Self Control, Risk Aversion, and the Allais Paradox

Drew Fudenberg* and David K. Levine**

First Version: May 12, 2006

This Version: January 24, 2009

This paper develops a dual-self model that is compatible with modern dynamic macroeconomic theory and evidence, and calibrates it to make quantitatively accurate predictions in experiments that display a wide range of behavioral anomalies concerning risk, including the Allais paradox. To obtain a quantitative fit, we extend the simpler "nightclub" model of Fudenberg and Levine [2006] by introducing one additional choice (the choice of a "nightclub," or more generally of anticipated consumption) and one additional parameter that needs to be calibrated. We find that most of the data can be explained with subjective interest rates in the range of 1-7%, short-run relative risk aversion of about 2, and a time horizon of one day for the short-run self.

* Department of Economics, Harvard University

** Department of Economics, Washington University in St. Louis

We thank Daniel Benjamin and Jesse Shapiro for helpful comments and a very careful reading of an early draft, and Eduardo Azevedo and Tao Jin for exceptional research assistance. We are also grateful to Juan Carillo, Ed Glaeser, Glenn Harrison, Yoram Halevy. John Kagel, Stefan Krasa, Drazen Prelec, and seminar participants at Harvard, Ohio State, the University of Illinois Champagne Urbana and the 2006 CEPR conference on Behavioral Economics.

1. Introduction

This paper develops a dual-self model that is compatible with modern dynamic macroeconomic theory and evidence, and calibrates it to make quantitatively accurate predictions in experiments that display a wide range of behavioral anomalies concerning risk: small-stakes risk aversion (the "Rabin paradox"), the Allais paradox, and the effect of cognitive load on risk taking. To obtain a quantitative fit, as opposed to simply showing that the model allows the anomalies, we extend the simpler "nightclub" dual-self model of Fudenberg and Levine [2006] by introducing one additional choice (the choice of a "nightclub," or more generally of anticipated consumption) and one additional parameter that needs to be calibrated.

A key motivation for the paper is our belief that models in behavioral economics, as in other areas of economics, should do more than simply organize the data from a given experiment, and that it is much better to have a small number of models that explain a large number of facts than the reverse. Ideally, a model should not only predict the data to which it was fit, but also make correct predictions about outcomes in other settings, including experiments that have not yet been run. Our first dual-self paper made a step towards that goal by studying a self-control problem derived from a game between a single long-lived self and a sequence of short-term myopic selves. The equilibrium of this game corresponds to a dynamic model of intertemporal choice consistent with those widely and successfully used in macroeconomics, and it qualitatively predicts many behavioral anomalies seen in the laboratory. In particular, while it is designed to be a time-consistent explanation of the facts used to motivate quasi-hyperbolic discounting, the dual-self model predicts the Rabin paradox of inconsistent risk aversion between small and large gambles. It also makes predictions about the impact of cognitive load on decision making. As we argued, existing data on cognitive load suggests that the cost of self-control is not linear but rather strictly convex.[1] This implies that choices involving

[1] The work of Baumeister and collaborators (for example, Muraven et al [1998,2000], Galiot et al [2008]) argues that self-control is a limited resource, moreover one that may be measured by blood glucose levels. The stylized fact that people often reward themselves in one domain (for example, food) when exerting more self control in another (for example, work) has the same implication. This is backed up by evidence from Shiv and Fedorikhin [1999] and Ward and Mann [2000] showing that agents under cognitive load exercise less self-control, for example, by eating more deserts. The first two observations fit naturally with the idea that a common "self-control function" controls many nearly simultaneous choices. The third fits naturally with the hypothesis that self-control and some other forms of mental activity draw on related mental systems or resources. Benahib and Bisin [2005], Bernheim and Rangel [2004], Brocas and Carillo

relatively less tempting options will be made according to the preferences of the long-run self, while choices involving more tempting options will be made according to the preferences of the short-run self. We then pointed out that when the cost function is convex, the dual-self model fails the independence axiom of expected utility theory. In this paper we show that convex cost of self-control can generate an Allais paradox, and that lab data supports the idea that the cost of self-control is convex.

