Does rational addiction imply irrational non-addiction?

John Quiggin*

August 17, 2001

Abstract

In this note, the implications of rational addiction theory for the acquisition of habits and addiction are explored. It is shown that, if age-specific death rates are uniform, habits should either be acquired at the beginning of life or not at all. By contrast, under the more realistic assumption that death rates are higher at higher ages, the rational addiction model implies that any addiction with positive initial benefits should be taken up as soon as the age-specific death rate is sufficiently high.

*Australian Research Council Fellow, Department of Economics, Faculty of Economics and Commerce, ANU ACT 0200 Australia.
The rational addiction model of Becker and Murphy (1988) has stimulated a number of theoretical and empirical contributions and has been widely cited in the public debate over issues such as gambling policy (Productivity Commission 1999). The crucial claim made by Becker and Stigler is that it may be rational for individuals to acquire habits such as drug addiction, taking full account (in discounted present value terms) of the future costs. Empirical analysis of the rational addiction hypothesis, has, in general, been based on statistical tests, in which the null hypothesis is that of myopic behavior, taking no account of the future consequences of current consumption of addictive goods. In most cases, the null hypothesis has been rejected (Becker, Grossman and Murphy 1994). However, as noted by the Productivity Commission (1999), the fact that behavior is not entirely myopic, does not imply that it is fully rational. It is, therefore, of interest to consider implications of the rational addiction hypothesis that may lead to more precise tests of the hypothesis.

Given that the acquisition of a habit may be consistent with the maximization of rational, dynamically consistent preferences, the question naturally arises as to when it is rational to acquire the habit. Casual observation suggests that most habits, and particularly most addictions are acquired in adolescence or early adulthood. However, given that the adverse health and financial consequences of addictions are frequently suffered many years after the initial consumption benefits are received, the question arises of whether it would not be more rational to defer the acquisition of addictions towards the end of the lifespan. Under this plan, the rational individual could enjoy the initial consumption benefits, but would not be alive to incur the costs.

The relationship between rational addiction models and lifespan has previously been considered by Levy (2000) whose focus is different from that of the present paper. Levy
considers increased risk of death as a cost of consuming addictive goods and shows that
the individual’s negative appreciation of the addictive stock is reinforced by the increased
risk of death. Hence, the rate of change of the consumption of the addictive commodity
is lower than that obtained when the effect of addiction on the probability of dying is
ignored.

In this note, attention is focused on the decision to acquire a habit. It is shown that,
if age-specific death rates are uniform, habits should either be acquired at the beginning
of life or not at all. By contrast, under the more realistic assumption that death rates are
higher at higher ages, the rational addiction model implies that any addiction with positive
initial benefits should be taken up as soon as the age-specific death rate is sufficiently high.
Since this implication is not consistent with observed behavior, the validity of the rational
addiction model is called into question.

1 Model

Consider an individual with a stochastic lifetime who has the opportunity to consume a
single habit-forming good. Consumption of the good in period $t$, $t = 1, 2 \ldots$ is denoted by
$h_t \in \mathbb{R}_+$. Utility in period $t$ is given by a time-independent utility function $u(h_t, H_t)$ where

$$H_t = \begin{cases} 0 & t = 1 \\ \sum_{s=1}^{t-1} \alpha_s h_{t-s} & t > 1 \end{cases}$$

is a stock measure of past consumption and the $\alpha_s \in \mathbb{R}_+$ are non-negative weights. We
normalize by assuming $u(0, 0) = 0$. We say that the individual acquires a habit in period
t if \( h_t > 0 \) and \( h_s = 0, s < t \). The habit is fatal at stock level \( H \) if \( u(\bullet, H) \equiv 0 \).

A habit is a ‘bad’ habit if for all non-zero sequences \( h_t, t = 1, 2 \ldots \), there exists some \( t^* > 1 \) such that
\[
u(h_t, H_t) \geq 0, t \leq t^*
\]
\[
u(h_t, H_t) < 0, t > t^*.
\]
and a ‘good’ habit if for all non-zero sequences \( h_t, t = 1, 2 \ldots \), there exists some \( t^* \) such that
\[
u(h_t, H_t) \leq 0, t < t^*
\]
\[
u(h_t, H_t) > 0, t \geq t^*.
\]

We will rule out trivial cases by assuming that there exists some sequence \( h \), and time-period \( t \) for which \( u(h_t, H_t) > 0 \). For simplicity, in the case of a bad habit, we will assume \( u(h_1, 0) > 0 \).

The individual’s probability of death at time \( t \) is denoted \( d_t \) and is assumed independent of \( h \) and \( H \). Also, for \( T \geq t \) we denote by
\[
D_t^T = \begin{cases} 
1 & T = t \\
1 - \prod_{s=t+1}^T (1 - d_s) & T > t
\end{cases}
\]
the probability that an individual alive at time \( t \) will be dead by time \( T \).

