Internalities and Paternalism: Applying the Compensation Criterion to Multiple Selves across Time

27 October 2008

Eric Rasmusen

Abstract

One reason to call an activity a vice and suppress it is that it reduces a person's future happiness more than it increases his present happiness. Gruber & Koszegi (2001) show how a vice tax can increase a person's welfare in a model of multiple selves with hyperbolic preferences across time. The present paper shows that an interself analogy of the compensation criterion can justify a vice ban whether preferences are hyperbolic or exponential, but subject to the caveat that the person has a binding constraint on borrowing.

Dan R. and Catherine M. Dalton Professor, Department of Business Economics and Public Policy, Kelley School of Business, Indiana University. BU 438, 1309 E. 10th Street, Bloomington, Indiana, 47405-1701. Office: (812) 855-9219. Fax:812-855-3354. erasmuse@indiana.edu, http://www.rasmusen.org. This paper: http://www.rasmusen.org/papers/internality-rasmusen.pdf.

Keywords: Internalities, sin tax, moral regulation, Kaldor-Hicks criterion, time inconsistency, hyperbolic preferences.

I thank Kevin Du and participants in seminars at Bath, Bern, and East Anglia and the Midwest Law and Economics 2008 meeting for their comments, and Nuffield College, Oxford, for its hospitality.

1. Introduction

A common justification for laws to discourage vice is that although the vice makes a person happy in the present, it will cause him even more unhappiness in the future. Gruber & Koszegi (2001) formalize this argument in the context of smoking and show how cigarette taxes can raise welfare. They build a model of multiple selves linked across time by altruism and the addictive consequences of smoking. They call these consequences "internalities" between different time-selves of the same person in analogy to "externalities" between different people. Their model assumes hyperbolic discounting and as a welfare criterion they maximize the utility of the first of the multiple selves.

What is most important in Gruber & Koszegi (2001), however, is not the hyperbolic discounting but the idea of multiple selves and the inefficiency that can arise from internalities. As will be shown below, if we are prepared to accept a multiple selves model and the idea of basing a welfare function on it, then paternalism can be desirable even if the person's discounting is orthodoxly exponential.

2. The Model

A person lives for three periods labelled 0, 1, and 2. He receives an endowment of W in each period.We will denote his consumption of the single good by C_0, C_1, and C_2. Our person can save at interest rate r, but he cannot borrow. We will denote wealth in each period as W_0, W_1, W_2, where $W_0 = W$ and the later wealths depend on earlier saving.

In period 1, the person chooses between smoking ($X = 1$) or not smoking ($X = 0$). We will assume smoking costs no money, for simplicity. If the person chooses to smoke he receives 1 unit of utility in period 1 and loses α in period 2, where $\alpha > 1$ so the loss is bigger than the gain.

Total utility is:

$$U_0 = U(C_0) + \beta\delta(X + U(C_1)) + \beta\delta^2(-\alpha X + U(C_2)), \tag{1}$$

with $0 \leq \beta < 1$ and

$$0 \leq \delta < \frac{1}{1+r}. \tag{2}$$

The "consume-early" assumption (2) ensures that the person's rate of time preference is greater than the interest rate, an assumption useful for the first part of this paper, where we will also assume that utility is linear in consumption ($U(C) = C$).

The utility function includes both the case of the conventional exponential discounting and quasi-hyperbolic discounting in the style of Laibson (1997): our person has an exponential discount factor of 0δ and a hyperbolic adjustment parameter of β. If $\beta = 1$ he has a standard exponential-discounting utility function. A value of β between 0 and 1 leads to the standard features of hyperbolic discounting: besides preferring earlier consumption to later, the person puts special value on immediate consumption: at period 0 his tradeoff between consumption in periods 1 and 2, both distant, is different from the tradeoff he would make in period 1, when period-1 consumption is immediate.

