $G \equiv \sum_{t=0}^{\infty} \beta^t g_t \in [\zeta_l, \zeta_h]$, then a unique positive solution for μ exists.*

For Proposition 7, we derive the solution for μ as follows. First, for the quadratic case, the first order conditions for the Ramsey problem (4.7)-(4.10) are now:

$$\frac{\phi_t}{\beta^t} = \omega(1 + \mu), \quad (7.18)$$

$$\frac{\phi_t}{\beta^t} = \nu(1 - \beta\gamma)(1 + \mu) - (1 + 2\mu)(s_t - \beta\gamma s_{t+1}), \quad (7.19)$$

$$\frac{\phi_t(1 - \theta)\left(\frac{k_t}{h_t}\right)^\theta}{\beta^t} = (e - 1)(1 + \mu) + (1 + 2\mu)h_t, \quad (7.20)$$

$$\phi_t \left(\theta \left(\frac{k_t}{h_t} \right)^{\theta-1} + 1 - \delta \right) = \phi_{t-1}. \quad (7.21)$$

Using equation (7.18) to eliminate ϕ_t gives:

$$\omega(1 + \mu) = (\nu(1 + \mu) - \mu)(s_t - \beta\gamma s_{t+1}), \quad (7.22)$$

$$\omega(1+\mu)(1-\theta)\left(\frac{k_t}{h_t}\right)^\theta = (e-1)(1+\mu) + (1+2\mu)h_t, \quad (7.23)$$

$$\beta\left(\theta\left(\frac{k_t}{h_t}\right)^{\theta-1} + 1 - \delta\right) = 1. \quad (7.24)$$

For the subtractive model equation (7.22) implies:

$$\omega(1+\mu) = \nu(1-\beta\gamma)(1+\mu) - (1+2\mu)(d_t - \gamma d_{t-1} - \beta\gamma(d_{t+1} - \gamma d_t)), \quad (7.25)$$

which simplifies to (5.4), and which we show in the next proof has general solution given by (5.5). Proposition 6 thus holds if a non-zero and finite solution for μ exists, which we now show.

Equation (7.24) implies the capital to labor ratio, denoted by A , is constant:

$$A \equiv \left(\frac{\theta}{\rho + \delta}\right)^{\frac{1}{1-\theta}}. \quad (7.26)$$

Thus, equation (7.23) implies

$$h_t = \frac{1+\mu}{1+2\mu}\hat{h}, \quad \hat{h} \equiv 1 - e + \omega(1-\theta)A^\theta, \quad (7.27)$$

is constant. Thus, $k_t = Ah_t$ is constant and equation (5.5) implies s_t and thus $MU_{d,t}$ is constant as well. Since elasticity of substitution of consumption over time is infinite, the planner absorbs all changes in g_t by varying c_t . Combining these results with resource constraint (4.1) yields a solution for c_t :

$$c_t = \frac{1+\mu}{1+2\mu}\left(\hat{h}(A^\theta - \delta A) - \hat{d}(1 - \gamma^{t+1})\right) - \gamma^{t+1}d_{-1} - g_t, \quad (7.28)$$

$$\hat{d} \equiv \frac{\nu(1-\beta\gamma) - \omega}{(1-\gamma)(1-\beta\gamma)}. \quad (7.29)$$

Now since h_0 enters into the left hand side of the IMC (4.2) and k_0 is given, the solutions for h_0 , k_0 and therefore c_0 generally differ from the solutions for $t \geq 1$. Therefore, we let $x \equiv \frac{1+\mu}{1+2\mu}$ and insert the solutions for c_t , d_t , and h_t into the IMC (4.2) for $t \geq 1$, so that:

$$\frac{R_0 k_0 + R_0^b b_0}{1 + \tau_{c,0}} = \sum_{t=1}^{\infty} \beta^t \left[x\omega \left(\hat{h} (A^\theta - \delta A) - \hat{d} (1 - \gamma^{t+1}) \right) - \omega \gamma^{t+1} d_{-1} + \right.$$

$$(1 - \beta\gamma) \left(\nu - (1 - \gamma) \hat{d}x \right) \left(\hat{d}x (1 - \gamma^{t+1}) + \gamma^{t+1} d_{-1} \right) - (e - 1 + \hat{h}x) \hat{h}x \Big] + \omega c_0 + MU_{d,0} d_0 - u_{l,0} h_0 + \omega g_0 - \omega G. \quad (7.30)$$