We further assume that the individual displays stationary discounting behavior. That is, there exists a strictly monotone decreasing vector \( \beta_s, s = 1, 2 \ldots \), with \( \beta_1 = 1 \) such that at each time \( t \), the individual chooses \( h^* = (h^*_1, h^*_2, \ldots) \) to solve
\[
Max_{\{h_s : s=1,2,\ldots\}} \sum_{s=1}^{\infty} \beta_s (1 - D_t^s) u(h_s, H_s)
\]
\[\text{\footnote{It would be straightforward to extend the model to make the effects of }H, \text{ including death risks, stochastic, as in Levy (2000). However, the additional complexity would not yield additional insight into the problem of interest here.}}\]
such that

\[ H_s = \sum_{r=0}^{t-1} \alpha_{s+r} h_{t-r} + \sum_{r=1}^{s-1} \alpha_r h^*_s - r \]

and sets \( h_t = h^*_t \). This formulation allows for both exponential and hyperbolic discounting. Only in the former case are preferences dynamically consistent.

We now observe that if the probability of death is constant, the individual will either acquire a habit in period 1 or she will never acquire a habit. Observe that if \( d_t \) is some constant \( \delta \), \( D_t^T = (1 - \delta)^{T-t} \)

**Theorem 1** If \( h_t, t = 1, 2 \ldots \) solves the problem above, then either

\[ h_t = 0 \quad \forall t; \text{ or} \]

\[ h_1 > 0 \]

**Proof:** Consider a consumption sequence in which the individual acquires the habit in period \( t > 1 \). Then some \( h^* \) with \( h^*_1 = h_t \), solves the problem at time \( t \)

\[ Max \{ h_s; s=1,2,\ldots \} \sum_{s=1}^{\infty} \beta_s (1 - (1 - \delta)^{s-t}) u(h_s, H_s) \]

where

\[ H_s = \sum_{r=1}^{s-1} \alpha_r h^*_s - r \]

But it follows immediately that the same \( h^* \) is a solution to the individual’s problem at time 1, contrary to the assumption that the habit is acquired in a later period. ■

In the case of exponential discounting, the individual maximizes a single dynamically consistent objective function and the proof above may be modified to show that, if acquiring the habit in period \( t \) is beneficial, acquiring the habit in period 1 instead of period \( t \) raises
the (necessarily positive) value of the objective function by a factor of $1/\beta_1$. Note that this argument applies to all habits, including both good and bad habits.

Theorem 1 appears to fit the stylized facts of habit-formation, in that most habits are acquired early in adult life or else not acquired at all. However, the assumption that the age-specific death rate is constant is not consistent with the data. Consider the alternative hypothesis that $d_t$ rises over time, approaching 1 at very high ages. It is easy to see that, in the case of constant death rates, the higher is the death rate $\delta$, the more attractive are bad habits and the less attractive are good habits. Intuitively, therefore we might expect that if the age-specific death rate rises with age, the acquisition of bad habits will become more attractive as individuals age. In fact, we have:

**Theorem 2** (i) Suppose the habit is a bad habit and $d_t$ is monotone increasing with $d_t \to 1$ as $t \to \infty$, and $h_t$, $t = 1, 2 \ldots$ solves the problem above, then $h_t > 0$ for some $t$.

Proof: Since the habit is a bad habit, for any $h^*$ there exists $u(h^*_1, 0) > 0$. Moreover, since $d_t \to 1, D^*_t \to 1, s > t$, and we can choose $t$ such that

$$\left\| \sum_{s=2}^{\infty} \beta_s (1 - D^*_t) u(h^*_s, H^*_s) \right\| > u(h^*_1, 0)$$

so

$$\sum_{s=2}^{\infty} \beta_s (1 - D^*_t) u(h^*_s, H^*_s)$$

It follows that the sequence in which the habit given by $h^*$ is acquired at time $t$, will be chosen in preference to any sequence in which the habit is acquired after time $t$ or not at all.

In most modern societies, and leaving aside deaths that may be attributed to addictive or risk-taking behavior, age-specific death rates are low and roughly constant at ages from
adolescence to middle-age (say 15 to 60) then rise steadily. In cases where the adverse consequences of addiction are evident within, say 20 years, the analysis above implies a bifurcated solution. Individuals with high rates of discount or a strong preference for the consumption benefits of bad habits will acquire those habits early in life. Those with lower, but still positive, consumption benefits will abstain until an advanced age, then acquire bad habits as the likelihood of incurring the negative consequences declines. Only those for whom the initial consumption benefits are negative will be lifelong abstainers. Assuming a continuous distribution of tastes, in cases where the first and third patterns are observed, so should the second.

The fact that older people do not appear to acquire dangerous habits with any frequency casts doubt on the rationality of their decisions with respect to the acquisition of habits. It would be possible to conclude from this that addiction decisions are rational, but non-addiction decisions are not. This seems unduly paradoxical however. Alternatively, it would be possible to ‘save the phenomena’ through the use of _ad hoc_ hypotheses. For example, it might be assumed that abstention from addictive commodities is itself addictive, or that young people gain greater utility from the consumption of addictive commodities than older people. Neither approach seems very appealing.

A more promising interpretation of the results may be derived from the fact that the same analysis applies both to dynamically consistent and dynamically inconsistent preferences. As a number of writers have observed, individuals who know themselves to behave according to dynamically consistent preferences, but who possess higher-order dynamically consistent preferences, may adopt devices to prevent themselves from making decisions they will regret in retrospect (Elster 1979). For example, individuals may seek to
convince themselves that smoking is morally wrong, thereby ensuring immediate negative consequences from consumption. Having adopted such self-binding devices to preclude youthful acquisition of bad habits, individuals may find themselves also precluded from taking up such habits in old age, even if acquisition of a habit would yield an improvement in expected welfare evaluated using dynamically consistent preferences.

2 References