This person onsists of three selves, Self 0, Self 1, and Self 2. Self 0's utility function is (1). Self 1 and Self 2 have the utility functions

$$U_1 = X + U(C_1) + \beta\delta(-\alpha X + U(C_2)), \tag{3}$$

and

$$U_2 = -\alpha X + U(C_2). \tag{4}$$

Self 0 controls the value of C_0, Self 1 controls X and C_1, and Self 2 controls C_2. This distribution of power, not the equal endowments across time, is the true distribution of property rights.

Self 2 would consume his entire wealth, so $C_2 = W + (1 + r)(W_1 - C_1)$.

Self 1 would maximize (3) by choice of X and C_1 subject to the constraint that $C_2 = W + (1+r)(W_1 - C_1)$. He derives utility from Self 2's consumption, but it is discounted by $\beta\delta$, which is less than $1/(1 + r)$ by the consume-early assumption. He would choose $X = 1$ if

$$\alpha < \frac{1}{\delta\beta}. \tag{5}$$

Self 0 will choose $C_0 = W$ because of the consume-early assumptino. If

Self 0 could control smoking, he would choose for Self 1 to smoke if

$$\alpha < \frac{1}{\delta} \tag{6}$$

3. Welfare Analysis: The Precommitment Criterion

Is the laissez faire outcome optimal? An obvious welfare criterion is intraself Pareto optimality: no change can be made which would make some self better off without making some other self worse off. This has frequently been applied to multiple selves models (e.g. in Bernheim & Rangel (2007)). The three selves, however, have contradictory preferences for smoking and consumption. When Selves 0 and 2 do not want smoking but Self 1 does, what should the social planner do?

A second welfare criterion is what I will call the "precommitment criterion": maximize Self 0's utility by seeing what would happen if he could directly commit to a future course of action, rather than having to try to manipulate his future selves. This is is the same criterion as we follow in a single-self model, simply recognizing that the social planner will have to somehow constrain the future selves to act in accord with Self 0's wishes. The precommitment criterion is commmonly used in the hyperbolic discounting literature, almost always in connection with decisions about consumption and savings (e.g. Laibson (1997), Krusell & Smith (2003)). It has two variants. In one, it is literally the earliest self whose utility is maximized. In the other, the modeller creates a "long-run self" who exists before Self 0 and who derives no utility from present sensations. The long-run self is effectively a self with exponential utility. He might have a β term in his utility function, but since he takes no immediate actions his present-orientedness is irrelevant.

Gruber & Koszegi (2001) use a model of a divided self to look at whether smoking should be taxed. A quasi-hyperbolic discounter lives for a finite number of periods and must decide the amount of smoking and other consumption for each period. They show that a large tax on cigarettes is optimal if the goal is to maximize Self 0's utility. O'Donoghue & Rabin (2003) also look at sin taxes with time- inconsistent consumers, to make the point that time-consistent consumers may be hurt very little and time-inconsistent consumers helped a great deal by sin taxes. Bhattacharya & Ladkdawalla (2004)

look instead at interself Pareto improvements that can result from government smoking policy. They suggest the use of voluntary smoking licenses that a person could purchase to constrain his future smoking. These licenses would subject the future self to high cigarette taxes, but would also entitle the future self to a lump-sum transfer that would make him better off overall.

What happens in our three-self model if we use the precommitment criterion?

Using the precommitment criterion, the laissez faire outcome fails to maximize welfare if $\alpha \in (1/\delta, 1/\delta\beta)$ because Self 0's optimality condition (6) and Self 1's optimality condition (5) conflict if there is hyperbolic discounting ($\beta < 1$). Self 1 trades off the present period 1 against the future period 2 at the high rate of 1 to $\beta\delta$, whereas Self 0 trades off the future period 1 against the distant-future period 2 at the lower rate of δ. The problem is not that sensations further in the future should not be treated as less important– a low δ by itself would not introduce a reason for paternalism– but that quasi-hyperbolic discounting makes the present period count too heavily.

If discounting is exponential, i.e. $\beta = 1$, there is no time inconsistency and the precommitment criterion will not call for a paternalistic ban on smoking. Self 1 will make the correct smoking decision even without coercion. It is still true that smoking hurts Self 2, possibly by much more than it helps Self 1, but if discounting via the exponential discount rate δ is heavy enough that Self 1 would choose smoking, then it is heavy enough that Self 0 cares much more about period 1 smoking utility than period 2 smoking disutility.