Next, recall from Proposition 1 that $\tau_{c,0}$ and $\tau_{h,0}$ are given. It follows from equations (3.12) and (3.13) that the planner cannot choose h_0 in this example, and instead takes the solution for h_0 from the competitive model as given. Further, the terms inside the summation depend on time only through γ^{t+1} and β^t , and the equation is quadratic in x . Therefore, after evaluating the summation we can write equation (7.30) as:

$$ch(x) \equiv -\zeta_1 x^2 + (\zeta_1 - \zeta_2)x + \zeta_2 - \zeta_3 = 0, \quad (7.31)$$

$$\zeta_1 \equiv \hat{h}^2 + \frac{(1 - \gamma)^2}{\beta} \hat{d}^2, \quad (7.32)$$

$$\zeta_2 \equiv \left(\frac{1 - \beta}{\beta} \right) (1 - \gamma) \gamma \hat{d} d_{-1}, \quad (7.33)$$

$$\zeta_3 \equiv \frac{1 - \beta}{\beta} \left(\omega G - \frac{(\omega(1 + \tau_{c,0}) - 1)(1 - \delta)k_0 - R_0^b b_0}{1 + \tau_{c,0}} - \right. \quad (7.34)$$

$$\left. \frac{\omega(1 + \tau_{c,0}) - \theta}{1 + \tau_{c,0}} k_0^\theta h_0^{1-\theta} + u_{l,0} h_0 \right). \quad (7.35)$$

A solution such that $\mu > 0$ is a solution in the range $\frac{1}{2} < x < 1$. Note that equation (7.31) can be written as:

$$ch(x) = (\zeta_1 x + \zeta_2)(1 - x) - \zeta_3 = 0. \quad (7.36)$$

Now $ch(x)$ attains a maximum at $x^* = (\zeta_1 - \zeta_2) / (2\zeta_1) < 1/2$ and $ch(0) > ch(1)$. Hence it is immediate that $\zeta_3 > 0$ is necessary for $x < 1$.

From equation (7.35), $\zeta_3 > 0$ if and only if:

$$G > \zeta_l \equiv \frac{(\omega(1 + \tau_{c,0}) - 1)(1 - \delta)k_0 - R_0^b b_0}{\omega(1 + \tau_{c,0})} + \frac{\omega(1 + \tau_{c,0}) - \theta}{\omega(1 + \tau_{c,0})} k_0^\theta h_0^{1-\theta} - \frac{u_{l,0} h_0}{\omega}. \quad (7.37)$$

This is the lower bound for G .

Condition (7.37) implies that $\zeta_3 > 0$ which implies both roots have modulus less than one. It remains to show that the roots are real and that one root is greater than one half.

Since $x^* < 1/2$, the smaller root has modulus less than one half. The larger root is real and greater than one half if and only if:

$$ch\left(\frac{1}{2}\right) = -\frac{\zeta_1}{4} + \frac{\zeta_1 - \zeta_2}{2} + \zeta_2 - \zeta_3 > 0, \quad (7.38)$$

$$\zeta_3 < \frac{\zeta_1 - 2\zeta_2}{4}. \quad (7.39)$$

Using equation (7.35), condition (7.39) holds if and only if:

$$G < \zeta_h \equiv \zeta_l + \frac{\beta}{1 - \beta} \left(\frac{\zeta_1 + 2\zeta_2}{4\omega} \right). \quad (7.40)$$

Defining ζ_l and ζ_h using equations (7.37) and (7.40) completes the proof. \square

For Proposition 6, we solve the second order difference equation (5.4):

$$d_{t+1} - \left(\frac{1 + \beta\gamma^2}{\beta\gamma} \right) d_t + \frac{1}{\beta} d_{t-1} = -\frac{1 + \mu}{1 + 2\mu} \left(\frac{\nu(1 - \beta\gamma) - \omega}{\beta\gamma} \right). \quad (7.41)$$

It is straightforward to show the general solution of the above difference equation is:

$$d_t = D_p + A_0\gamma^t + A_1(\beta\gamma)^{-t}, \quad (7.42)$$

$$D_p \equiv \frac{1 + \mu}{1 + 2\mu} \left(\frac{\nu(1 - \beta\gamma) - \omega}{(1 - \beta\gamma)(1 - \gamma)} \right). \quad (7.43)$$

