Volume Title: Risky Behavior among Youths: An Economic Analysis

Volume Author/Editor: Jonathan Gruber, editor

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-31013-2

Volume URL: http://www.nber.org/books/grub01-1

Publication Date: January 2001

Chapter Title: Risky Behavior among Youths: Some Issues from Behavioral Economics

Chapter Author: Ted Oâ€™Donoghue, Matthew Rabin

Chapter URL: http://www.nber.org/chapters/c10686

Chapter pages in book: (p. 29 - 68)

1

Risky Behavior among Youths
Some Issues from Behavioral Economics

Ted O'Donoghue and Matthew Rabin

The goal of this volume is to provide an economic analysis of *risky behavior among youths,* loosely defined to be behavior by people under age nineteen that might have important future ramifications. Examples of such behaviors include smoking, drinking, having unprotected sex, and engaging in crime. The traditional approach used by economists would seem to have important shortcomings in this realm. The rational-choice model provides a powerful tool for understanding behavior and has yielded an array of insights across a broad range of human activities. But a growing number of economists have come to recognize that the rational-choice model is inaccurate in some systematic and important ways, and that to take full advantage of the economic insights and methodology, economists must embrace insights from psychology and other social sciences so as to make our models more relevant and realistic.

While the shortcomings of the rational-choice model are relevant for people of all ages, they seem particularly acute for youths. In this chapter, we discuss how recent efforts combining psychology and economics can be used to help understand risky behavior by adolescents. We are not (in the least) experts in youthful risky behavior and do not provide a very broad perspective on all the psychology relevant to this topic. Our goal here is less ambitious and more specific: to explore what some of the main

Ted O'Donoghue is assistant professor of economics at Cornell University. Matthew Rabin is professor of economics at the University of California, Berkeley, and a member of the Russell Sage Foundation Behavioral Economics Roundtable.

The authors thank George Loewenstein and Ken Train for useful conversations and Jon Gruber and participants in the NBER conference for useful feedback. For financial support, the authors thank the National Science Foundation (award 9709485) and the NBER project; Rabin thanks the Russell Sage, Alfred P. Sloan, and MacArthur Foundations.

insights and issues raised by recent research in behavioral economics suggest about risky behavior by adolescents.[1] Our focus is on the potential for applying formal behavioral-economic models to theoretical and empirical research on youthful behavior.

Why should economists be motivated to study risky behavior among youths? It could be that we have only a "positive" interest—we are interested merely because we would like to understand society better and estimate or predict drug use, criminal behavior, suicide, and so forth. For most people, however, the interest in youths' behavior is motivated by "normative" considerations. Parents, citizens, policy makers, and even many economists are interested not merely in predicting whether sixteen-year-olds start smoking, use cocaine, get pregnant, or kill themselves, but also in understanding the welfare consequences of these behaviors.

One important normative question is whether risky behavior among youths creates negative externalities that affect other members of society. Negative externalities are obviously an important facet of many of the behaviors studied in this volume—for example, crime, or such behaviors as alcohol and drug use that can lead to crime and automobile accidents, or any behavior that leads to increased dependence on the state. Similarly, a major concern in preventing early pregnancy among girls is the harm done to society (and to the children born). A reasonable guess is that youths have a higher preference than adults for activities that create negative externalities and that society may therefore be especially keen to curtail these activities.

But, clearly, most of us are concerned about risky adolescent behavior in large part because we believe that adolescents are not behaving in their own best interests and because we feel that something should be done to help them. This concern is warranted by the clear evidence that even adults often do not behave in their own best interests. Of course, even if this concern motivates our research, we could help study suicide, drug use, sex, and so forth without taking an explicit stand on whether these behaviors are good or bad and then let policy makers and other audiences use our behavioral conclusions to further their normative concerns. We intend that this chapter help in this way. We believe, however, that behavioral economics provides some valuable insights into the precise nature of the harm that youths may cause themselves, and, hence, the most central contribution of behavioral economics may be helping policy makers understand the connection between behavior and welfare. This will be our main emphasis.

Our welfare emphasis may be controversial. Over the years, economists have developed an aggressive agnosticism with regard to welfare analysis for individual choice, refusing to make any judgments that people are not

1. For a general overview of some of the topics studied by behavioral economics, see Thaler (1992), Camerer (1995), and Rabin (1998).

behaving in their own best interests. Caution is, of course, warranted because, more often than not, people probably have a better idea of what is in their own best interests than do economists, other social scientists, and policy makers. But this caution has largely transformed itself into an a priori presumption that people always behave in their own best interest. There are some realms where common sense, compassion, and intellectual curiosity all tell us that we should consider the possibility that people may not be behaving in their own best interests. Risky behavior by youths is one of those realms.[2]

Of course, we should not replace welfare agnosticism with a "promiscuous paternalism" that provides undisciplined assertions that others' behaviors are not good for them and that we know better what they should do. Rather, we need a principled way in which to study when and how people make errors, what types of interventions might help mitigate these errors, and when we can have some confidence that these interventions help more than they harm. When considering risky behavior among youths, it is important to avoid both opinionated moralism as to what is the right behavior and naive faith that sixteen-year-olds make no predictable mistakes in their choices. By identifying systematic patterns in errors that people make, behavioral economics provides just such an approach.

The development of behavioral economics has not been targeted at analyzing the behavior of adolescents. The literature has developed with a belief that people make errors at all ages. It is, indeed, worth stressing the similarities in the mistakes made at different ages. A fifty-year-old may sacrifice too much for sexual gratification just as a fifteen-year-old may, or a thirty-six-year-old may drive too soon after drinking just as a sixteen-year-old may. Errors associated in the common imagination with one's youth are often made throughout life, and bad decisions attributed to youth may not be as strongly associated with age as is often claimed.[3]

That said, there are likely broad differences between adolescents and adults in many of the realms that we discuss. Young people almost surely make more mistakes. In section 1.1, we briefly discuss psychological evidence on how youths make decisions and how youths differ from adults. As we proceed, we shall relate some of our theoretical analysis back to

2. Given our focus on welfare, we shall not devote this essay to proving beyond (an economist's) doubt that behavior predicted outside the rational-choice framework is fundamentally inconsistent with rational choice. We doubt the general usefulness of the widespread methodology of employing post hoc attempts to fit behavior into the rational-choice framework without any inquiry as to whether it is the correct explanation. In making welfare assessments, this approach is clearly inappropriate.

3. This interpretation has been endorsed by some of the leading political figures of our day. As many of those attacking Bill Clinton for lying about his extramarital affairs were exposed for their own misbehavior, *youthful indiscretion* became something of a catchphrase for bad behavior at virtually any age. This is, e.g., precisely the term that seventy-four-year-old Congressman Henry Hyde used to describe the extramarital affair he had had at age forty-one.

this evidence, but more often we speculate on how some of the obvious but little-researched differences between youths and adults relate to the behavioral phenomena that we consider.

Before proceeding to this evidence, however, we briefly outline the other sections of this chapter. We focus on three types of questions that can all be usefully thought of in terms of their relation to a rational-choice base model. Throughout the chapter, we assume that a person's overall well-being is determined by adding up her well-being at each moment. We refer to a person's well-being in period t as her *instantaneous utility* in period t, which we denote by u_t. To allow for the possibility that the person's instantaneous utility in period t is stochastic, let S_t be the set of possible states for period t, and for $s \in S_t$, let $p_t(\cdot, s)$ and $u_t(\cdot, s)$ be the probability of state s and instantaneous utility function in state s, respectively. The person's expected instantaneous utility in period t is therefore $\Sigma_{s \in S_t} p_t(\cdot, s) u_t(\cdot, s)$. Finally, we assume that the person's overall well-being from the perspective of period t, which we denote by W^t, is given by

$$W^t \equiv \sum_{\tau=t}^{T} \left[\sum_{s \in S_\tau} p_\tau(\cdot, s) u_\tau(\cdot, s) \right].$$

Section 1.2 is devoted to discounting. The reader will notice that our base model assumes no discounting: the expected instantaneous utilities for all periods are weighted equally. We begin our discussion in section 1.2 by arguing that, from a normative perspective, there should be no discounting. Just because an adolescent cares very little for her thirty-five-year-old self, it does not follow that *we* should care very little for her thirty-five-year-old self. We then discuss some recent approaches that formalize the ways in which people underweight the future consequences of their actions and the lessons that such approaches have for youthful behavior. We discuss excessive myopia per se (pure underweighting of the future) and the tendency to have a time-inconsistent preference for immediate gratification (pursuing immediate gratification on a moment-by-moment basis in a way that does not match the person's own long-run best interests). We also discuss the closely related error of overoptimism about future self-control problems, which implies an underestimate of future misbehavior.

Section 1.3 discusses ways in which people incorrectly predict future instantaneous utilities. Hence, while section 1.2 explores ways in which people pay too little attention to the future consequences of their actions, section 1.3 explores ways in which people incorrectly predict how they will feel in the future about those consequences. We describe some systematic ways in which youths may underestimate the future harm caused by their current behavior because they do not fully recognize the extent of day-to-day fluctuations in tastes, or the extent to which peer pressure will tempo-

rarily influence their preferences, or just how much their preferences when older will differ from their youthful preferences.

In section 1.4, we discuss some issues with respect to the probability function p_τ, focusing on the logic of repeated risky choices. Since both past and future risky behavior may change the consequences of current risky behavior—in particular, behaving in a risky way at other times may affect the marginal risk accrued by behaving in a risky way now—certain types of risky behavior can be understood only in an intertemporal context. We flesh out the logic of repeated risky choices largely in a rational-choice setting but then explore how those implications might differ when people make errors in assessing risks, have self-control problems, or incorrectly predict their own future preferences.

We conclude in section 1.5 with a more general discussion of the issues raised and lessons learned from the analysis in this chapter.

1.1 Evidence on Adolescent Decision Making

In this section, we review some evidence from psychology and related fields concerning how youths make decisions and how youths differ from adults.

The paradigm of psychological research most closely related to the economic approach is behavioral decision theory, which examines people's actual decision-making processes and how these compare to "normative" (Bayesian) decision making. Behavioral decision theory often breaks down decision making into a sequence of steps so that performance on individual steps can be analyzed in isolation. There are extensive literatures that identify weaknesses in the ways in which adults perform these steps (see, e.g., Camerer 1995; and Fischhoff 1988).

There is a smaller literature that attempts to analyze the decision-making performance of adolescents and how adolescents differ from adults. The general themes in this literature seem to be that there is little evidence, particularly evidence that makes direct comparisons between adolescents and adults, that much of the evidence is, as assessed by researchers whose analysis most resembles the perspective of economists, weak owing to methodological problems, and that, while what little evidence there is suggests a few differences between adolescents and adults, on the whole they are remarkably similar. Indeed, a review article by Furby and Beyth-Marom (1992) emphasizes that many common conceptions of how youths differ from adults do not seem to be borne out by the evidence.[4]

Many studies ask subjects to formulate lists of potential consequences of various behaviors. Beyth-Marom et al. (1993) is one of the few studies

4. For reviews in this vein, see also Fischhoff (1992) and Beyth-Marom and Fischhoff (1997).

of this type that directly compares adolescents and adults. Teens and parents were asked to generate possible consequences of several decisions (e.g., you were [your child was] at a party where marijuana was passed around and decided to smoke). Although there were a few differences—for instance, on average, adults generated slightly more consequences, and adolescents were slightly more likely to mention consequences involving social reactions—overall the most striking conclusion was the similarity between adults and adolescents.