The main objection to the precommitment criterion is its arbitrariness. Caplin & Leahy (2004) and Whitman (2006) criticize focussing on the welfare of the "long-run " or earliest self– the "dictatorship of the present". Why privilege Self 0 rather than some later self? Of course, the criterion is not meant to be taken literally: what is special about Self 0 is that he is making a choice about something before it becomes a present decision. The idea is that the present-time orientation of a hyperbolically discounting person is illegitimate and we are willing to override consumer sovereignty. That is not an unreasonable argument, but we do not ordinarily argue that tastes are illegitimate. Moreover, the criterion disregards not only the desire of Self 1, the self making the choice about smoking, but also the desires of Self 2, the post-action self.

4. The Kaldor-Hicks Criterion

An alternative to the precommitment criterion is to apply the Kaldor-Hicks criterion, but within the self, requiring that a change would benefit all selves if appropriate cash transfers were made, but not requiring the transfers themselves. If we accept this logic for policy consequences between different individuals, it seems all the more acceptable within one individual, since we are no longer trading off utilities interpersonally.

The intraself Kaldor-Hicks criterion takes the laissez faire outcome as the base and asks whether the winners from a change would be willing to compensate the losers. This captures the idea with which we started that a vice causes more future disutility than it causes present utility.

Like the precommitment criterion, it does not suggest changing consumption alone from laissez faire. Because the discount rate exceeds the interest rate, a later self could not compensate an earlier self for saving more.

The smoking choice is different. It adds $(X - \beta\delta\alpha X)$ to Self 1's utility and $-\alpha X$ to Self 2's. Would Self 2 pay enough to change Self 1's laissez faire decision of $X = 1$? Selves 1 and 2 never meet, so the transaction is impossible, but Kaldor-Hicks only requires a potential Pareto improvement, not an actual one. There does exist a complication: Self 1 cares about Self 2's consumption level, so a dollar payment from Self 2 to Self 1 raises Self 1's utility by less than one dollar. Moreover, Self 0's utility is also affected by any transaction between Self 1 and Self 2. Thus, when we start to consider transfers, the question of whether a payment would compensate is more complicated than it would be between strangers.

Since Selves 0 and 2 both might benefit from the smoking ban, let us imagine both of them making payments to Self 1 in exchange for the ban, payments to Self 1 of P_0 and P_2 as measured in giving- period dollars. Since our starting point is $U(C_t) = W$ in each period, the differences between the utilities when the smoking bargain is made $(X = 0)$ and when it is not $(X = 1)$ are:

$$\Delta U_0 = \left\{[W - P_0] + \beta\delta\left(W + (1+r)P_0 + \frac{P_2}{1+r}\right) + \beta\delta^2(W - P_2)\right\} - \{W + \beta\delta(W+1) + \beta\delta^2(W-\alpha) \tag{7}$$

$$\Delta U_1 = \{W + (1+r)P_0 + \frac{P_2}{1+r} + \beta\delta(W - P_2)\} - \{W + 1 + \beta\delta(W - \alpha)\} \quad (8)$$

and

$$\Delta U_2 = \{W - P_2\} - \{W - \alpha\}. \quad (9)$$

Self 0's net benefit from the bargain to ban smoking is

$$Net\ Benefit(Self\ 0) = -P_0 + \beta\delta\left((1+r)\,P_0 + \frac{P_2}{1+r} - 1\right) + \beta\delta^2(-P_2 + \alpha) \quad (10)$$

Self 1's net benefit is

$$Net\ Benefit(Self\ 1) = (1+r)P_0 + \frac{P_2}{1+r} - 1 + \beta\delta(-P_2 + \alpha) \quad (11)$$

Self 2's is

$$Net\ Benefit(Self\ 2) = -P_2 + \alpha \quad (12)$$

We can find a condition for Pareto improvement by picking P_0 and P_2 to make Selves 1 and 2 indifferent about the bargain and then seeing if Self 0's utility rises under the resulting P_0. Working through the algebra yields the necessary condition:[1]

$$\alpha > 1 + r. \quad (13)$$

If the difference between Self 2's loss from smoking and Self 1's gain exceeds the interest rate, smoking should be banned.