Following convention, we rule out the explosive, bubble solutions which requires $A_1 = 0$. Letting $t = -1$ implies $A_0 = \gamma(d_{-1} + D_p)$. Substituting for A_0 and simplifying gives the desired solution. \square

Finally, to derive equation (5.8), we rewrite equation (5.3) using the definition of x , which implies:

$$\frac{1 + \tau_{d,t}}{1 + \tau_{c,t}} = \frac{MU_{d,t}}{\omega} = x + \frac{\nu(1 - \beta\gamma)}{\omega} (1 - x). \quad (7.44)$$

Hence:

$$\frac{\partial}{\partial \gamma} \left(\frac{1 + \tau_{d,t}}{1 + \tau_{c,t}} \right) = \frac{\partial x}{\partial \gamma} \cdot \left(1 - \frac{\nu(1 - \beta\gamma)}{\omega} - \frac{\nu\beta}{\omega} (1 - x) \right) < 0 \text{ iff}, \quad (7.45)$$

$$\frac{\partial x}{\partial \gamma} > \frac{\beta \nu (1 - x)}{\nu (1 - \beta \gamma) - \omega}. \quad (7.46)$$

Next, using the implicit function theorem on equation (7.36), we see that:

$$\frac{\partial x}{\partial \gamma} = \frac{(\zeta_{1\gamma} x + \zeta_{2\gamma})(1 - x)}{2\zeta_1 x - \zeta_1 + \zeta_2}. \quad (7.47)$$

We have $x \in (\frac{1}{2}, 1)$, and so the denominator of equation (7.47) is positive. Thus, substituting equation (7.47) into condition (7.46) and cross multiplying results in:

$$\left[\zeta_{1\gamma} (\nu (1 - \beta \gamma) - \omega) + 2\beta \nu \zeta_1 \right] x + \zeta_{2,\gamma} (\nu (1 - \beta \gamma) - \omega) + \beta \nu \zeta_2 > \beta \nu \zeta_1. \quad (7.48)$$

The coefficient on x equation (7.48) is positive. Thus, it is sufficient to show:

$$\left[\zeta_{1\gamma} (\nu (1 - \beta \gamma) - \omega) + 2\beta \nu \zeta_1 \right] \frac{1}{2} + \zeta_{2,\gamma} (\nu (1 - \beta \gamma) - \omega) + \beta \nu \zeta_2 > \beta \nu \zeta_1, \quad (7.49)$$

$$\frac{1}{2} \zeta_{1\gamma} (\nu (1 - \beta \gamma) - \omega) + \zeta_{2,\gamma} (\nu (1 - \beta \gamma) - \omega) + \beta \nu \zeta_2 > 0. \quad (7.50)$$

Finally substituting in the definitions of ζ using equations (7.32) and (7.33) and the derivatives of ζ :

$$\zeta_{1,\gamma} = -\frac{2\omega (\nu (1 - \beta \gamma) - \omega)}{(1 - \beta \gamma)^3}, \quad \zeta_{2,\gamma} = \frac{1 - \beta}{\beta} \left(\frac{\nu (1 - \beta \gamma)^2 - \omega}{(1 - \beta \gamma)^2} \right) d_{-1}, \quad (7.51)$$

and simplifying yields the desired result.

7.5 Additively Separable Utility

In this section we consider the case in which utility is additively separable. Following a similar procedure as with Proposition 2, we have:

PROPOSITION 8 *Let assumptions (A1)-(A3) hold. In addition, let $u(\cdot)$ be additively separable in c , s , and l . Then $\tau_{d,t} > \tau_{c,t}$ if and only if:*

$$\alpha \sigma_{s,t} + 1 - \alpha - \sigma_{c,t} > J(\alpha \Delta \sigma_s). \quad (7.52)$$

Proof: If utility is separable, equations (7.11) and (7.12) become:

$$IMC_{ct} = \beta^t u_{c,t} (1 - \sigma_c), \quad (7.53)$$

$$IMC_{d_t} = \beta^t MU_{d,t} (\alpha - \alpha\sigma_{s,t} + J(\alpha\Delta\sigma_s)). \quad (7.54)$$

Equations (4.7), (4.8), (7.53), and (7.54) together imply:

$$\frac{MU_{d,t}}{u_{c,t}} = \frac{1 + \mu(1 - \sigma_{c,t})}{1 + \mu(\alpha(1 - \sigma_{s,t}) + J(\alpha\Delta\sigma_s))}. \quad (7.55)$$