Some list-the-consequences studies focus on the question of how future-oriented youths are. Lewis (1981) conducted a study that compares adolescents in three grade categories (grades 7–8, 10, and 12). In simulated peer-counseling sessions, subjects were presented hypothetical dilemmas and asked what advice they would give to a peer who faced these dilemmas. One of the main results was a significant increase with grade level in the mention of the potential risks and future consequences of decisions, which supports the hypothesis that there is an increase in future orientation through adolescence.[5] Further evidence of this hypothesis is reviewed in Greene (1986), who concludes that "adolescents, as compared to younger children: (1) demonstrate greater depth and extension of temporal perspective . . . ; (2) project a more complex, differentiated set of future expectations . . . ; and (3) describe future aspirations with greater planfulness, organization, and realism" (p. 100).[6]

Comparisons have been conducted between adolescents and adults not only in awareness of consequences but also in perceptions of the likelihood of those consequences. Quadrel, Fischhoff, and Davis (1993) test the conventional wisdom that youths are prone to feelings of invulnerability by asking subjects to assess the likelihood that various negative events would occur to themselves, an acquaintance, a close friend, and their parent or child. Subjects typically assessed similar likelihoods for themselves and for others. There was some evidence for feelings of invulnerability—conditional on assessing different likelihoods, subjects were twice as likely to assess lower likelihoods for themselves—but this invulnerability was not stronger for adolescents than for adults.[7]

In fact, there is evidence that youths are in some ways overly pessimistic about their future. Fischhoff et al. (2000) survey youths about personal probabilities of dying young. For a representative sample of fifteen- and sixteen-year-olds, the mean response to how likely it is that they would die

5. A contrary result, however, was that there seemed to be no difference across grade levels in recommendations as to whether peers or parents should be consulted for advice.

6. Greene conducts an experiment to determine whether such changes are correlated with the emergence of Piaget's formal-operations reasoning and finds at best very weak evidence.

7. In fact, the invulnerability was stronger among adults than among adolescents, but this seemed merely to reflect the plausible consensus among youths and adults that the adults *were* less vulnerable to many of the risks under consideration.

in the next year was 18.6 percent, whereas the statistical estimate is 0.08 percent. The mean response to how likely it is that they would die by age twenty was 20.3 percent, whereas the statistical estimate is 0.4 percent.

Researchers have also asked whether adolescents are competent decision makers. For instance, Weithorn and Campbell (1982) presented adolescents with hypothetical medical- and psychological-treatment decisions, finding that fourteen-year-olds scored as well as eighteen- and twenty-one-year-olds in competency, and Lewis (1987) concludes that, in terms of pregnancy and contraceptive decisions, adolescents may equal adults in their competence to reason.

While the studies discussed above examine hypothetical decisions, another literature examines how adolescents' perceptions of consequences, the likelihood of consequences, and the importance of consequences predict their own behavior. For instance, Bauman (1980) presented seventh graders with fifty-four potential consequences that might occur if they used marijuana and asked them to rate on a five-point scale both the likelihood and the importance of those consequences to themselves. These ratings were then found to be predictive of self-reported marijuana use by the same individuals one year later. A similar technique has been used by a variety of researchers to study cigarette smoking, drinking, and sexual intercourse. Furby and Beyth-Marom (1992) summarize (and criticize) these studies and conclude that, "in sum, what little evidence there is (with all its mentioned weaknesses) suggests that to at least some small extent teens choose to engage in behaviors which are more likely to bring consequences they perceive as positive and less likely to bring consequences they perceive as negative" (pp. 16–17).

Many of the studies in this volume also support this conclusion. For instance, Gruber and Zinman (chap. 2) find that youth smoking depends negatively on price, at least for older teens; Pacula et al. (chap. 6) find that youth marijuana use depends negatively on both price and the perceived risk of future harm; and Levine (chap. 4) finds that teenage women are less likely to have sex and more likely to use contraceptives when labor market conditions are good and when the perceived risk of HIV infection is high. These findings that adolescents react to costs and benefits suggest that youths are to some degree rational in pursuing their well-being. But, since all behavioral-economics models with which we are familiar assume that people respond to costs and benefits, such findings say nothing about the validity of the extreme rational-choice model.

The evidence reviewed above suggests that adolescents are similar to adults in terms of their ability to carry out the decision-making process. Youths seem to differ more from adults in how they value the consequences of decisions. In fact, research in developmental psychology that studies adolescent behavior focuses not on the decision-making process, but rather on what considerations matter most to adolescents. Much of

this research focuses on adolescent concerns with such things as identity formation, sexual-identity formation, and establishing autonomy and independence. This research suggests that adolescents may make decisions based primarily on these considerations, not on "objective" consequences. For instance, an adolescent male may drive fast so as to confirm his masculine identity, virtually ignoring the potential negative consequences. Baumrind (1987) even argues that many risk-taking behaviors by adolescents play an integral role in identity formation and making the transition to adulthood.[8]

There are, of course, other reasons why youths and adults might value consequences differently. Over the years, many studies have found that youths tend to score higher than adults on sensation-seeking and risk-taking behavior (e.g., Zuckerman, Eysenck, and Eysenck 1978; and Arnett 1994). And, presumably, youths are more concerned than adults are with how their peers will react to their behavior.

Differences in how youths and adults value consequences reflect differences in preferences, which in our model means differences in the instantaneous utility functions. If a young male engages in some risky behavior because it satisfies a need to confirm his masculine identity, or because it yields desirable sensations, or because it will provoke positive reactions from his peers, it seems natural to conclude that he has positive marginal instantaneous utility from engaging in the behavior.

The theoretical analysis developed below does not focus per se on how adolescent preferences differ from adult preferences. Instead, we focus on the ways in which youths fail to behave in their own best interests and for the most part remain agnostic about what those best interests are. But the fact that youths care a lot about such things as identity formation, sensation seeking, and peer reactions that tend to increase short-term benefits in a highly variable way may imply that the errors that we discuss are particularly problematic for youths, even if youths and adults do not differ in their inherent propensity for these errors.

The evidence on risk perceptions and differential preferences discussed above suggests some ways in which our analysis in section 1.3 of incorrectly predicting preferences and in section 1.4 of repeated risky choices may be especially applicable to youths. While the evidence discussed above comparing the future orientation of adolescents and that of adults is relevant to our analysis of self-control problems in section 1.2, we discuss more direct evidence on the relation between age and self-control problems when we discuss evidence on self-control problems more generally.

8. For recent research by economists that explores the role of identity, see Akerlof and Kranton (2000).

1.2 Trading Off Present versus Future Consequences

Most of the risky behaviors addressed in this volume involve a trade-off between short-term benefits and long-term costs. The decision whether to have sex involves a trade-off between the short-term benefit of sexual pleasure and the long-term cost of possibly getting pregnant or acquiring a sexually transmitted disease. The decision whether to commit a crime involves a trade-off between the short-term thrill or material benefits of the crime and the long-term costs of possibly going to jail. The decision whether to drink alcohol, smoke cigarettes, or use other drugs involves a trade-off between the short-term benefits of consumption and the long-term costs to future health, job prospects, and personal satisfaction. Each of these domains is quite rich, and, clearly, no single factor can explain the misbehavior of youths in these domains. But there is a simple error that plays a significant role across all these domains: excessive myopia in trading off present versus future consequences.

In order to discuss errors in trading off present versus future consequences, we must begin with a normative standard of how a person *should* trade off present versus future consequences to maximize her true well-being, which requires a discussion of the appropriateness of discounting. Economists use the notion *discounting* in a variety of ways.[9] Most common and most literal is to assume that discounting is part of a person's preferences. However, given the fundamental disposition toward revealed-preference theory and the assumption that whatever informed people do must be optimal for them, economists often take the view that, if such discounting is merely a matter of preference, then it is not to be questioned. We take a different view, one that we think is more in accord with the intuition that everybody—including most economists when we let our methodological guard down—views as a more sensible welfare criterion: that we should wish on ourselves, our children, our neighbors, and society the equal weighting of the expected hedonic well-being at different moments.

Most of us would find it morally repugnant if, controlling for uncertainties, personality, physical differences, etc., a parent admitted openly that he cares about the well-being of his five-year-old son ten times more than he cares about the well-being of his ten-year-old daughter. We would similarly be repulsed if he admitted that, again controlling for uncertainties, he cares ten times more about his five-year-old's current well-being than he cares about the same child's well-being five years from now, when he becomes a ten-year-old. For a parent to apply such differential weights

9. We ignore the "financial discounting" of money owing to interest earned, which has nothing to do with issues of hedonic discounting of the sort that we consider here.

to the hedonic well-being of his different children or to the hedonic well-being of one of his children at different ages is morally insane.

We believe that people should be just as repulsed if a fifteen-year-old says that he cares ten times more about his own current well-being than he cares about his own well-being five years from now, when he becomes a twenty-year-old. That is, we ought to be willing to make the same sort of normative judgments that we make about how a parent weights the hedonic well-being of his children about how an individual weights his hedonic well-being at different times in his life. Just as we must at times make reasoned judgments about whether a particular thirty-five-year-old is a fit guardian for a particular fifteen-year-old, so, too, we must at times make reasoned judgments about whether the fifteen-year-old is a fit guardian for her thirty-five-year-old self.

This argument for no hedonic discounting provides the basis for not including a discount function in our base model. But, as most economists (implicitly or explicitly) recognize, a discount function often serves as a useful reduced form to capture unmodeled uncertainties such as the probability of death or severe illness. As a simple illustration, suppose that the only uncertainty that matters is whether a person is alive or dead and that the person's utility is a function of consumption if living and a constant if dead. If the probability of dying between periods τ and $\tau + 1$ is q, then, from the perspective of period t, the probability of being alive in period τ is $(1 - q)^{\tau - t}$—that is, $p_\tau(\text{alive}) = (1 - q)^{\tau - t}$. Normalizing $u_\tau(\cdot, \text{dead}) = 0$, we can then usefully conceptualize the person's true well-being as being given by $W^t = \sum_{\tau=t}^{T} \delta^{\tau - t} u_\tau(\cdot, \text{alive})$, where $\delta \equiv 1 - q$.

In a variety of settings, therefore, a discount factor ought to be incorporated into the model as a sort of heuristic used by the people we are modeling, or as a heuristic being used by the modelers, to capture unmodeled contingencies. People should discount the future in the same sense that they discount a rumor of a coming appearance by (say) Johnny Depp—because they doubt whether it will happen. Our claim is that, from a normative perspective, such "heuristic discounting" is the only proper source of discounting.

Of course, when one looks at the choices that people make, substantial discounting beyond plausible uncertainties seems to be a fundamental behavioral reality—that is, people are excessively myopic relative to what would maximize their true well-being.[10] To model excessive myopia, we begin with the simple exponential-discounting formulation that is commonly used by economists. Suppose that a person makes choices that affect her well-being in periods $1, 2, \ldots, T$, and let u_τ denote her instanta-

10. For an overview of how to go about separating assessment of hedonic return to activities from revealed-preference theory, see Kahneman, Diener, and Schwartz (1999).

neous utility in period τ. Suppose further that the person's true overall well-being from the perspective of period t is given by

$$W^t \equiv \sum_{\tau=t}^{T} \hat{\delta}^{\tau-t} u_\tau .$$

In this formulation, $\hat{\delta}$ is the person's heuristic discount factor, capturing unmodeled uncertainties as discussed above.

As a simple form of inappropriate discounting, we suppose that, in period t, the person chooses her behavior to maximize period t intertemporal preferences

$$U^t \equiv \sum_{\tau=t}^{T} \delta^{\tau-t} u_\tau ,$$

where δ is the discount factor that she uses when making decisions. Simple excessive myopia says that the discount factor that the person uses when making decisions is smaller than her true heuristic discount factor—that is, $\delta < \hat{\delta}$. As a result, at each moment, the person gives too little weight to her future well-being.

Because the risky behaviors studied in this volume tend to generate positive short-term benefits and negative long-term consequences, simple excessive myopia makes people overly likely to engage in risky behaviors relative to the normative standard. Suppose, for instance, that a young person must decide whether to engage in sexual intercourse. In making this decision, she takes into account both the sexual pleasure that she would derive from the act and the expected costs that she might bear later in life. Formally, suppose that there are two periods, youth and adulthood, and that engaging in sexual intercourse in period 1 yields immediate pleasure $u_1 = 10$ but causes an expected future cost of 15, so $u_2 = -15$. Abstinence yields $u_1 = u_2 = 0$. If the person's true heuristic discount factor is $\hat{\delta} = 1$, the youth should choose abstinence, but if to make decisions the youth uses any discount factor $\delta \leq 2/3$, she would choose sex.