Nothing in this argument has relied on hyperbolic discounting. We could have assumed $\beta = 1$ and the conclusion would be the same. Thus we have Result 1.

Result 1: The intraself Kaldor-Hicks criterion can justify paternalistic banning of a vice even if discounting is exponential.

[1] If it is Self 2 alone who pays Self 1, this requires that Self 2's income be large enough: $W \geq \frac{(1+r)(1-\alpha\beta\delta)}{1-\beta\delta(1+r)}$. Otherwise, Self 0 must be able to pay the remainder.

Unlike the precommitment criterion, the intraself Kaldor-Hicks criterion does not rely on conflicting time preferences to justify helping a future self, but on low interest rates. Result 1 relies not on the altruism of earlier selves to generate utility, but on the willingness of the future self to sacrifice consumption. The altruism of earlier selves still helps make a smoking ban more attractive, but it is no longer a necessary condition. This can be seen by setting $\beta = 0$, so the earlier selves care nothing for the later self. The precommitment criterion no longer has anything to say about a smoking ban, since Self 0 cares nothing about either Self 1 or Self 2. The intraself Kaldor-Hicks criterion, however, still bans smoking if $\alpha > 1 + r$, which is to say if Self 2 loses more from Self 1's smoking than can be compensated by the rate of interest.

We can think of there being market failure within the self, a missing market for trade of future money for present abstention. If the social planner created that market, he would create gains from trade, and both buyer and seller would benefit when Self 1 sells a smoking ban to Self 2.

Contrast the conditions under which a smoking ban raises welfare under our two welfare criteria.

$$Precommitment : \alpha \in \left(\frac{1}{\delta}, \frac{1}{\delta\beta}\right) \tag{14}$$

$$Kaldor - Hicks : \alpha > 1 + r. \tag{15}$$

In both, greater smoking damage α makes a ban more attractive. Under the precommitment criterion, the smoking damage has to be large relative to the time preference discount factor, because under hyperbolic discounting ($\beta < 1$) Self 1 attaches more weight on period 1 actions than Self 0 would like. Under the intraself Kaldor-Hicks criterion, the smoking damage has to be large relative to the interest rate, because it then outweighs the fact that Self 1 gets less consumption out of having a period 1 dollar than Self 2, able to wait a period and earn interest on it, would get from that same dollar.

5. Borrowing and Saving in the Linear-Utility Model

Should the Social Planner Permit Borrowing?

I have assumed borrowing was impossible. What if the social planner could replicate borrowing across selves by a system of age-dependent taxes and transfers? Under laissez faire, each self consumed his entire endowment rather than save. Would they now borrow?

The precommitment criterion says that the social planner should permit borrowing, which would transfer all consumption to period 0. Self 0 would choose to consume all three selves' endowments because the rate of time preference is high and utility is linear in consumption.

The intraself Pareto and Kaldor-Hicks criteria lead to a different result: they would not instruct the government to replicate borrowing. Permitting borrowing would help Self 0, but hurt Selves 1 and 2. Moreover, there is no exchange of goods for money on which to base a potential Pareto improvement. These criteria are concerned with the production and allocation of goods, not with their distribution.

The multiple selves paradigm is a bigger change from standard normative economics than one might realize. The paradigm provides a plausible reason for banning smoking, but it weakens the argument against usury laws. The argument for banning smoking is that the ban hurts the present self less than it helps the future self. If the two could bargain across time, they would do so and both would be better off. The conventional argument against usury laws is that if someone accepts a loan, the benefit in current consumption must be worth the loss in future consumption. That is just to say that Self 0 benefits from borrowing, however, despite the loss to Selves 1 and 2. A usury law banning borrowing would hurt Self 0 but benefit Selves 1 and 2, so under the intraself Kaldor-Hicks criterion, we cannot say that the change is either a gain or a loss.