Hence, d is taxed at a higher rate if and only if the right hand side is greater than one, or:

$$1 - \sigma_{c,t} > \alpha(1 - \sigma_{s,t}) + J(\alpha\Delta\sigma_s), \quad (7.56)$$

which simplifies to the desired result. \square

For the intuition of (7.52), first assume the terms multiplied by d_{t-1} in equation (4.5) and (4.3) are zero. Further, assume all terms multiplied by β in equation (4.5) equal zero. Then all dynamic effects are removed and condition (7.52) becomes:

$$\sigma_{s,t} \frac{s_{1,t}d_t}{s_t} - \frac{s_{11,t}d_t}{s_{1,t}} - \sigma_{c,t} > 0 \quad (7.57)$$

This condition is satisfied if and only if d_t is more complementary with leisure (or income inelastic) than c_t in the model without dynamic effects. For example, assume again that $s_{11,t} = 0$ and $s_{1,t}d_t = s_t$. Then condition (7.57) holds if and only if s is more complementary with leisure than c . If these conditions do not hold, then d_t and s_t have different income elasticities in the model without dynamic effects, and thus condition (7.57) results.

The difference between the left hand side of conditions (7.57) and (7.52) is the tax anticipation effect. The tax anticipation effect reduces the optimal addictive tax if and only if:

$$\sigma_{s,t} \frac{s_{1,t}d_t}{s_t} - \frac{s_{11,t}d_t}{s_{1,t}} - \sigma_{c,t} > \alpha\sigma_{s,t} + 1 - \alpha - \sigma_{c,t}. \quad (7.58)$$

Exploiting homogeneity and $s_2 < 0$, condition (7.58) reduces exactly to condition (3.6). Thus, the tax anticipation effect reduces the tax rate on d_t if and only d_t is addictive.

As in Proposition (2), the dynamic effects in $t - 1$ and t reduce to a single current period term which depends only on the elasticities and homogeneity, and should therefore be straightforward to check in empirical applications.

For the addiction stock effect, we remove the left hand side of condition (7.52) which consists of the tax anticipation effect and the usual static period t effect. The addiction stock

effect is thus $J(\alpha\Delta\sigma_s) < 0$, which implies d_t should be taxed at a higher rate if and only if $\sigma_{s,t+1} < \sigma_{s,t}$. Suppose s_t and leisure are substitutes, and are becoming more strongly so over time ($\sigma_{s,t+1} < \sigma_{s,t}$), relative to c_t and l_t . Then an increase in $\tau_{d,t}$ results in a decrease in d_t and s_t and thus an increase in leisure and a loss of labor tax revenues in t . However, the decrease in d_t lowers the addiction stock in $t+1$ and thus raises effective consumption in $t+1$. Since s_{t+1} and leisure are substitutes, leisure falls and labor tax revenues rise in $t+1$. Further, labor tax revenues rise more in $t+1$, because effective consumption and leisure are more strongly substitutable. Thus, taxing d_t is relatively attractive, which shows up as $J(\alpha\Delta\sigma_s) < 0$.

As shown in the following proposition, however, the common CRR specification has constant elasticities for which the addiction stock effect vanishes.

PROPOSITION 9 *Let the conditions of Proposition 8 hold, and let $u(\cdot) = v^1(c) + v^2(s) + v^3(l)$, with v^1 and $v^2(\cdot)$ CRR. Then $\tau_{d,t} > \tau_{c,t}$ for all t if and only if:*

$$\alpha\sigma_s + 1 - \alpha > \sigma_c. \tag{7.59}$$

Proof: For CRR preferences, $\sigma_{i,t} = \sigma_{i,t+1}$ for all $i \in \{s, c\}$, so the result follows immediately from condition (7.52). \square

Condition (7.59) combines the steady state tax anticipation effect with any difference in income elasticities between d_t and c_t . Condition (7.59) requires that d_t and leisure to be sufficiently strong complements relative to c in the model without dynamic effects to overcome the dynamic tax anticipation effect.

The homogeneity of the addiction function affects the strength of the tax anticipation effect. For $\alpha = 1$, as in the subtractive model, the dynamic tax anticipation effect exactly offsets the difference in income elasticities between d and s in the model without dynamic effects. In this case, proposition (9) says $\tau_{d,t} > \tau_{c,t}$ if and only if s_t is more complementary with leisure in a static model where s could be taxed. However, in the multiplicative model, higher tolerance implies a lower α and thus makes addictive goods less complementary with leisure, lowering the optimal tax rate, for $\sigma_{s,t} > 1$.