It is instructive to explore calibrationwise what magnitudes of discounting are consistent with heuristic discounting of the sort that we discuss above. Suppose, for instance, that a fifteen-year-old has a 50 percent chance of being alive and well at age thirty-five (which is obviously conservative). Then, under the simple formulation with $W^t = \Sigma_{\tau=t}^{T} \hat{\delta}^{\tau-t} u_\tau$, the person's yearly $\hat{\delta}$ should be something on the order of 0.966 (i.e., $0.966^{20} = 0.5$). Hence, even yearly discount factors of 0.95 should perhaps be considered excessive myopia.

The discussion presented above suggests that youths might engage in too much risky behavior because they attach too little weight to their well-being as adults. A related question arises: In terms of the discount factor that they use in making decisions, are youths more impatient than adults?

Indeed, researchers sometimes claim that youths discount the future at a higher rate than do adults and that this difference might explain certain differences in behavior between youths and adults (e.g., Becker and Murphy 1988).[11] We think that there is truth to this but that there are some subtleties involved of which researchers are not fully aware. In particular, people seem to have in mind using the exponential-discounting formulation combined with an assumption that δ gets larger as a person gets older—for example, a sixteen-year-old has preferences $\sum_{\tau=16}^{T} (0.8)^{\tau-16} u_\tau$, whereas a thirty-year-old has preferences $\sum_{\tau=30}^{T} (0.9)^{\tau-30} u_\tau$. But this formalization would imply that preferences are time inconsistent, and, moreover, the form of the time inconsistency seems intuitively wrong: it would imply, for instance, that people systematically plan to be indulgent in their distant future and then change their minds when the moment arrives. As we discuss below, people tend to exhibit exactly the opposite behavior.[12]

Realistically, youths are more impatient than adults. But it is probably best to model such differences by assuming date-specific per period discount factors. That is, for each k, there exists a discount factor δ_k between periods k and $k + 1$, and in period t the person chooses her behavior, to maximize the intertemporal preferences represented by

$$U^t \equiv \sum_{\tau=t}^{T} \left(\prod_{k=t}^{\tau-1} \delta_k \right) u_\tau .$$

Unfortunately, this formulation is not very easy to work with. Both empirical research and theoretical research become more difficult when there are multiple discount parameters to consider.

While the formulation presented above assumes time-consistent discounting, evidence is clear that people tend to have a time-inconsistent preference for immediate gratification. That is, when people make decisions that have both short-run and long-run consequences, they tend to satisfy their immediate wants in ways that they do not like from a long-run perspective.[13] Such preferences imply that people have self-control problems wherein they are unable on a moment-by-moment basis to behave in their own long-run best interests.

While the phenomenon is more general, there is a particularly simple

11. An alternative explanation is that youths "look" more impatient than adults because their perceived instantaneous utilities differ. We discuss this possibility in sec. 1.3 below.

12. Another subtlety goes hand in hand with the time inconsistency: Is the person aware of how her preferences will change? We discuss this issue below in the context of a preference for immediate gratification.

13. See, e.g., Ainslie (1975, 1991, 1992), Ainslie and Haslam (1992a, 1992b), Loewenstein and Prelec (1992), Thaler (1991), and Thaler and Loewenstein (1992). While the rubric *hyperbolic discounting* is often used to describe such preferences, the qualitative feature of the time inconsistency is more general (and more generally supported by empirical evidence) than the specific hyperbolic functional form.

model of preferences that captures the notion of a time-inconsistent taste for immediate gratification:[14]

$$U^t \equiv u_t + \beta \sum_{\tau=t+1}^{T} \delta^{\tau-t} u_\tau ,$$

where $\beta < 1$. This formulation is a simple modification of the standard model of exponential discounting, where the parameter β represents the time-inconsistent preference for immediate gratification—at any given moment, the person has an extra bias for the present over the future.

The assumption that people have a preference for immediate gratification accords with introspection, folk wisdom, and the psychological evidence. The most prevalent form of psychological evidence is evidence of declining discount rates. As an illustration, consider the findings in Green, Fry, and Myerson (1994). As part of their study (we shall return to other parts), twelve undergraduates at Washington University in St. Louis (averaging twenty years of age) were asked to make a series of choices between a delayed reward of $1,000 and an immediate reward ranging between $1 and $1,000. The length of the delay and the amount of the immediate reward were varied.[15] From this procedure, and from the relatively consistent behavior observed, discount functions can be inferred by comparing the "immediate equivalent" of $1,000 delayed by different durations.

Green, Fry, and Myerson do not report the raw data but visually present the median immediate equivalents. By our own visual inspection of this picture (their fig. 1), the immediate equivalent of $1,000 in one year was $625, the immediate equivalent of $1,000 in five years was $350, and the immediate equivalent of $1,000 in twenty-five years was $100. These numbers correspond to discount rates of 60 percent per year for year 1, but only 16 percent per year for years 2–5, and only 6 percent per year for years 6–25.[16] While the specific discount rates are not closely matched with other discount rates reported in the same study—for example, discounting was less severe for a $10,000 delayed reward—or in other studies, the general feature of declining discount rates is universal.[17]

14. These preferences were originally developed by Phelps and Pollak (1968) in the context of intergenerational altruism and later used by Laibson (1994) to model time inconsistency within an individual. This model has since been used by Laibson (1996, 1997), Laibson, Repetto, and Tobacman (1998), O'Donoghue and Rabin (1999a, 1999b, 1999c, 1999d, in press), Fischer (1997), and others.

15. Delays of one week, one month, six months, one year, three years, five years, ten years, and twenty-five years were used. Thirty different values between $1 and $1,000 were used for the immediate rewards.

16. For example, $\$625(1.6) = \$1{,}000$, and $\$350(1.6)(1.16)^4 = \$1{,}000$.

17. Extreme caution should be used in making too much of these results. In addition to having a small sample size, this study infers discount functions from trade-offs involving money amounts, which should not logically serve as proxies for utility discounting. Even so, there have been dozens of studies over the years that find that variants of hyperbolic discounting fit human and nonhuman choice better than exponential discounting (for additional

A time-inconsistent taste for immediate gratification implies that people might want to engage in some indulgent behavior at the present moment while at the same time they would prefer not to engage in the same indulgent behavior in the future. Hence, when people think about having sex, drinking alcohol, taking drugs, and so forth, their desire to do the activity now is greater than their current desire to do it in the future. This, in turn, implies that a person is more likely to engage in an indulgent activity at the moment of action than she would have preferred at some prior moment.

To illustrate, consider a modification of our earlier sex example. Suppose that having sex yields immediate benefits of 10 but has expected long-term costs of 15, and suppose that a person has the (β, δ) preferences described above with $\beta = 1/2$ and $\delta = 1$. Consider how a person feels in period t about having sex now, in period t, versus having sex in the future, in some period $t' > t$. Because having sex now yields utility $10 - (1/2)15 > 0$, the person would like to have sex now; and, because having sex in the future yields utility $(1/2)10 - (1/2)15 < 0$, the person in period t would prefer not to have sex in period t'. But, when period t' arrives, the person will then view having sex as yielding utility $10 - (1/2)15 > 0$ and will therefore then prefer to have sex.

The implications of having a preference for immediate gratification often depend on the person's beliefs about her own future behavior. Most research has focused on two extreme assumptions about beliefs. *Sophisticated* people are fully aware of their future self-control problems and therefore correctly predict how their future selves will behave, and *naive* people are fully unaware of their future self-control problems and therefore believe that their future selves will behave exactly as they currently would like them to behave.[18] But, clearly, this is a continuum; O'Donoghue and Rabin (1999d, in press) model a person who is *partially naive*—she is aware that she has future self-control problems, but she underestimates their magnitude.

A simple way in which to formalize these different beliefs is to suppose that a person is characterized not only by her true preference for immediate gratification, as reflected by β, but also by her beliefs as to what her future preference for immediate gratification will be, which we denote by $\hat{\beta}$. A sophisticated person, who knows exactly her future preference for immediate gratification, has perceptions $\hat{\beta} = \beta$. A naive person, who believes that she will not have a preference for immediate gratification in the

examples with human subjects, see Solnick et al. [1980], Myerson and Green [1995], and Kirby and Marakovic [1995]). To the best of our knowledge, there have been no studies in which behavior was found that fits exponential discounting better than hyperbolic discounting.

18. Strotz (1956) and Pollak (1968) carefully lay out these two assumptions (and develop the labels) but do not much consider the implications of assuming one versus the other. Fischer (1997) and Laibson (1994, 1996, 1997) assume sophisticated beliefs. O'Donoghue and Rabin (1999c) consider both and explicitly contrast the two.

future, has perceptions $\hat{\beta} = 1$. A partially naive person has perceptions $\hat{\beta} > \beta$ but $\hat{\beta} < 1$.

To illustrate the importance of beliefs, consider yet another version of our sex example. We again suppose that having sex yields immediate benefits of 10 but long-term costs of 15 and that a person has (β, δ) preferences with $\beta = 1/2$ and $\delta = 1$. But, instead of analyzing a person who is deciding whether to have sex now, we analyze a person who is planning to go on a date tonight (which we think of as the next period), during which she might have sex. Our discussion above implies that the person would prefer not to have sex tonight, and, hence, if she predicts that she would have sex, she may make some "commitment" so as to prevent herself from having sex—for example, she might cancel the date, ask a friend or parent to come along as a chaperone, etc. What is her prediction? Given beliefs $\hat{\beta}$, the person believes that this evening she will feel that having sex yields utility $10 - \hat{\beta}(15)$, and, thus, she predicts that she will have sex this evening if $\hat{\beta} < 2/3$. We can conclude that if $\hat{\beta} < 2/3$, she would avoid having unwanted sex, whereas if $\hat{\beta} > 2/3$, she would not make a commitment and end up having unwanted sex.

Beliefs about one's future preference for immediate gratification need not influence behavior. There is a class of situations in which beliefs do not affect behavior, namely, when decisions are disconnected in the sense that both the short-term and the long-term consequences of any specific decision are unrelated to those of any other decision. But, when one decision constitutes a commitment that changes later choice sets, or when the benefits or costs of decisions in different periods are tied together in some way, beliefs typically matter. This is likely the case for many of the behaviors examined in this volume. For instance, there may be decreasing returns to the number of times a person has sex in a week, or the expected future costs of having sex may be nonlinear in the number of times a person has sex. The very essence of addictive behaviors such as smoking cigarettes, drinking alcohol, and taking other drugs is that the utility from both consumption and nonconsumption depends on how much a person has consumed in the past.

To further illustrate the implications of sophistication versus naïveté, we return once again to our sex example. Suppose that a person has multiple opportunities to have sex and that the expected future cost of having sex is a nonlinear function of the number of sexual encounters. For any individual encounter, the perceived future cost of having sex on this particular occasion depends on both how often the person has had sex in the past and how often she expects to have sex in the future. Suppose, for simplicity, that the person has two opportunities to have sex, in periods 1 and 2, and then possibly experiences some cost in period 3. The benefits from having sex in periods 1 and 2 are V_1 and V_2, respectively. Let C_n be the expected period 3 cost if the person has sex n times, in which case the

expected future cost is nonlinear if $C_2 - C_1 \neq C_1$. Finally, we assume that $\delta = 1$ and $V_2 < \min\{C_1, C_2 - C_1\}$, which implies that, from a long-run perspective, the person should not have sex in period 2.