We would have to resort to some other argument to justify the allocation of property rights among the three selves. Richard Posner (1985, p. 93) says "If you asked a random group of economists how to assign property rights in a new society with a literate population so as to maximize the prices (time quantities), explicit and implicit ("shadow"), asking and offer, in the society they would almost certainly begin by giving each mentally competent adult the property rights to his own labor." Here, that would imply a ban on

borrowing (unless the future self would also favor borrowing), so that each self would be able to keep (or pass along) the fruits of his own period's labor.

The Effect of Borrowing on the Optimality of Smoking

What is the effect of borrowing on smoking? The laissez faire result is that Self 0 would borrow and consume the entire endowment, and Self 1 would smoke. The extreme consumption result occurs because of the assumptions of linear utility and the rate of time preference being higher than the interest rate.

Under the precommitment criterion, zero consumption for Selves 1 and 2 is optimal, but whether Self 1 should smoke depends on whether condition (14) is satisfied.

The intraself Pareto improvement and Kaldor-Hicks criteria, agree with the precommitment criterion in not requiring any change from zero consumption by Selves 1 and 2. Any such change would harm Self 0, and the harm to Self 0 would be exactly balanced by benefit to the other two selves, a matter of redistribution rather than any improvement in social surplus.

What is more surprising is that unlike when borrowing was impossible, the Kaldor-Hicks criterion now does not require a ban on smoking whenever the present value of utility harm exceeds the utility benefit ($\alpha > 1 + r$).

There is no scope for mutually beneficial trade of wealth for wealth, but the possibility of borrowing has also eliminated the possibility of mutually beneficial trade of wealth for a smoking ban. As a result of Self 0's borrowing, Selves 1 and 2 have zero wealth. Self 2 cannot offer any wealth to Self 1 in exchange for the smoking ban.

Self 0 could offer some wealth to Self 1 in exchange for a smoking ban. Would he offer enough?

The effect on Self 0's utility of a bargain in which he gives P to Self 1 in exchange for a smoking ban is

$$-P + \beta\delta(-1 + P) + \beta\delta^2\alpha \tag{16}$$

The minimum P that Self 1 would accept is $P = \frac{1-\beta\delta\alpha}{1+r}$. Substituting

that P into the expression for Self 0's utility change and solving for α yields the condition for Self 0's utility to rise from the bargain:

$$\alpha > \frac{1 + \delta\beta r}{\beta\delta(1 - \beta\delta + \delta(1 + r))} \tag{17}$$

Condition (17) is stronger than the precommitment criterion's condition of $\alpha > 1/\delta$, but weaker than Self 1's condition for not smoking, $\alpha > 1/(\beta\delta)$. If $\beta = 1$, it is equivalent to $\alpha > 1/\delta$, and regulation is unnecessary because Self 1 will abstain from smoking whenever Self 0 would be willing to pay him not to smoke.[2]

Result 2: If the person borrows enough against his future income, the Kaldor-Hicks criterion can support a smoking ban only if discounting is quasi-hyperbolic.

The result is even stronger if we drop the consume-early assumption and replace it with a "consume-late" assumption:

$$\delta\beta > \frac{1}{1 + r}. \tag{18}$$

In equilibrium Selves 0 and 1 would save their entire incomes, and Self 1 would smoke if $\alpha < \frac{1}{\delta\beta}$. Foreshadowing Result 3 below, however, Kaldor-Hicks would no longer support a smoking ban. In our hypothetical transaction, Selves 1 and 2 transfer consumption to Self 1, but now Self 1 would find this undesirable, a penalty rather than a reward. Under the consume-late assumption, Self 1 gains utility of 1 from his own consumption of 1 unit of income, but utility of $\frac{1+r}{\delta\beta} > 1$ from Self 2's consumption. Self 0 would be passing along his entire endowment to Self 1 anyway, and so could not credibly condition the payment on Self 1 not smoking. Hence, potential gains from trade are absent.