Thus, under the conditions outlined above, neglecting the dynamic tax anticipation effect results in overly high tax rates for addictive goods, relative to the optimum, especially for addictive goods that exhibit strong tolerance.

7.6 Non-weakly separable utility

For the homothetic and separable cases studied above, weak separability implies that the static income elasticity and the substitutability with leisure are identical. Here we consider a utility function for which the substitutability with leisure does not necessarily equal the income elasticity.

To see this in a concise way, let us consider the following class of utility functions:

$$u(s_t, d_t, l_t) = q(c_t) + v(s_t, l_t). \quad (7.60)$$

For this specification, we find:

PROPOSITION 10 *Let assumptions (A1)-(A3) hold. In addition, let $u(\cdot)$ be of the form given by (7.60). Then $\tau_{d,t} > \tau_{c,t}$ if and only if:*

$$1 - \alpha + \alpha\sigma_{s,t} + \sigma_{hs,t} - \sigma_{c,t} > J(\alpha\Delta\sigma_s + \Delta\sigma_{hs}). \quad (7.61)$$

Proof: If utility is given by equation (7.60), equations (7.11) and (7.12) become:

$$IMC_{c_t} = \beta^t u_{c,t} (1 - \sigma_{c,t}), \quad (7.62)$$

$$IMC_{d_t} = \beta^t MU_{d,t} (\alpha - \alpha\sigma_{s,t} - \sigma_{hs,t} + J(\alpha\Delta\sigma_s + \Delta\sigma_{hs})). \quad (7.63)$$

Equations (4.7), (4.8), (7.62), and (7.63) together imply:

$$\frac{MU_{d,t}}{u_{c,t}} = \frac{1 + \mu(1 - \sigma_{c,t})}{1 + \mu(\alpha(1 - \sigma_{s,t}) - \sigma_{hs,t} + J(\alpha\Delta\sigma_s + \Delta\sigma_{hs}))}. \quad (7.64)$$

Hence, d is taxed at a higher rate if and only if the right hand side is greater than one, or:

$$1 - \sigma_{c,t} > \alpha(1 - \sigma_{s,t}) - \sigma_{hs,t} + J(\alpha\Delta\sigma_s + \Delta\sigma_{hs}), \quad (7.65)$$

which simplifies to the desired result. \square

For the intuition of (7.61), suppose again that neither the tax anticipation effect nor the addiction stock effect are present, then condition (7.61) reduces to:

$$\sigma_{s,t} \frac{s_{1,t} d_t}{s_t} - \frac{s_{11,t} d_t}{s_{1,t}} + \sigma_{hs,t} - \sigma_{c,t} > 0. \quad (7.66)$$

This condition is now a mixture of the difference in income elasticities and substitutability of leisure. But the tax anticipation effect works in an exactly similar manner as in the

separable case, since the difference between the left hand side of condition (7.61) and (7.66) is exactly (3.6). Thus the tax anticipation effect reduces the optimal tax on d_t if and only if d_t is addictive.

The addiction stock effect also works in a similar manner. A change in the current addictive consumption induced by the addictive tax changes the period $t + 1$ stock of addiction, which affects addictive and labor tax revenues in period $t + 1$.

The addiction stock effect vanishes in the steady state or with constant elasticities.

PROPOSITION 11 *Let the conditions of Proposition 10 hold, and let $v(\cdot) = \frac{(s^\xi(1-l)^{-(1-\xi)})^{1-\sigma} - 1}{1-\sigma}$, with $1 - \sigma < \frac{-1}{1-2\xi}$ and $\xi < \frac{1}{2}$ to ensure concavity. Then $\tau_{d,t} < \tau_{c,t}$ if and only if:*

$$1 + (1 - \sigma)(1 - \xi(1 + \alpha)) > \sigma_c. \quad (7.67)$$

Proof: note that $\sigma_s = 1 - \xi(1 - \sigma)$, and $\sigma_{sl} = (1 - \xi)(1 - \sigma)$. Since both terms are constant, we have $J = 0$. Substituting in these conditions into equation (7.61) gives (7.67). \square

Note that the concavity restrictions imply the left hand side of (7.67) is less than one. Since $\sigma > 1$ and $\xi < \frac{1}{2}$ (both conditions are necessary for concavity) and $\alpha \leq 1$, the second term on the left hand side is negative and so $\sigma_{c,t} \geq 1$ is sufficient for the condition to be violated, and thus $\tau_{d,t} \leq \tau_{c,t}$.