Consider the period 1 decision when $V_2 > \max\{\beta C_1, \beta(C_2 - C_1)\}$, which implies that in period 2 the person will have sex regardless of what she did in period 1. A sophisticated person in period 1 recognizes that she will have sex in period 2 and therefore views her decision in period 1 as having sex twice versus having sex once. Hence, she chooses to have sex in period 1 if $V_1 > \beta(C_2 - C_1)$. In contrast, a naive person in period 1 believes that she will not have sex in period 2 and therefore views her decision in period 1 as having sex once versus not having sex at all. Hence, she chooses to have sex in period 1 if $V_1 > \beta C_1$. Whether a naive person is more or less likely than a sophisticated person to indulge in period 1 depends on whether the costs are convex or concave. With convex costs, pessimism about the future makes a sophisticated person perceive the cost of current indulgence to be larger than does a naive person, and, hence, a sophisticated person is less likely to indulge. With concave costs, pessimism about the future makes a sophisticated person perceive the cost of current indulgence to be smaller than does a naive person, and, hence, a sophisticated person is more likely to indulge.

In fact, more matters than just the effects of pessimism. Let us reconsider this example when $\beta C_1 > V_2 > \beta(C_2 - C_1)$, which implies that in period 2 the person will have sex only if she had sex in period 1. The mindset of a naive person is the same as before: she expects not to have sex in period 2 no matter what she does now and therefore chooses to have sex in period 1 if $V_1 > \beta C_1$. A sophisticated person, in contrast, recognizes in period 1 that her decision is effectively between having sex twice versus not having sex at all and therefore chooses to have sex in period 1 if $V_1 > \beta[C_1 + (C_2 - C_1) - V_2]$. Inspection reveals that, for this case, a naive person is unambiguously more likely to choose to have sex in period 1.

The two cases in this example illustrate two effects of sophistication. On one hand, there is a *pessimism effect:* a sophisticated person is more pessimistic than a naive person about her future behavior, and this pessimism may affect the perceived future consequences of current indulgence. As the example illustrates, this effect can go in either direction. Second, there is an *incentive effect:* a sophisticated person might recognize that to avoid future indulgence, she must restrain herself now.[19]

Although, in theory, sophistication about future self-control problems can mitigate or exacerbate misbehavior, we suspect that for realistic environments, sophistication more often than not mitigates the harm caused

19. These two effects are first identified (and named) in O'Donoghue and Rabin (1999a, 1999b) for the realm of addiction. Those papers show that, for addictive goods, the pessimism effect leads a sophisticated person to indulge more than a naive person while the incentive effect leads a sophisticated person to indulge less.

by a preference for immediate gratification. This suspicion is somewhat impressionistic, but it is based on a conjecture that, in real-world environments, the incentive effect is likely quite important. Naive people are hurt not just because optimism about the future leads them to perceive small costs of current misbehavior, but also because they fail at self-management—they do things in the present that lead them to do more indulgent things in the future.

In what ways might self-control problems be particularly relevant for youths? There are two immediate points of comparison between youths and adults. First, do youths have a larger preference for immediate gratification than do adults—that is, do youths have a smaller β? Second, are youths less good at predicting their future preference for immediate gratification—that is, do youths have a larger $\hat{\beta}$? While the answers to these questions are both intuitively yes, there is limited evidence. We know of no experimental evidence on how adolescents compare to other age groups in terms of awareness of self-control problems, but the Green, Fry, and Myerson (1994) study cited above provides some suggestive evidence in terms of the magnitude of the preference for immediate gratification. In addition to gathering evidence on twenty-year-old college students, Green, Fry, and Myerson also report data on sixth graders (averaging twelve years old) from private religious schools and elderly subjects (averaging sixty-eight years old) from a subject pool maintained by the Washington University Psychology Department for the study of aging. Using the same methods that we used above to derive the discount rates for college students of 60, 16, and 6 percent per year for horizons of one year, one to five years, and five to twenty-five years, the comparable numbers are 111, 21, and 2 percent for sixth graders and 14, 8, and 8 percent for elderly subjects. We are highly skeptical of the external validity of the precise estimates, but they suggest that near-term discounting becomes less severe through adolescence (comparing the twenty-year-olds and the twelve-year-olds) and even less severe through adulthood.[20]

Of course, even if youths were no different from adults in terms of both their preference for immediate gratification and their awareness, the implications of these errors could still be quite different in youths and in adults. A preference for immediate gratification is relevant only to the extent that the person faces temptations, and youths and adults may differ on this dimension in a number of ways. First, there may be differences in inherent preferences—in the types of activities that youths and adults actually enjoy. Second, even for activities that yield equal intrinsic benefits for youths and adults, youths may perceive even larger immediate benefits from en-

20. Again, extreme caution should be used in making too much of these results. In addition to our earlier reservations, here the money amounts likely mean different things to the different age groups.

gaging in the activity owing to secondary considerations unique to adolescents, such as peer pressure, identity formation, or establishing autonomy. As a result, such activities become even more tempting for youths. Third, youths and adults face different opportunity sets that affect the types of tempting situations with which they are faced. Given the professional and personal constraints that adults face, many activities that might be tempting are so costly as not to pose a problem. Youths may be tempted to use illicit drugs, for instance, when the costs are too high for adults.

1.3 Mispredicting Future Utility

Section 1.2 above discusses errors in intertemporal trade-offs owing to excessive discounting. This section discusses a second way in which people make errors in weighing intertemporal trade-offs: they misperceive how they will feel about the future consequences of their actions.

The source of such errors is changing preferences, changes due to such factors as past behavior, temporary fluctuations in tastes, and changes in the environment. The notion that "states"—factors other than contemporaneous consumption—can affect preferences is not new to economics. For instance, many models over the years posit that people become accustomed to past consumption levels (see, e.g., Duesenberry 1952; Ryder and Heal 1973; and Bowman, Minehart, and Rabin 1999), Becker and Murphy (1988) build a model of addiction based on earlier research on habit formation (e.g., Pollak 1970, 1978) wherein the utility from consuming an addictive product depends on past consumption, and Laibson (in press) studies how changes in exogenous states, which he calls *cues,* can affect well-being.

Most models of state-dependent utility assume that a person can perfectly predict how changes in future states will affect her future preferences. For example, if a person must make summer vacation plans during the winter, it is assumed that she can predict how she will feel in the summer; and, if a person must decide whether to try crack cocaine for the first time, it is assumed that she can correctly predict how this consumption will influence her future enjoyment of activities, including consuming more crack cocaine. In a recent paper, Loewenstein, O'Donoghue, and Rabin (1999) formalize and explore the implications of a general bias in such predictions, which they label *projection bias:* people tend to underappreciate the effects of changes in their states and hence falsely project their current consumption preferences onto their future preferences. In this section, we review the evidence in support of such projection bias, present a simplified version of their model, and discuss its implications both in general and for youths in particular.

Research shows that people underappreciate short-term, transient changes in preferences, such as those induced by fluctuations in hunger or

the presence of environmental cues, and slowly developed but longer-lasting changes, such as those induced by addiction or changes in one's accustomed standard of living. Moreover, people underappreciate endogenous changes in preferences that depend on prior choices, such as drug addiction, and exogenous changes in preferences that do not depend on prior choices, such as those associated with aging. We discuss here a few representative studies.[21]

The prototypical example of incorrectly predicting short-term fluctuations in tastes is underappreciating the effects of hunger. Read and van Leeuwen (1998), for instance, asked office workers to choose between healthy snacks and unhealthy snacks that they would receive in one week, at a time when they should expect to be either hungry (late in the afternoon) or satiated (immediately after lunch). Subjects were approached to make the choice either when they were hungry (late in the afternoon) or when they were satiated (immediately after lunch). In general, people who expected to be hungry the next week were more likely to opt for unhealthy snacks than were those who expected to be satiated. But the key finding was that people who were hungry when they made the choice were also more likely to opt for unhealthy snacks than were those who were satiated, suggesting that people were projecting their current preferences onto their future selves.

Loewenstein, Nagin, and Paternoster (1997) provide evidence of projection bias with regard to sexual arousal. Male undergraduates were randomly assigned to view sexually arousing or nonarousing photographs. Subjects were then exposed to a vivid first-person date scenario in which "their date" suddenly requested a termination of physical intimacy, and asked to report their likelihood of behaving in a sexually aggressive fashion in this situation. Aroused subjects reported substantially higher likelihoods (70 percent) than did nonaroused subjects (50 percent), suggesting again that people's current preferences affect their predictions of future preferences.

The prototypical example of projection bias in predicting long-term changes in tastes is the underappreciation of adaptation. There is a plethora of evidence that adaptation is a central component of human well-being (see Helson [1964]; and for a recent review, see Frederick and Loewenstein [1999]). This literature consistently shows that people adapt to major changes in life circumstances. But there is also a great deal of evidence that people underestimate the extent to which they will adapt to new circumstances and hence overestimate the effect of major changes in circumstances on their long-run level of happiness. For instance, Loewenstein and Frederick (1997) compared the predictions by survey respon-

21. For a more extensive review of the evidence, see Loewenstein, O'Donoghue, and Rabin (1999).

dents of how changes in various environmental (e.g., a decline in sport fishing), social (e.g., increases in coffee shops), and personal (e.g., increases in body weight or income) factors would affect their well-being over the next decade to the reports of others about how actual changes in the last decade had affected their well-being. A clear pattern of underprediction of adaptation emerged in the data: people expected future changes to affect their well-being much more than others believed that matched changes in the past had affected their well-being.

To model such incorrect predictions, suppose that a person's true instantaneous utility in period t is given by $u(\boldsymbol{c}_t, \boldsymbol{s}_t)$. The vector $\boldsymbol{c}_t$ is the person's period t consumption vector, which includes all period t behavior relevant for current or future instantaneous utilities. The vector $\boldsymbol{s}_t$ is the person's "state" in period t. An individual state could be determined by past consumption (e.g., a person's addiction level) or by exogenous factors that might be internal (e.g., depression) or environmental (e.g., peer pressure). In addition, calendar time could be a state variable; we suspect that projection bias over states associated with aging may be quite important for youths.

Let $\tilde{u}(\boldsymbol{c}_\tau, \boldsymbol{s}_\tau|\boldsymbol{s}_t)$ denote the prediction of a person currently in state $\boldsymbol{s}_t$ of what her instantaneous utility would be from consuming $\boldsymbol{c}_\tau$ in state $\boldsymbol{s}_\tau$ in period $\tau > t$. If a person were fully rational, her prediction would be correct—that is, $\tilde{u}(\boldsymbol{c}_\tau, \boldsymbol{s}_\tau|\boldsymbol{s}_t) = u(\boldsymbol{c}_\tau, \boldsymbol{s}_\tau)$. But the evidence presented above suggests that people tend to exhibit *projection bias,* which, roughly speaking, means that predicted utility $\tilde{u}(\boldsymbol{c}_\tau, \boldsymbol{s}_\tau|\boldsymbol{s}_t)$ lies "in between" true utility $u(\boldsymbol{c}_\tau, \boldsymbol{s}_\tau)$ and utility in the current state $u(\boldsymbol{c}_\tau, \boldsymbol{s}_t)$. For the purposes of the discussion here, we shall consider a particularly simple formulation of projection bias:[22]

Definition. *Predicted utility exhibits* simple projection bias *if there exists* $\alpha \in [0, 1]$ *such that, for all* $\boldsymbol{c}$, $\boldsymbol{s}_\tau$, *and* $\boldsymbol{s}_t$, $\tilde{u}(\boldsymbol{c}_\tau, \boldsymbol{s}_\tau|\boldsymbol{s}_t) = (1 - \alpha)u(\boldsymbol{c}, \boldsymbol{s}_\tau) + \alpha u(\boldsymbol{c}, \boldsymbol{s}_t)$.

If $\alpha = 0$, the person predicts her future instantaneous utility correctly and therefore has no projection bias. If $\alpha > 0$, the person has projection bias, where the bigger is α, the stronger is the bias.