6. The Model with Concave Utility and Money Transfers across Time

[2] The right-hand- side is less than $1/\beta\delta$ because $(1 + \delta\beta r)\beta\delta < \beta\delta(1 - \beta\delta + \delta(1 + r))$ because $\beta\delta < \beta\delta(1 - \beta\delta + \delta$ and $(\delta\beta r)\beta\delta < \delta\beta r$. The right-hand-side is at least as great as $1/\delta$ because $(1 + \delta\beta r) \geq \beta(1 - \beta\delta + \delta(1 + r))$ because $1 \geq \beta(1 - \beta\delta + \delta)$ because that reaches a maximum at$\beta = 1$, when $\beta(1 - \beta\delta + \delta) = 1$.

So far, consumption decisions have been corner solutions because of the assumption of linear utility. What if we introduce a motive for partial, interior-solution "bequests" from one self to a later self? One might guess from Barro (1974) (on "Ricardian equivalence" and intergenerational debt) and Becker (1974) (on intrafamily transfers) that bequests would have an important effect, because both of those authors find that money transfers can neutralize behavior changes.

Let us now make the utility of consumption the strictly concave $U(C)$ with $U' > 0$, $U'' < 0$, and $C \overset{limit}{\to} 0$ $U'(C) = \infty$. Assume that borrowing is limited to an amount $\overline{B}$ in borrowing period dollars. Let us also abandon the consume-early assumption.

Imagine a Kaldor-Hicks transfer of P from Self 2 to Self 1 to ban smoking, with P measured in period 2 dollars. This payment will cost Self 2 amount P and increase Self 1's consumption by $P/(1+r)$. In addition, Self 1 may be borrowing amount $B \leq \overline{B}$, which will cost Self 2 amount $(1+r)B$ in consumption.

Let W_1 equal Self 1's income of W plus whatever Self 0 saves, minus whatever Self 1 borrows from him. Self 1 is constrained to spend no more than his income on present consumption, so

$$C_1 \leq W_1 + \overline{B} + \frac{P(1-X)}{1+r}. \tag{19}$$

If this is binding, Self 1 is borrowing the full amount $\overline{B}$. Otherwise, he is saving or borrowing less than $\overline{B}$.

Self 2's consumption equals his income of W plus $(1+r)$ times the amount Self 1 saves, $(W_1 - C_1)$:

$$C_2 = W + (1+r)(W_1 - C_1) \tag{20}$$

Self 1's maximization problem is to choose X and C_1 to maximize

$$U_1(C_1, C_2, X) = X + U(C_1) + \beta\delta(-\alpha X + U(C_2)) \tag{21}$$

subject to the constraints

$$\begin{array}{lll} (a) & C_1 & \leq W_1 + \overline{B} + \frac{(1-X)P}{1+r} \\ (b) & C_2 & = W + (1+r)(W_1 - C_1) \end{array} \tag{22}$$

Consider first the case where constraint (a) is not binding. Instead, Self 1 finds it optimal to save, or to borrow less than $\overline{B}$. Then we can substitute for C_2 from constraint (b) and write the problem as to maximize by choice of C_1 and X

$$X + U(C_1) + \beta\delta(-\alpha X + U(W + (1+r)(W_1 - C_1)). \tag{23}$$

The problem needs no budget constraint, because income does not constrain the choice of C_1.

The condition for choosing $X = 1$ is

$$1 - \beta\delta\alpha > 0, \tag{24}$$

and the first order condition with respect to C_1 is

$$U'(C_1) + \beta\delta(1+r)U'(W + (1+r)(W_1 - C_1) = 0 \tag{25}$$

The price P appears in neither of the two optimality conditions. Hence, Self 2's offer of P in exchange for the smoking ban makes no difference. Either Self 1 would abstain from smoking anyway, or he would not be persuaded by the offer of money from Self 2. If discounting is exponential, then it is only Self 2 who would be willing to offer money to Self 1 in exchange for not smoking, because Self 0 shares Self 1's incentives, as we saw earlier. Thus, Kaldor-Hicks no longer can support a smoking ban, as Result 3 states.

Result 3: If discounting is exponential and a person's borrowing constraint is non-binding (including if he is saving a positive amount), then the intraself Kaldor-Hicks criterion does not support a smoking ban.