PROPOSITION 12 *Let the conditions of Proposition 10 hold. Then $\bar{\tau}_d > \bar{\tau}_c$ if and only if:*

$$\alpha(\bar{\sigma}_s - 1) + \bar{\sigma}_{hs} > \bar{\sigma}_c - 1. \quad (7.68)$$

Proof: $\sigma_{i,t} = \sigma_{i,t+1}$ for all $i \in \{s, hs\}$ in the steady state, so the result follows immediately from condition (7.61). \square

It is possible to construct examples for which d_t is more complementary with leisure and yet the tax anticipation effect implies a lower tax rate for addictive goods. Suppose subtractive model, then in the static model with utility as in proposition 11, s is more complementary with leisure than c . Yet if (7.67) is violated it is optimal to tax d_t at a lower rate. Further, in the steady state $\bar{s} = \bar{d}(1 - \gamma)$, and so condition (7.66) becomes:

$$\frac{1 - \xi(1 - \sigma)}{1 - \gamma} + (1 - \xi)(1 - \sigma) > \bar{\sigma}_c, \quad (7.69)$$

and condition (7.68) becomes:

$$1 - \xi(1 - \sigma) + (1 - \xi)(1 - \sigma) > \bar{\sigma}_c. \quad (7.70)$$

So for $\bar{\sigma}_c$ satisfying (7.69), but not (7.70), $\bar{\tau}_d > \bar{\tau}_c$ in the model without dynamic effects but $\bar{\tau}_d < \bar{\tau}_c$ when the tax anticipation effect is accounted for. Note the range of values satisfying (7.69) but not (7.70) is increasing in γ . Strong tolerance tends to decrease the optimal addictive tax, by strengthening the tax anticipation effect.

8 Appendix: Tables and Figures

Parameter	Value	Results for Constant Variables				
		variable	$\gamma = 0$	$\gamma = 0.45$	$\gamma = 0.5$	$\gamma = 0.55$
β	0.9	k_t	2.19	1.66	1.58	1.54
ν	2	h_t	0.76	0.57	0.55	0.53
θ	0.4	y_t	1.16	0.88	0.83	0.81
g_t	0.3	i_t	0.22	0.17	0.16	0.15
δ	0.1	μ	0.39	2.23	3.48	4.74
e	0.95	x	0.78	0.59	0.56	0.55
d_{-1}	0.02	τ_c	0.07	0.07	0.07	0.07
$\tau_{c,0}$	0.07	τ_d	0.30	0.15	0.12	0.075
b_0	0	τ_h	0.18	0.39	0.42	0.44
ω	1					

Table 1: Parameter values and results for variables which are constant over time. The parameters h_0 and k_0 are set equal to h_t using equation (7.27) and $k_t = Ah_t$, respectively. The parameter g_t is set equal to 30% of GDP for all t .

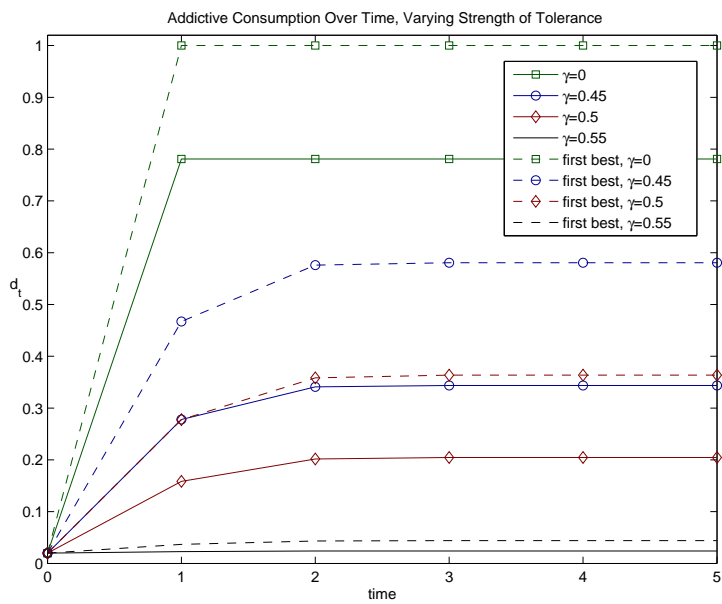


Figure 1: Dynamics of first and second best addictive consumption for various values of γ .