For any period t and initial state $\boldsymbol{s}_t$, a fully rational person would choose a path of consumption $(\boldsymbol{c}_t, \boldsymbol{c}_{t+1}, \ldots, \boldsymbol{c}_T)$ to maximize true intertemporal utility $U^t = \sum_{\tau=t}^{T} \delta^{\tau-t} u(\boldsymbol{c}_\tau, \boldsymbol{s}_\tau)$, taking into account how the consumption path affects the evolution of future states. A person with projection bias attempts to maximize her intertemporal utility: for any period t and initial state $\boldsymbol{s}_t$, a person with projection bias chooses a path of consumption $(\boldsymbol{c}_t, \boldsymbol{c}_{t+1}, \ldots, \boldsymbol{c}_T)$ to maximize her perceived intertemporal utility $\tilde{U}^t =$

22. This simple formulation incorporates two key features. First, the person understands the qualitative nature of changes in her preferences but underestimates the magnitude of these changes. Second, the more the person's future preferences differ from her current preferences, the further her prediction is from her true future utility. For a more general formulation that incorporates these two features, see Loewenstein, O'Donoghue, and Rabin (1999).

$\Sigma_{\tau=t}^{T} \delta^{\tau-t}\tilde{u}(\boldsymbol{c}_\tau, \boldsymbol{s}_\tau|\boldsymbol{s}_t)$, taking into account how the consumption path affects the evolution of future states. In other words, she behaves exactly as a fully rational person would except that she attempts to maximize $\tilde{U}^t \neq U^t$.

Projection bias can have important implications in a broad array of environments. There are three categories of errors to which projection bias can lead, which we illustrate with some simple examples along the lines of the behaviors discussed in this volume.

The first category of projection-bias errors involves choosing a suboptimal behavior owing simply to incorrect predictions of the future utility consequences of the behavior. To illustrate this type of error, suppose that a person is contemplating suicide owing to extreme depression. Suppose that the person is currently depressed and that the depression is sufficiently painful that, if it were to last a long time, it would be optimal for the person to commit suicide. Because projection bias can lead a person to underestimate the true utility of happy times, it can lead a depressed person to conclude incorrectly that suicide is optimal even when it is not.

We formalize this situation with a two-period model in which the person is "depressed" in period 1 and "happy" in period 2—that is, no matter what her behavior, the person's state will be D in period 1 and H in period 2. At the start of period 1, the person decides whether to commit suicide, where we normalize the utility from committing suicide to be 0. If the person chooses to live, then she receives utility $u(\text{life}, D) = -2$ in period 1 and utility $u(\text{life}, H) = 3$ in period 2. Assuming no discounting, the person's optimal choice is clearly to live because the eventual happy times are sufficiently happy to make it worth enduring the depression. But, with simple projection bias α, at the time she is deciding whether to commit suicide the person perceives her period 2 payoff from life to be $\tilde{u}(\text{life}, H \mid D) = (1 - \alpha)(3) + \alpha(-2)$. If her projection bias is big enough ($\alpha > 3/5$), then the person makes the incorrect choice to commit suicide.[23]

In the example given above, the person's state in each period is independent of her behavior; she chooses suboptimally only because her state at the time she makes the decision clouds her evaluation of the available options. Projection bias has more complicated and more damaging effects when the person's future state depends on her current behavior. In particular, if engaging in some activity causes future preferences to change in a deleterious way, then projection bias leads to overindulgence in that activity. For instance, if becoming addicted to cigarettes decreases a person's overall well-being, a person with projection bias will overindulge in cigarettes.

The first category of projection-bias errors is driven by incorrect predic-

23. A simple extension of this example shows the benefits of cooling-off periods for potential suicide victims. If the person delays until the depression at least partially subsides, she will perceive the utility of happy times to be closer to its true value and therefore will be less likely to commit suicide when she should not. This theme is discussed further in Loewenstein, O'Donoghue, and Rabin (1999).

tions of the future utility consequences of today's behavior. Incorrectly predicting future preferences can also cause a person to incorrectly predict future behavior, and—just as we saw for a preference for immediate gratification—incorrect predictions of future behavior can lead to bad decisions today. The second and third categories of projection-bias errors both revolve around how incorrect predictions of future behavior can lead to suboptimal decisions now.

The second category of projection-bias errors occurs when incorrect predictions of future behavior cause incorrect predictions about the future consequences of current choices. For instance, if projection bias leads a person to underestimate how often she will have sex in the future, then she may have an incorrect prediction about the future cost of having sex now. This logic is similar to that behind the pessimism effect in the realm of self-control problems and, just as for the pessimism effect, can lead to more or less current indulgence depending on the specifics of the environment. For instance, if a person who is not pregnant incorrectly predicts that, if she were pregnant, she would not get an abortion, then she would perceive the potential costs of having sex to be larger than they truly are and hence might be prone toward having too little sex (of course, this effect might be offset by other errors which we discuss that lead to having too much sex).

The third category of projection-bias errors involves *state mismanagement.* Because a person's future behavior depends on her future state, avoiding certain behaviors may require avoiding certain states.[24] But, since projection bias can lead a person not to recognize how a certain state would influence behavior, the person might end up engaging in unanticipated indulgent behavior. This failure to avoid situations in which indulgence is irresistible is similar to the lack of the incentive effect in people who are naive about self-control, and both errors tend to cause overindulgence.

To illustrate this type of error, suppose that, in the early evening, a person who is not sexually aroused must decide whether to go to a bar. If she decides not to go to the bar, then she remains at home and sexually unaroused for the entire night. If, instead, she goes to the bar, she meets someone, becomes sexually aroused, and then chooses whether to have sex. Suppose that optimal state management involves not going to the bar because she would choose to have sex if she did go to the bar, whereas, from an ex ante perspective, she would prefer to stay at home all night than go out and have sex.

To formalize this situation, consider a model with two periods, early

24. Although his model assumes exogenous states, Laibson (in press) discusses how in a more general model it is important to manage cues so as to avoid some undesirable behaviors. He refers to this phenomenon as *cue management,* which motivates our term *state management.*

evening and later that night. In period 1, the person decides either to go to the bar or stay home, and her state is sexually unaroused, which we denote by N. If she decides to stay home in period 1, then she remains at home and sexually unaroused for period 2. If she decides to go to the bar, in contrast, then her period 2 state is sexually aroused, which we denote by A, and in this state she must make the decision whether to have sex or to go home. We assume the instantaneous utilities to be as follows:

In Unaroused State	In Aroused State
$u(\text{home}, N) = 0$	$u(\text{home}, A) = -2$
$u(\text{bar}, N) = 1$	$u(\text{sex}, A) = -1.5$
$u(\text{sex}, N) = -3$	

The person is better off in period 1 if she goes to the bar rather than staying home. But going to the bar can lead to undesirable period 2 behavior, and, indeed, for this example, going to the bar will lead to having sex. Optimal behavior is to stay home for the entire night because $u(\text{bar}, N) + u(\text{sex}, A) < u(\text{home}, N) + u(\text{home}, N)$. Now suppose that the person has simple projection bias α. First note that, no matter what is α, the person would prefer to stay home for the entire evening if she thought she would have sex in period 2. But, since she perceives $\tilde{u}(\text{sex}, A|N) = (1 - \alpha)(-1.5) + \alpha(-3)$ and $\tilde{u}(\text{home}, A|N) = (1 - \alpha)(-2) + \alpha(0)$, for $\alpha > 1/7$ the person perceives that she would choose not to have sex in the event that she goes to the bar, and, moreover, for $\alpha > 1/2$ she would choose to go to the bar.

Such state mismanagement owing to projection bias can arise in a variety of other domains. For instance, if having one beer creates an increased desire for additional beers, then optimal state management may require avoiding bars altogether, whereas projection bias might lead the person to go to a bar expecting to have only one beer and end up drinking too many. Similarly, if being in a smoke-filled room provokes a strong desire to smoke, then optimal state management may require avoiding smoke-filled rooms, whereas projection bias might undermine this decision.

This third category of projection-bias error includes errors that are not state mismanagement per se; it might also involve a person failing properly to prepare for some behavior because she did not expect to engage in that behavior. To return to the sex example, for instance, if the person goes to the bar expecting not to have sex, she might not bother to bring a condom. As a result, even if optimal behavior were to have sex with a condom, the person might have sex without a condom because she planned not to have sex.

How might projection bias be particularly relevant for youths? The first question to ask is whether youths are more susceptible to projection bias than are adults. While the answer is almost certainly yes, we know of no

good evidence on this issue. Suggestive evidence is that people are clearly aware of projection-bias problems and develop rules to help overcome them over the course of their lives—for example, the common wisdom "never shop on an empty stomach." At the same time, the evidence cited earlier and in Loewenstein, O'Donoghue, and Rabin (1999) makes it clear that adults do not fully appreciate changes in their preferences even on dimensions such as hunger where they have accumulated a great deal of experience. Hence, much as for self-control problems, our impression is that the major differences between adults and adolescents in terms of projection bias are driven not so much by different degrees of the bias, but by how a given degree of bias operates on the different preferences and situations that confront youths and adults.

As discussed in section 1.1 above, youths differ from adults in their concern for identity formation, establishing autonomy, and maintaining the regard of peers. The natural way to incorporate such concerns into a decision-making framework is in the utility function; if a person suddenly feels pressure from his peers to smoke marijuana, then his perceived marginal utility of smoking marijuana increases. Similarly, if a young person is driving and suddenly feels a drive to confirm his masculine identity, then his perceived marginal utility from driving fast increases.

Because such forces influence utilities, they are subject to projection bias. For instance, while youths likely recognize that their friends will influence their desire to engage in various activities, they likely underestimate the magnitude of these forces when they are not with their friends and, as a result, end up engaging in unwanted behaviors. State mismanagement becomes an important problem. A young person may go out with his friends expecting to resist peer pressure but then fail to do so. He may, for instance, go to a bar expecting to have one beer but then be pressured by his friends to drink many more. Giving in to peer pressure or drives to confirm identity and autonomy may also, in turn, affect the management of other states. For instance, a college student might choose to give in to peer pressure and drink alcohol every weekend, planning to quit when the peer pressure subsides, but not realizing how much more she will crave alcohol after she has become addicted.

Projection bias over the states associated with aging may be quite important for youths. To evaluate the long-term consequences of many risky behaviors, adolescents must predict how they will feel as adults. But youths and adults clearly have different preferences, and projection bias predicts that youths will underestimate how much their preferences will change as they age.

To illustrate the importance of this youth-to-adult projection bias, consider a young man who is debating whether to drop out of school, a decision that will affect whether he has a good job or a bad job as an adult.

Suppose that when he becomes an adult, he will care a lot about having a good job, but as a youth, he does not care at all. If the young man has projection bias with respect to differences between his youthful and his adult preferences, he will underestimate the long-run costs of dropping out of school and will therefore be too likely to drop out of school. An analogous conclusion would hold for any risky behavior, such as taking drugs, that might influence his future job prospects.

This youth-to-adult projection bias is perhaps even more important in the light of changing constraints. For instance, while youths have significant free time, adults must work five days a week. If youths are aware of these changing constraints, and if, because of projection bias, they predict that their utility function will not change very much, then they may predict that their actual utilities will change a lot. If youths currently think that working five days a week would be horrible and project this preference onto their adult selves, they might think that adulthood is going to be one, long miserable existence and may therefore care very little about imposing additional costs on their adult selves.