The reason that our model predicts the Allais paradox is that the convexity of the cost function leads to a particular sort of violation of the independence axiom: Agents should be "more rational" about choices that are likely to be payoff-irrelevant. This is exactly the nature of the violation of the independence axiom in the Allais paradox. In the Allais paradox there are two scenarios each involving two options. Under expected utility theory, the same option must be chosen in each scenario, but in practice people choose different options in the two scenarios. A key element of the paradox is that one of the scenarios involves a much smaller probability of winning a prize. That means that there is less temptation to the short-run self. With a convex cost of self-control, less temptation lowers the marginal cost of self-control, so that the long-run self uses more self-control and thus chooses the lottery with the highest long-run utility. Because the long-run self is patient and the lottery is a small share of lifetime wealth, this will be the lottery with the higher expected value. When the chance of winning a prize is high, the temptation great, so the marginal cost of self-control high. In this case the long-run self should allow the short-run self to choose the lottery. Since short-run utility will be essentially the same regardless of whether the prize is large or small, the short-run self prefers the lottery with the highest probability of winning a prize. This is exactly the sort of reversal observed in the Allais paradox. We should emphasize that our theory does not explain all possible violations of the independence axiom: If the choices in each of the two Allais scenarios were reversed, the independence axiom would still be violated, but our explanation would not apply.

Our goals in this paper are not only to formalize the analysis of the Allais paradox, but to construct a form of the dual-self model that can be calibrated to explain a

---

[2005], Loewenstein and O'Donoghue [2005] and Ozdenoren et al [2006] present similar dual-self models, but they do not derive them from a game the way we do, and they do not discuss risk aversion, cognitive load, or the possibility of convex costs of self-control.

range of data about choices over lotteries. To do this, we extend the nightclub model we used in Fudenberg and Levine [2006] by adding an additional choice of "consumption technology." This is necessary not only to explain the Allais paradox with plausible parameter values, but to explain why there is substantial risk aversion even to the very small stakes used in some experiments. Specifically, while our earlier model can explain the examples in Rabin [2000], those examples (such as rejecting a bet that had equal probability of winning $105 or losing $100) understate the degree of risk aversion in small-stakes experiments, where agents are risk averse over much smaller gambles, and fitting our earlier model to these small gambles requires parameter values that conflict both with intuition and with other data.

The idea of the bank/nightclub model is that agents use cash on hand as a commitment device, so that on the margin they will consume all of any small unexpected winnings. However, when agents win large amounts, they choose to exercise self-control and save some of their winnings. The resulting intertemporal smoothing make the agents less risk averse, so that they are less risk averse to large gambles than to small ones. When calibrating the model to aggregate data, we took the underlying utility function to be logarithmic and the same for all consumers. In the present paper we show that this simple specification is not consistent with experimental data on risk aversion and reasonable values of the pocket cash variable against which short-term risk is compared.[2] For this reason we introduce an extension of the nightclub model in which the choice of venue at which short-term expenditures are made is endogenous. This reflects the idea that over a short period of time, the set of things on which the short-run self can spend money is limited, so the marginal utility of consumption decreases fairly rapidly and risk aversion is quite pronounced. Over a longer time frame there are more possible ways to adjust consumption, and also to learn how to use or enjoy goods that have not been consumed before, so that the long-run utility possibilities are the upper envelope of the family of short-run utilities. With the preference that we specify in this paper, this upper

[2] Since the first version of this paper was written, Cox et al [2007] conducted a series of experiments to test various utility theories using relatively high stakes. They also observe that the simple logarithmic model is inconsistent with observed risk aversion, and they argue that the simple linear-logarithmic self-control model does not plausibly explain their data. We will be interested to see whether their data is consistent with the more complex model developed here.

envelope, and thus the agent's preferences over steady state consumption levels, reduces to the logarithmic form we used in our previous paper.

After developing the theory of "endogenous nightclubs," we then calibrate it in an effort to examine three different paradoxes. Specifically, we analyze Rabin paradox data from Holt and Laury [2002], the Kahneman and Tversky [1979] and Allais versions of the Allais paradox, and the experimental results of Benjamin, Brown, and Shapiro [2006], who find that exposing subjects to cognitive load increases their small-stakes risk aversion.[3]

Our procedure is to find a set of sensible values of the key parameters, namely the subjective interest rate, income, the degree of short-term risk aversion, the time-horizon of the short-run self, and the degree of self-control, using a variety of external sources of data. We then investigate how well we can explain the paradoxes using the calibrated parameter values and the dual-self model. How broad of set of parameter values in the calibrated range will explain the paradoxes? To what extent can the same set of parameter values simultaneously explain all the paradoxes? Roughly speaking, we can explain most of the data if we assume an annual subjective interest rate in the range of 1-7%, a short-run "relative risk aversion of consumption" of about 2, and a daily time horizon for the short-run self. We find that the Rabin paradox is relatively insensitive to the exact parameters assumed; the Allais paradox is sensitive to choosing a plausible level of risk aversion; and the Chilean cognitive load data is very sensitive to the exact parameter values chosen. In particular, the Chilean data requires a subjective interest rate of 7%, which is at the high end of the range consistent with macroeconomic data.