If Self 1 cares enough about Self 2 to pass savings along to him, then the intraself Kaldor-Hicks criterion does not require Self 1 to give up smoking for the sake of Self 2. If Self 1 is selfish enough not to save, the criterion does

require him to give up smoking. There is no Kaldor-Hicks inefficiency, but it is not because the altruist abstains from seemingly inefficient behavior, but because he will not accept compensation for abstaining.

The intuition behind Result 3 is that when Self 1 is saving, the marginal dollar of saving results in period 2 consumption that gives him as much utility (discounted though it is) as if he spent that dollar on period 1 consumption. He would be hurt by anything that transferred that consumption back to period 1. Thus, if he consumes the Kaldor-Hicks payment given him by Self 2, he is actually worse off.

Similarly, if Self 1 is borrowing less than the limit, he has decided that though overall he would like to reduce Self 2's consumption to increase his own, at the margin, once he has borrowed the appropriate amount, his utility is the same from having Self 2 consume a little as from having himself consume more.

If constraint (a) is binding, there might or might not be a Kaldor- Hicks transfer that would induce Self 1 to stop smoking. Consider what happens if our person care only about the present, so $\delta = 0$. Then P^* is such that

$$1 + U(W_1 + \overline{B}) = U\left(W_1 + \overline{B} + \frac{P^*}{1+r}\right). \tag{26}$$

Self 2 will be willing to pay P^* if

$$U(W - (1+r)\overline{B} - P^*) \geq -\alpha + U(W - (1+r)\overline{B}). \tag{27}$$

If U'' is close to zero, so utility is almost linear, and if r is low then the gain in Self 1's consumption utility from the transfer of P^* will be close to the loss in Self 0's, and Self 2 would indeed be willing to pay P^*. If, on the other hand, the utility of cosumption flattens beyond Self 1's initial consumption level of $W_1 + \overline{B}$ but increases enough for lower levels of consumption, then Self 1 will be unwilling to deprive Self 2 of consumption via the transfer of P^*. Thus, there is a possibility of regulation being optimal for borrowing-constrained individuals.

This may be easier to see in a numerical example. Assume that $W = 1$,

$K = 6$, $r = 0$, $\overline{B} = .2$, and

$$\begin{aligned} U'(C) &= 2 \quad \text{if} \quad C < 2 \\ &= 1 \quad \text{if} \quad C > 2. \end{aligned} \tag{28}$$

(1) Suppose $\beta\delta < .5$. There is no saving, and Self 2 will pay Self 1 not to smoke. The laissez faire outcome will be $C_0 = 7.2, C_1 = 1, X = 1$, and $C_2 = .8$. If Self 0 or 1 deviates by saving more, he will lose utility at a rate of 1 per reduced consumption unit and gain it at only a rate of $2\beta\delta$ from his successor's increased consumption. Self 0 will borrow up to the limit of $B = .2$ from Self 1, and Self 1 will borrow up to the limit from Self 2. Smoking raises Self 1's utility by 1 directly and reduces it only by $\beta\delta\alpha$ through its effect on Self 2.

The Kaldor-Hicks Criterion can support a smoking ban if $\beta\delta < .5$ even if $\beta = 1$ because Self 1 would accept as little as $(.5 - \beta\delta\alpha$ to stop smoking, whereas smoking only increases utility by 1) and Self 2 would be willing to pay up to $\alpha/2$, an amount greater than .5 since $\alpha > 1$.

(2) Suppose instead that $.5 < \beta\delta < 1$. The laissez faire outcome will be $C_0 = 5, C_1 = 2, X = 1$, and $C_2 = 2$. If Self 0 or 1 deviates by saving more, he will lose utility at a rate of 1 per increased consumption unit and gain it at the lower rate of $\beta\delta$ from his successor's increased consumption. If he deviates by saving less, he will gain utility at a rate of 1 per increased consumption unit and lose it at the higher rate of $2\beta\delta$.