An implicit theme throughout the discussion presented above is that youths may exhibit behaviors that appear to be extremely myopic when in fact they are not. Consider a person who gives in to peer pressure or a drive to confirm his masculine identity. At first glance, we might be inclined to interpret this behavior as a sudden increase in myopia—to assume that peer pressure made the person neglect the future. We feel that this interpretation is incorrect. An alternative, projection-bias interpretation is that, when the pressure to conform is aroused, projection bias causes the person to exaggerate the persistence of this pressure. Hence, while he acts as if he cares only about his current well-being, he thinks that pursuing his immediate well-being is also what he must do for his long-run well-being.[25]

1.4 Repeated Risky Choices

Sections 1.2 and 1.3 above abstract away from the probabilistic nature of the long-term consequences of risky behavior. We now focus on riskiness per se, examining situations in which a person repeatedly chooses while young whether to engage in some behavior that might cause a future

25. Some actions in pursuit of immediate gratification may well be usefully thought of not in terms of either extreme discounting or projection bias but rather in terms of something more akin to the "visceral" model of choice discussed in Loewenstein (1996). If, e.g., a teenage boy makes Meatlovian promises with lifelong consequences in the pursuit of immediate sexual gratification, this is perhaps not to be thought of in terms of either an active belief that his current state of sexual frustration will last a lifetime or a conscious decision that his near-term gratification is more important to him than an entire lifetime of consequences; rather, he may simply not be attending to these future consequences at all.

bad outcome and does not learn while young whether his behavior thus far has assured the bad outcome. Many risky behaviors match this abstract description to some degree. Even when a young person is aware that smoking can lead to lung cancer or emphysema, he must repeatedly choose whether to smoke without knowing how much future harm he has already caused himself. Or a young person might face decisions whether to have sex on multiple occasions before knowing whether past sexual experiences have led to pregnancy or AIDS.

We begin our analysis within the rational-choice framework. While the rational-choice analysis is of interest in its own right, we also use it as a template to study the implications of the errors discussed in sections 1.2 and 1.3 above in the context of repeated risky choices, and to study the role that incorrect perceptions of risk can play in risky behavior.

Suppose that a fully rational person chooses up front the number of times, n, in which to engage in an activity. Let $V(n)$ be the total pleasure received from engaging in the activity n times, where V is increasing and concave. We interpret concavity as a proxy for the fact that enjoyment of the activity varies over time and the person indulges only at the n most enjoyable times. But we emphasize that, while concavity is convenient, it is not necessarily a good assumption here.[26]

Engaging in the activity carries with it the probabilistic cost of some bad outcome that might occur in the future. We assume that the bad outcome either occurs or does not occur, and if it occurs, then the person incurs a cost $\Gamma > 0$. In other words, the risk of harm accumulates, but the extent of harm does not. This feature is central to our results. While unrealistic in its extreme, it captures well the qualitative nature of many of the risky activities examined in this volume.

Formally, we assume that each time the person engages in the activity, there is an independent probability p that doing so causes the bad outcome to occur (had it not already been caused). Hence, if the person engages in the activity once, then the bad outcome occurs with probability p; if the person engages in the activity twice, then the bad outcome occurs with probability $p + (1 - p)p$; and, if the person engages in the activity n times, then the bad outcome occurs with probability $\pi(n; p) \equiv 1 - (1 - p)^n$. The expected cost of engaging in the activity is therefore

$$C(n; p) \equiv \pi(n; p)\Gamma \equiv [1 - (1 - p)^n]\Gamma.$$

Assuming for simplicity that there is no discounting, and assuming for analytic ease that n is a continuous variable, the person chooses n to maximize his net payoff $V(n) - C(n; p)$. Let $n^*(p, \Gamma)$ denote the person's opti-

26. In particular, convexity might be a better assumption for the consumption of addictive products; consuming moderate amounts of the product can be a horrible mix that yields many moments of the pain of withdrawal.

mal choice as a function of his perception of the risk p and the severity of the bad outcome Γ.

The perceived severity of the bad outcome has a simple and straightforward effect on the person's behavior: the more costly is the bad outcome—the larger is Γ—the less the person will engage in the activity. For instance, the worse a person perceives pregnancy or acquiring HIV to be, the less unprotected sex the person will have.

Owing to some interesting features of the cost function $C(n; p)$, however, the perceived riskiness p has considerably more complicated effects on the person's behavior. Because $C(n; p)$ is concave in n, $V(n) - C(n; p)$ can be convex, which gives rise to possible "corner solutions." For some parameter values, it can be that the person wants to engage in either none of the activity or a lot of the activity, never in between. But the more notable feature of $C(n; p)$ is that increasing the riskiness p can decrease the marginal cost of engaging in the activity—that is, can decrease $\partial C(n; p)/\partial n$. Formally, define $\pi_{np} \equiv \partial^2\pi/(\partial n \partial p)$, which is the cross-partial of π. It is straightforward to derive $\pi_{np} = (1 - p)^{n-1}[1 + \ln(1 - p)]$ so that $\pi_{np} = 0$ for $n = 1/[-\ln(1 - p)] \equiv \hat{n}(p)$. Moreover, $\pi_{np} > 0$ for $n < \hat{n}(p)$, in which case increasing p increases the marginal cost of engaging in the activity, and $\pi_{np} < 0$ for $n > \hat{n}(p)$, in which case increasing p decreases the marginal cost of engaging in the activity. Figure 1.1 illustrates how the total cost and marginal cost depend on n and p.

As indicated by panel A of figure 1.1, the larger is the perceived riskiness, the higher is the expected total cost for any given n. Hence, there is one straightforward prediction: if, for a given p, the person would not engage in the activity at all, then, for any larger riskiness, the person also would not engage in the activity at all. If a teenager is refraining entirely from unprotected sex or drug use and then comes to believe that the bad consequences of that activity are even more likely than she had earlier thought, she clearly will not start engaging in the activity.

But, if, for a given p, the person would engage in at least some of the

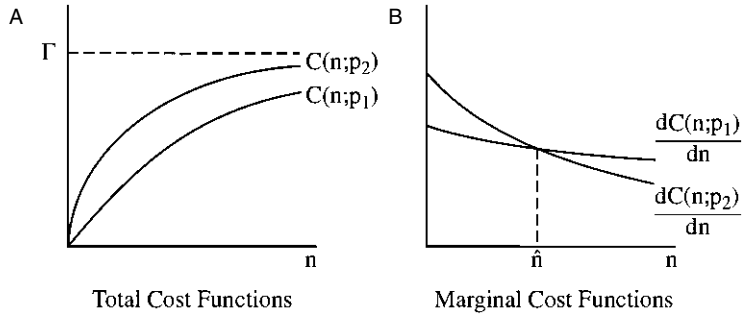

Fig. 1.1 Cost functions for $p_2 > p_1$

activity, an increase in the riskiness might either decrease or increase her level of that activity. To illustrate, consider the following special functional form for the benefits $V(n)$:

$$V(n) = \begin{cases} rn & \text{if } n \leq n_L, \\ rn_L + s(n - n_L) & \text{if } n \in [n_L, n_H], \\ rn_L + s(n_H - n_L) & \text{if } n \geq n_H, \end{cases}$$

where $r > s$. This function implies that the marginal utility of engaging in the activity is given by

$$V'(n) = \begin{cases} r & \text{if } n \leq n_L, \\ s & \text{if } n \in [n_L, n_H], \\ 0 & \text{if } n \geq n_H. \end{cases}$$

This contrived functional form is useful because, given the concavity of costs, the person chooses $n \in \{0, n_L, n_H\}$. In other words, the person chooses either abstinence, low indulgence n_L, or high indulgence n_H. Because we are interested in whether an increase in riskiness leads to more or less indulgence, we consider the case where initially the person chooses low indulgence, or $n^*(p, \Gamma) = n_L$.

If $n_H < \hat{n}(p)$, then an increase in riskiness raises the marginal cost of changing from n_L to n_H and therefore will not lead to increased indulgence. If, however, $\hat{n}(p) < n_L$, then an increase in riskiness lowers the marginal cost of changing from n_L to n_H, and therefore the person may now prefer high indulgence to low indulgence. Of course, the increase in riskiness also raises the total cost of engaging in the activity at levels n_L and n_H, so now abstinence may be optimal. But, if r is sufficiently high, the person will not respond to greater risk by switching to abstinence, and, if s is sufficiently high—for instance, if initially the person was just indifferent between engaging in the activity n_L versus n_H times—then the increase in riskiness will lead the person to switch from low indulgence to high indulgence.

This example illustrates the more general point that there are two possible reactions that a person might have to an increase in perceived riskiness. First, there is the intuitive reaction wherein the person reduces his indulgence so as to avoid the bad outcome whose likelihood has increased. But, second, there is a fatalistic reaction: the person instead might decide that, because he is not willing to choose very low indulgence, the bad outcome is now essentially unavoidable, and therefore he might as well increase indulgence. If a young person is sufficiently committed to some level of sexual activity, then, if he comes to believe that the risk is greater, he may increase his activity. If an adolescent perceives that, once she uses drugs, she will surely become an addict, then she is likely to use drugs a great deal if she uses them at all. More generally, for any activity in which

a person will engage to some degree no matter what, an increase in the perceived riskiness can potentially lead to increased indulgence.

This fatalistic reaction relies on two key assumptions. First, it must be that, given the amount in which the person might plausibly want to indulge, the probability of harm done is not negligible. Second, as we assume throughout, the eventuality being risked must be all or nothing rather than cumulative—that is, if the bad thing happens once, then it either cannot happen again or will not cause much further harm if it does occur.

We do not have a strong empirical sense of whether, for realistic benefits and risk levels, the fatalistic reaction identified above is likely to be important within the rational-choice setting. As we shall discuss shortly, however, such fatalism is more likely to be important when a person suffers from an overly strong taste for immediate gratification or from projection bias. Moreover, even within the rational-choice framework, a similar logic applies on a more realistic level when adolescents choose among alternative activities, where the more pleasurable activities are also riskier. To illustrate, suppose that an adolescent can engage in activities 1 and 2, at levels n_1 and n_2, and suppose that the benefits are

$$V(n_1, n_2) = \begin{cases} r \min\{n_1, \bar{n}\} & \text{if } n_1 > n_2, \\ s \min\{n_2, \bar{n}\} & \text{if } n_1 \leq n_2. \end{cases}$$

This contrived functional form implies that the activities are substitutes for each other. In addition, again because of concave costs, the person will do one of three things: engage in neither activity, engage in level $\bar{n}$ of activity 1, or engage in level $\bar{n}$ of activity 2. Suppose that the risks associated with the two activities are p_1 and p_2 and that the expected cost to the person is given by $C(n_1, n_2; p_1, p_2) = [1 - (1 - p_1)^{n_1}(1 - p_2)^{n_2}]\Gamma$. Hence, these activities are substitutes not only in terms of the benefits that they supply, but also in terms of the risks that they carry. Finally, assume that $s > r$ but that $p_2 > p_1$ so that activity 2 is both more enjoyable and riskier.

Suppose that, given some initial risks (p_1, p_2), the person chooses to engage in level $\bar{n}$ of activity 1, and consider what happens when p_1 increases. If initially the person strictly prefers no activity to level $\bar{n}$ of activity 2, then the only possible change following an increase in the riskiness of activity 1 is to no activity. But, if initially the person prefers level $\bar{n}$ of activity 2 to no activity, then the only possible change following an increase in the riskiness of activity 1 is to level $\bar{n}$ of activity 2.

This example illustrates the more general point that an increase in the perceived riskiness of less risky behaviors can lead people not to abstain, but rather to engage in riskier substitute behaviors. For instance, if young gay men suddenly perceive oral sex to carry a larger risk of HIV infection than they had thought, they may start engaging in riskier activities, such as anal sex. Similarly, if young heterosexuals suddenly learn that even con-

doms do not protect fully against the risk of pregnancy, they may start having unprotected sex.

While the rational-choice analysis of repeated risky choices yields some important implications, additional insights are yielded by relaxing the assumptions that people rationally perceive the probability consequences of their actions and that people react optimally to such perceptions. We now discuss these additional insights in general intuitive terms rather than with a formal analysis.

Consider first the case where people rationally foresee the probability consequences of their actions but, along the lines discussed in sections 1.2 and 1.3 above, do not react optimally to these beliefs. When people make repeated risky choices, overindulgence due to self-control problems and projection bias becomes more pronounced. More precisely, when self-control problems and projection bias cause overindulgence, the concavity of $C(n; p)$ exacerbates this overindulgence because the person perceives further indulgence as less costly.