There is no payment by Self 2 that will persuade Self 1 to stop smoking. A payment of P by Self 2 will have a net effect of $(P - 2\beta\delta P) < 0$ on Self 1's utility, so Self 1 would refuse the payment even if it were not contingent on not smoking.

Self 0, on the other hand, could pay P to Self 1 in addition to the laissez faire bequest of 3 and Self 1 would have no objection. Self 0's utility increase by

$$-P + \beta\delta(-1 + P) + \beta\delta^2(\alpha). \tag{29}$$

Self 1 would accept as little as $P = (1 - \beta\delta\alpha)$ to not smoke. Substituting for P in expression (29) tells us that Self 0's utility would rise after making

that payment to stop smoking if

$$\alpha > \frac{1}{\delta\beta(1-\beta\delta+\delta)}. \tag{30}$$

There exists a range of values of α for which Self 1 would not abstain from smoking unilaterally but would after a payment that Self 0 would be willing to pay. This relies on hyperbolic discounting, since if $\beta = 1$condition (30) reduces to $\alpha > 1$, the same condition as for Self 1 to unilaterally abstain from smoking. Thus we have Result 4.

Result 4: *The intraself Kaldor- Hicks criterion can support a smoking ban for someone whose borrowing constraint is binding if discounting is hyperbolic.*

9. Concluding Remarks

Applying the Kaldor-Hicks criterion to a multiple selves model does provide a justification for a social planner to ban vice, though with caveats. First, the vice must cause more immediate harm than it creates future benefit. Second, the extent of the harm must outweigh the cost of transferring consumption to an earlier period. Third, if discounting is exponential, the person must be constrained in his borrowing. Otherwise, the earlier self would not value transfers to himself from the future self and there are no potential gains from trade. Even if the person engaged in the harmful behavior is making a bequest, however, if discounting is hyperbolic the Kaldor-Hicks criterion can justify restricting his behavior for the sake of selves prior to himself in time who care more than he does for far-future consequences.

References

Barro, Robert J. (1974) "Are Government Bonds Net Wealth?" *Journal of Political Economy*, 82(6): 1095-1117 (November/December 1974).

Becker, Gary (1974), "A Theory of Social Interactions," *Journal of Political Economy*, 82: 1063-1093 (November/December 1974).

Bernheim, B. Douglas & Antonio Rangel (2007) "Beyond Revealed Preference: Choice Theoretic Foundations for Behavioral Welfare Economics," http://papers.nber.org/papers/w13737 (December 2007).

Bhattacharya, Jay & Darius Ladkdawalla (2004) "Time-Inconsistency and Welfare," NBER Working Paper No. W10345, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=515239 (March 2004).

Gruber, Jonathan & Botond Koszegi (2001) "Is Addiction "Rational"? Theory and Evidence," *Quarterly Journal of Economics*, 116(4): 1261-1303 (November 2001).

Krusell, Per & Anthony A. Smith Jr. (2003) "Consumption-Savings Decisions with Quasi-Geometric Discounting," *Econometrica*, 71(1): 365-375 (January 2003).

Laibson, David (1997) "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 112(2): 443-477 (May 1997).

Loewenstein, George (1992) "The Fall and Rise of Psychological Explanations in the Economics of Intertemporal Choice," pp. 3-34 of George Loewenstein and Jan Elster, eds., *Choice over Time*, New York:Russell Sage Foundation (1992).

O'Donoghue, Ted & Matthew Rabin (2003) "Studying Optimal Paternalism, Illustrated by a Model of Sin Taxes," *American Economic Review, Papers and Proceedings,* 93(2): 186-191 (May 2003).

Posner, Richard A. (1985) "Wealth Maximization Revisited," *Notre Dame Journal of Law, Ethics & Public Policy*, 2: 85-106.

Strotz, R. H. (1955) "Myopia and Inconsistency in Dynamic Utility Maximization," *Review of Economic Studies*, 23(3): 165-180 (1955-56).

Whitman. Glen (2006) "Against the New Paternalism: Internalities and the Economics of Self-Control," Policy Analysis no. 563, Cato Institute, http://www.cato.org/pub_display.php?pub_id=5531&print=Y (February 22, 2006).