This basic idea can play itself out in a couple of ways. Sections 1.2 and 1.3 emphasize how naïveté about self-control problems and projection bias can lead to failures of self-management: people fail to avoid situations in which they will not be able to resist indulgence. Because each choice to indulge in a series of risky choices reduces the costs of future indulgences, each failure at self-management is potentially more harmful than if the choice were made in isolation. For instance, while initially there might be only a few situations in which adolescents cannot resist having sex, once they have succumbed in such episodes, they will realize that having further sex is less costly and will therefore succumb in even more situations. Hence, naive self-control problems and projection bias can lead to an eventual fatalism and therefore high levels of indulgence.

Sophistication about self-control problems overcomes such descents into fatalistic consumption due to failures at self-management. But sophistication can lead to overconsumption due to an up-front feeling of fatalism. A person who is worried about future self-control problems may expect not to behave optimally in the future. Hence, even when the person does not want to indulge moderately, he may pessimistically (but correctly) predict that there will be a number of future occasions on which he will not be able to control himself. Given the concavity of $C(n; p)$, such pessimism reduces the perceived cost of indulgence at any given moment and hence makes indulgence more likely.[27]

We next consider the possibility that people have irrational beliefs about the probability consequences of their actions. Suppose that, given some level of an activity n, the true likelihood of the bad outcome is $\pi(n; p)$ but

27. As in our analysis in sec. 1.2 above, whether awareness of self-control problems helps or hurts a person here depends on the specifics of the environment.

that a young person perceives the likelihood to be $\hat{\pi}(n) \neq \pi(n; p)$. We initially focus on errors induced by the incorrect perception of risk per se and therefore assume that the person does not depart in any other way from full rationality.

The simplest form of incorrect risk perceptions is that adolescents may incorrectly perceive the riskiness, p, of each episode of the activity. If they believe that each indulgence carries with it probability $\hat{p} \neq p$ of incurring the bad outcome, then their perceived risk function will be $\hat{\pi}(n) = \pi(n; \hat{p})$. For instance, the folk wisdom that adolescents have a false sense of invulnerability might be conceptualized as perceiving $\hat{p} < p$.[28] For an otherwise fully rational person, this error would cause the same effects on behavior as a decrease in the true riskiness. Hence, following the logic outlined above, while feelings of invulnerability might have the intuitive effect of causing the person to engage in more of the activity, it can in fact decrease how often he engages in the activity if he becomes (falsely) convinced that he has not doomed himself by engaging in moderate amounts of the activity.

Unfortunately, there is evidence that people, especially youths, have a more extreme form of irrational belief: quite apart from incorrectly perceiving the true risk per episode of an activity, adolescents think that the probability function is more concave in n than it actually is—in our notation, $\hat{\pi}(n)$ is more concave than $\pi(n; p)$. For instance, Linville, Fischer, and Fischhoff (1993) find that college students wildly exaggerate the chance of acquiring HIV from one sexual encounter but then think that the chance of acquiring it from ten sexual encounters is not much higher. In reality, for the levels of heterosexual activity that are most common, the probability of contracting HIV is approximately proportional to the number of sexual encounters. This finding that people form probabilistic beliefs that are too insensitive to the number of times in which they engage in an activity accords well with the more general psychological phenomenon of diminishing sensitivity and what Kahneman (1994) has called *extension neglect.* In a variety of domains, people are much less sensitive to the magnitude of variables than they rationally ought to be. For instance, in deciding ahead of time how burdensome having to walk to a hotel with suitcases will be, people are more sensitive to having to walk three blocks versus two blocks than they are to having to walk twenty blocks versus nineteen blocks. This heuristic that the difference between nineteen and twenty is smaller than the difference between two and three may be correct in most environments, but in some environments—such as walking with suitcases—it is quite inappropriate. In the context of risky choice, diminishing

28. We remind the reader that, while such feelings of invulnerability are commonly associated with youths, there is very little evidence supporting this conventional wisdom, and there are some ways in which youths seem to overestimate the risks.

sensitivity, taken literally and to its extreme, suggests that people are likely vastly to underestimate the risk of all but the first time they engage in some activity.

This error can be expected to have the same general effect that overestimating p has. Because people overestimate the perceived risk from engaging in the activity a few times, it may induce abstinence. But, if people cannot resist indulging to some degree, an undersensitivity to the number of times that they indulge may lead them to indulge a great deal. Taking the results in Linville, Fischer, and Fischhoff (1993) literally, for instance, and assuming that acquiring HIV is the only risk from sex, then we would expect their subjects to have unsafe sex quite a lot or not at all.

Because the discussion presented above of irrational beliefs assumes that people react optimally to these beliefs, our conclusions about behavior represent departures from what would be optimal given the true risk $\pi(n; p)$. In other words, if we let $W[\hat{\pi}(n)|\pi(n; p)]$ be the true well-being of a person who optimizes with respect to beliefs $\hat{\pi}(n)$ when the true risk is $\pi(n; p)$, then $W[\hat{\pi}(n)|\pi(n; p)] \leq W[\pi(n; p)|\pi(n; p)]$. If, however, a person suffers from other psychological errors, false risk beliefs can interact with these other errors. While the behavioral implications of the false beliefs discussed above do not change, these effects can exacerbate or counteract other errors that people make. Indeed, a common theme in psychological research is that irrational beliefs can help rather than hurt a person in many situations. Along these lines, we believe that it is important to understand how false beliefs interact with self-control problems and projection bias.

Because self-control problems and projection bias generally lead to overindulgence, false beliefs help when they lead to less indulgence and hurt when they lead to more indulgence. Whether falsely believing that the risk function is more concave than it really is (because people exaggerate the riskiness of the activity or suffer from extension neglect) helps or hurts depends crucially on the degree to which the person is enticed by the activity. Exaggerated beliefs about risk may help if they make it so that the person is able to resist even the strongest temptations. If, for instance, the perceived risk of a single sexual encounter is sufficiently exaggerated, it may help a person with self-control problems refrain altogether. But, if there are likely to be a number of occasions on which the person cannot resist, then false beliefs in the direction of more concavity can exacerbate overindulgence.

1.5 General Discussion

In this section, we attempt to tie together our analysis by briefly discussing some empirical and policy implications.

Our focus in this chapter has been the potential for applying formal

behavioral-economic models to theoretical and empirical research on youthful behavior. While there is a broad range of issues from both psychology and economics that are relevant to the behavior of adolescents, we discuss only a few specific topics. Nonetheless, we believe that the errors that we discuss may go a long way toward understanding excessive risk taking by adolescents.

Behavioral-economic models abandon those features of the classical economic model that psychological evidence indicates are wrong, such as the assumption that discounting is time consistent. But behavioral-economic models incorporate the features of the rational-choice model that are realistic, such as the assumption that people generally pursue satisfying rather than unsatisfying activities. Indeed, almost all the qualitative empirical results that are lauded as support for the rational-choice model—such as prices affecting consumption of addictive goods—are consistent with the behavioral models that we discuss and with all other behavioral models with which we are familiar.[29]

Although in many instances the rational-choice model and behavioral models both make correct *qualitative* predictions, in those same instances behavioral models make sounder *quantitative* predictions and can better explain observed behaviors with reasonable parameter values. For example, while the rational-choice model can explain certain observed patterns of consumption of addictive products, this explanation often seems to require absurd levels of impatience. Because even small departures from the rational-choice model can lead to significant quantitative changes in behavior, the same observed patterns of consumption that require absurd levels of impatience when viewed through the lens of rational choice can be explained, for instance, by reasonable levels of impatience and a small self-control problem.

There is, however, an important qualitative implication of our models that is different from the rational-choice model: systematic incorrect predictions of future behavior. The rational-choice model permits incorrect predictions due to uncertainty but rules out any systematic bias in these incorrect predictions. By contrast, the behavioral models that we discuss above suggest systematic incorrect predictions. Naïveté about self-control problems yields a systematic underestimation of future indulgence. Projec-

29. Research proposing behavioral alternatives to the classical economic model is often designed to persuade the audience of the truth and usefulness of the behavioral alternatives and therefore typically emphasizes how the behavioral alternatives make distinct comparative-static predictions from the classical economic model. Since such persuasion is one of our goals, we, too, discuss comparisons. But we also emphasize that there is overwhelming support for the assumption that people have a time-inconsistent preference for immediate gratification and strong support for the assumption that people suffer from projection bias. Hence, we do not limit ourselves solely to comparisons of comparative-static predictions but instead focus more on the direct goal of understanding the implications of these true behavioral assumptions.

tion bias also yields systematic incorrect predictions, although the direction is more environment specific. Consider, for example, projection bias over the addictiveness of cigarettes. An unaddicted person projects her current low craving onto her future preferences and therefore tends to underestimate future consumption. An addicted person, on the other hand, projects her current high craving onto her future preferences, which moves her toward overestimation of future consumption.

Evidence of systematic incorrect predictions therefore provides some support for these models. A direct approach to investigating such incorrect predictions is to ask people to predict future behavior and then later compare predictions to actual behavior. This approach is commonly used in psychological research, and, in fact, systematic incorrect predictions are often found; much of the evidence for projection bias was exactly such evidence.

For economists unwilling to infer anything from self-reported predictions, however, a second approach is to look for situations in which a person's behavior provides information about her predictions of future behavior. As a contrived example, consider how a fifteen-year-old might react differently to two different changes in the pricing scheme for cigarettes:

PLAN A. *The price of cigarettes is permanently increased by fifty cents per pack.*

PLAN B. *There is no change in the price of cigarettes, but to buy cigarettes, a person must purchase an access card for an up-front fee of $500.*

If the fifteen-year-old were asked which of these plans she would prefer, her answer would reveal her beliefs about how much she expects to smoke. In particular, since $500 is less than three years of fifty cents per day, preferring option A suggests that the fifteen-year-old does not plan to be smoking for more than three years, and preferring option B suggests that she does not plan to be smoking for fewer than three years.

Suppose that there are two types of states, those that implement plan A and those that implement plan B, and consider how the behavior of fifteen-year-olds would differ across the two types of states. In particular, consider what the rational-choice model has to say about the states in which we would expect to see more fifteen-year-olds become long-term smokers. Rational actors facing little uncertainty plan to be either nonsmokers, short-term smokers, or long-term smokers and then stick to their plans. Because the price of long-term smoking is lower and the price of short-term smoking is higher in plan B states, the rational-choice model predicts that more fifteen-year-olds become long-term smokers in plan B states.

In contrast, if people underestimate future consumption owing either to naïveté about self-control problems or to projection bias, then we could expect to see fewer fifteen-year-olds become long-term smokers in plan B

states. In addition to people who carry out their plans to be nonsmokers, short-term smokers, or long-term smokers, our models suggest that there will also be people who plan to be short-term smokers but end up being long-term smokers. Although, just as in the rational-choice model, there is a tendency for more people to plan long-term smoking in plan B states, the higher cost of short-term smoking in plan B states deters people from planning short-term smoking and ending up as long-term smokers. This latter effect could dominate when underestimation is large.

More realistic empirical tests comparing our models to rational-choice models revolve around how people react both to changes in short-term costs and benefits and to changes in long-term costs and benefits. For instance, suppose that there is an improvement in detox programs that reduces the long-term cost of becoming addicted. According to the rational-choice model, such a change should lead to many people becoming addicted because many prospective addicts should now be enticed by the lower cost of addiction. When people underestimate their likelihood of becoming addicted, however, such a change would have less effect because the marginal consumer who ends up in the detox program is unlikely to have been planning to get addicted.

We conclude with some brief and speculative discussion of the policy implications of our analysis. While the simplified nature of our analysis prevents its confident use for specific policy prescriptions, some general themes emerge for each of the two approaches that society often considers for combatting excessive risk taking by adolescents, regulation and education.

The main theme of our analysis is that youths do not react optimally to the intrinsic costs and benefits that they face. Because so many of the risky behaviors studied in this volume involve short-term benefits and long-term costs, providing short-term rewards for good behavior may be quite important. Policy can be designed to manipulate incentives to better align their perceived immediate incentives with their actual long-term goals.

Our analysis also suggests that one of the major problems that youths have is that they fail to recognize when current indulgence will lead to future indulgence. Hence, policy makers should perhaps create incentives that encourage youths to take into account the linkages between decisions. By making youths pay a large sum up front for the right to begin smoking, for instance, plan B discussed above may provide an alternative tax-incentive scheme that will force adolescent smokers to decide whether they want to be smokers when they begin their habit rather than paying incrementally as in plan A. While there are, of course, some practical limits to this approach, an appealing feature of such incentives is that they would presumably do relatively little harm to those youths who are rationally taking up the career of smoking—who correctly realize that $500 up front is worth the lifetime of pleasure—while preventing smoking among those who might be developing the smoking habit unintentionally.

The second category of efforts to reduce risky behavior is education. If adolescents were fully rational and caused no harm to others by their risky behavior, then the optimal education policy would be to provide them with as much accurate information as possible. If youths make errors, in contrast, there may be a role for "propaganda" aimed at misleading adolescents about the true risks involved in their activities. There are many reasons that society and government might wish only to tell the truth, from moral disposition to worries about loss of credibility. But, even under the constraints of only telling the truth, there is usually scope for what to emphasize. Our analysis can provide some insights into the types of truthful or untruthful education campaigns that might be most useful.

Raising perceptions of the severity of bad outcomes is likely to have a direct and desired effect. Exaggerating the costs of being pregnant, acquiring AIDS, being jailed for committing a crime, or being addicted is likely to diminish risky behaviors and, hence, improve the welfare of adolescents. In a real-world example of this approach, a recent poster used by the state of California to discourage young men from smoking shows a cigarette drooping, connoting a high risk of sexual dysfunction from smoking. Presumably, the state of California believed that sexual dysfunction will register to the target audience as a more salient cost than (say) lung cancer.

Whereas raising perceptions of the severity of bad outcomes is likely to have the desired effects, our analysis in section 1.4 above suggests that raising perceptions of the likelihood of bad outcomes is not guaranteed to reduce risk-taking behavior. A wariness of exaggerating dangers as a means of discouraging risky behavior is especially warranted in situations in which an adolescent is choosing among substitute risky behaviors. Preaching the dangers of marijuana use can cause more harm than good if it induces young people who use marijuana to exaggerate the degree to which their lives are ruined already and hence to underestimate the additional harm of cocaine or, worse yet, if it leads them to substitute cocaine for marijuana to begin with. This issue is very much on the minds of at least some of those working to discourage risky behavior. There is frequent debate in the gay press, for instance, about whether to emphasize the dangers of acquiring HIV through unprotected oral sex. Some have argued that even truthful emphasis on the dangers of oral sex might lead gay men to engage in unprotected anal sex under the premise that they are at great risk unless they refrain from sexual activity altogether, which they are unwilling to do.

Our policy prescriptions are not very specific or concrete because our analysis draws out general principles rather than specific implications. Research aimed at developing practical policies to combat risky behaviors must be based on analyses that are far more context specific and empirically grounded than ours has been. We hope, however, that by discussing some important lessons from behavioral economics and their potential im-

plications for risky behavior, we have helped lay the groundwork for such research.

References

Ainslie, G. 1975. Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin* 82, no. 4:463–96.

———. 1991. Derivation of "rational" economic behavior from hyperbolic discount curves. *American Economic Review* 81, no. 2:334–40.

———. 1992. *Picoeconomics: The strategic interaction of successive motivational states within the person.* New York: Cambridge University Press.

Ainslie, G., and N. Haslam. 1992a. Hyperbolic discounting. In *Choice over time,* ed. G. Loewenstein and J. Elster. New York: Russell Sage.

———. 1992b. Self-control. In *Choice over time,* ed. G. Loewenstein and J. Elster. New York: Russell Sage.

Akerlof, G., and R. Kranton. 2000. Economics and identity. *Quarterly Journal of Economics* 115:715–53.

Arnett, J. 1994. Sensation seeking: A new conceptualization and a new scale. *Personality and Individual Differences* 16:289–96.

Bauman, K. E. 1980. *Predicting adolescent drug use: Utility structure and marijuana.* New York: Praeger.

Baumrind, D. 1987. A developmental perspective on adolescent risk taking in contemporary America. In *Adolescent social behavior and health* (New Directions for Child Development, no. 37), ed. C. E. Irwin Jr. San Francisco: Jossey-Bass.

Becker, G., and K. Murphy. 1988. A theory of rational addiction. *Journal of Political Economy* 96:675–700.

Beyth-Marom, R., L. Austin, B. Fischhoff, C. Palmgren, and M. J. Quadrel. 1993. Perceived consequences of risky behaviors: Adults and adolescents. *Developmental Psychology* 29:549–63.

Beyth-Marom, R., and B. Fischhoff. 1997. Adolescents' decisions about risks: A cognitive perspective. In *Health risks and developmental transitions during adolescence,* ed. J. Schulenberg, J. L. Mags, and K. Hurrelmann. New York: Cambridge University Press.

Bowman, D., D. Minehart, and M. Rabin. 1999. Loss aversion in a consumption-savings model. *Journal of Economic Behavior and Organization* 38:155–78.

Camerer, C. 1995. Individual decision making. In *Handbook of experimental economics,* ed. J. Kagel and A. E. Roth. Princeton, N.J.: Princeton University Press.

Duesenberry, J. S. 1952. *Income, saving, and the theory of consumer behavior.* Cambridge, Mass.: Harvard University Press.

Fischer, C. 1997. Read this paper even later: Procrastination with time-inconsistent preferences. Ph.D. diss., University of Michigan.

Fischhoff, B. 1988. Judgment and decision making. In *The psychology of human thought,* ed. R. J. Sternberg and E. E. Smith. Cambridge: Cambridge University Press.

———. 1992. Risk taking: A developmental perspective. In *Risk taking,* ed. J. F. Yates. New York: Wiley.

Fischhoff, B., A. M. Parker, W. Bruine de Bruin, J. Downs, C. Palmgren, R. Dawes, and C. F. Manski. 2000. Teen expectations for significant life events. Carnegie-Mellon University. Mimeo.

Frederick, S., and G. Loewenstein. 1999. Hedonic adaptation. In *Well-being: The foundations of hedonic psychology,* ed. D. Kahneman, E. Diener, and N. Schwarz. New York: Russell Sage.

Furby, L., and R. Beyth-Marom. 1992. Risk taking in adolescence: A decision-making perspective. *Developmental Review* 12:1–44.

Green, L., A. Fry, and J. Myerson. 1994. Discounting of delayed rewards: A life-span comparison. *Psychological Science* 5, no. 1:33–36.

Greene, A. L. 1986. Future-time perspective in adolescence: The present of things future revisited. *Journal of Youth and Adolescence* 15:99–113.

Helson, H. 1964. *Adaptation-level theory: An experimental and systematic approach to behavior.* New York: Harper & Row.

Kahneman, D. 1994. New challenges to the rationality assumption. *Journal of Institutional and Theoretical Economics* 150:18–36.

Kahneman, D., E. Diener, and N. Schwartz, eds. 1999. *Well-being: The foundations of hedonic psychology.* New York: Russell Sage.

Kirby, K., and N. Marakovic. 1995. Modeling myopic decisions: Evidence for hyperbolic delay-discounting within subjects and amounts. *Organizational Behavior and Human Decision Processes* 64, no. 1:22–30.

Laibson, D. 1994. Essays in hyperbolic discounting. Ph.D. diss., Massachusetts Institute of Technology.

———. 1996. Hyperbolic discount functions, undersaving, and savings policy. Working Paper no. 5635. Cambridge, Mass.: National Bureau of Economic Research.

———. 1997. Hyperbolic discounting and golden eggs. *Quarterly Journal of Economics* 112, no. 2:443–77.

———. In press. A cue-theory of consumption. *Quarterly Journal of Economics.*

Laibson, D., A. Repetto, and J. Tobacman. 1998. Self-control and saving for retirement. *Brookings Papers on Economic Activity,* no. 1:91–196.

Lewis, C. C. 1981. How adolescents approach decisions: Changes over grades seven to twelve and policy implications. *Child Development* 52:538–44.

———. 1987. Minors' competence to consent to abortion. *American Psychologist* 42:84–88.

Linville, P. W., G. W. Fischer, and B. Fischhoff. 1993. AIDS risk perceptions and decision biases. In *The social psychology of HIV infection,* ed. J. B. Pryor and G. D. Reeder. Hillsdale, N.J.: Erlbaum.

Loewenstein, G. 1996. Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes* 65:272–92.

Loewenstein, G., and S. Frederick. 1997. Predicting reactions to environmental change. In *Environment, ethics, and behavior,* ed. M. Bazerman, D. Messick, A. Tenbrunsel, and K. Wade-Benzoni. San Francisco: New Lexington.

Loewenstein, G., D. Nagin, and R. Paternoster. 1997. The effect of sexual arousal on predictions of sexual forcefulness. *Journal of Crime and Delinquency* 34: 443–73.

Loewenstein, G., T. O'Donoghue, and M. Rabin. 1999. Projection bias in predicting future utility. Mimeo.

Loewenstein, G., and D. Prelec. 1992. Anomalies in intertemporal choice: Evidence and an interpretation. *Quarterly Journal of Economics* 107, no. 2:573–97.

Myerson, J., and L. Green. 1995. Discounting of delayed rewards: Models of individual choice. *Journal of Experimental Analysis of Behavior* 64, no. 3:263–76.

O'Donoghue, T., and M. Rabin. 1999a. Addiction and present biased preferences. Mimeo.

———. 1999b. Addiction and self-control. In *Addiction: Entries and exits,* ed. J. Elster. New York: Russell Sage.

———. 1999c. Doing it now or later. *American Economic Review* 89, no. 1:103–24.
———. 1999d. Procrastination in preparing for retirement. In *Behavioral dimensions of retirement economics,* ed. H. Aaron. Washington, D.C.: Brookings; New York: Russell Sage.
———. In press. Choice and procrastination. *Quarterly Journal of Economics.*
Phelps, E. S., and R. A. Pollak. 1968. On second-best national saving and game-equilibrium growth. *Review of Economic Studies* 35:185–99.
Pollak, R. A. 1968. Consistent planning. *Review of Economic Studies* 35:201–8.
———. 1970. Habit formation and dynamic demand functions. *Journal of Political Economy* 78, no. 4:745–63.
———. 1978. Endogenous tastes in demand and welfare analysis. *American Economic Review* 68, no. 2:374–79.
Quadrel, M. J., B. Fischhoff, and W. Davis. 1993. Adolescent (in)vulnerability. *American Psychologist* 48:102–16.
Rabin, M. 1998. Psychology and economics. *Journal of Economic Literature* 36, no. 1:11–46.
Read, D., and B. van Leeuwen. 1998. Predicting hunger: The effects of appetite and delay on choice. *Organizational Behavior and Human Decision Processes* 76:189–205.
Ryder, H. E., and G. M. Heal. 1973. Optimal growth with intertemporally dependent preferences. *Review of Economic Studies* 40:1–33.
Solnick, J., C. Kannenberg, D. Eckerman, and M. Waller. 1980. An experimental analysis of impulsivity and impulse control in humans. *Learning and Motivation* 11:61–77.
Strotz, R. H. 1956. Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies* 23, no. 3:165–80.
Thaler, R. H. 1991. Some empirical evidence on dynamic inconsistency. In *Quasi rational economics.* New York: Russell Sage.
———, ed. 1992. *The winner's curse: Paradoxes and anomalies of economic life.* New York: Free Press.
Thaler, R., and G. Loewenstein. 1992. Intertemporal Choice. In *The winner's curse: Paradoxes and anomalies of economic life,* ed. R. Thaler. New York: Free Press.
Weithorn, L. A., and S. B. Campbell. 1982. The competency of children and adolescents to make informed treatment decisions. *Child Development* 53:1589–98.
Zuckerman, M., S. Eysenck, and H. J. Eysenck. 1978. Sensation seeking in England and America: Cross-cultural, age, and sex comparisons. *Journal of Consulting and Clinical Psychology* 46:139–49.