We should emphasize that the subject populations we study are heterogeneous and range from Chilean high school students to U.S. and Dutch undergraduates. Moreover, only some subjects exhibit reversals. In short, there is no reason to believe that

---

[3] The main focus of Benjamin, Brown and Shapiro [2006], like that of Frederick [2005], is on the correlation between measures of cognitive ability and the phenomena of small-stakes risk aversion and of a preference for immediate rewards. Benjamin, Brown and Shapiro find a significant and substantial correlation between with each of these sorts of preferences and cognitive ability. They also note that the correlation between cognitive ability and time preference vanishes when neither choice results in an immediate payoffs, and that the correlation between small-stakes risk aversion and "present bias" drops to zero once they control for cognitive ability. This evidence is consistent with our explanation of the Rabin paradox, as it suggests that that small-stakes risk aversion results from the same self-control problem that leads to a present bias in the timing of rewards. They also discuss the sizable literature that examines the correlation between cognitive ability and present bias without discussing risk aversion.

the reversals we observe in the data can all be explained by a single common set of "representative" parameter values. Consequently instead of trying to focus on the single parameter constellation that would best fit all the data, we focus on the range and robustness of the parameter values that can be used to explain reversals, and the fact that even across subjects and populations this range is roughly the same.

After showing that the base model provides a plausible description of data on attitudes towards risk, gambles, and cognitive load, we examine the robustness of the theory. Specifically, in the calibrations we assume that the opportunities presented in the experiments are unanticipated, so we consider what happens when gambling opportunities are foreseen.

2. Self-Control, Cash Constraints, and Target Consumption

We consider an infinite-lived consumer making a savings decision. Each period $t = 1, 2, \ldots$ is divided into two sub-periods, the *bank* sub-period and the *nightclub* sub-period. Wealth at the beginning of the bank sub-period is denoted by w_t. During the "bank" sub-period, consumption is not possible, and wealth is divided between savings s_t, which remains in the bank, pocket cash x_t which is carried to the nightclub, and durable consumption c_t^d which is paid for immediately and is consumed in the second sub-period of period t.[4] In the nightclub consumption $0 \le c_t \le x_t$ is determined, with $x_t - c_t$ returned to the bank at the end of the period. Wealth next period is just $w_{t+1} = R(s_t + x_t - c_t)$. For simplicity money returned to the bank bears the same rate of interest as money left in the bank. No borrowing is possible, and there is no other source of income other than the return on investment.

Consumption Commitment: So far, we have followed Fudenberg and Levine [2006]. Now we consider an extension of that model that we will use to explain the degree of risk aversion we observe in experimental data. Specifically, we suppose that there is a choice of nightclubs to go to in the nightclub sub-period. These choices are indexed by the

[4] Durable and/or committed consumption is a significant fraction (roughly 50%) of total consumption so we need to account for it in calibrating the model, but consumption commitments are not our focus here. For this reason we use a highly stylized model, with consumption commitments reset at the start of each time period. A more realistic model of durable consumption would have commitments that extend for multiple periods, as in Grossman and Laroque [1990].

quality of the nightclub $c_t^* \in (0,\infty)$. In a nightclub of quality c_t^* we assume that the utility of the short-run self has the form $u(c_t \mid c_t^*)$ depending on the amount consumed c_t there and the quality of the nightclub.

The utility function at the nightclub is assumed to satisfy $u(c_t \mid c_t^*) \leq u(c_t \mid c_t)$. This means that when planning to consume a given amount c_t it is best to choose the nightclub of the same index. Intuitively, the quality c_t^* of a nightclub represents a "target" level of consumption expenditure at that nightclub. That is, if you are going to consume a low level of c_t you would prefer to spend it at a nightclub with a low value of c_t^* and you would like to consume a high level of c_t at a high quality nightclub.

It is useful to think of a low value of c_t^* as representing a nightclub that serves cheap beer, while a high value of c_t^* represents a nightclub that serves expensive wine. At the beer bar c_t represents expenditure on cheap beer, while at the wine bar it represents the expenditure on expensive wine. The assumption that $u(c_t \mid c_t^*) \leq u(c_t \mid c_t)$ captures the idea that spending a large amount at a low quality nightclub results in less utility than spending the same amount at a high quality nightclub: lots of cheap beer is not a good substitute for a nice bottle of wine. Conversely, spending a small amount at a high quality nightclub results in less utility than spending the same amount at a low quality nightclub: a couple of bottles of cheap beer are better than a thimble-full of nice wine. People with different income and so different planned consumption levels will choose consumption sites with different characteristics. The quality of a nightclub can also be interpreted as a state variable or capital stock that reflects experience with a given level of consumption: a wine lover who unexpectedly wins a large windfall may take a while both to learn to appreciate differences in *grands crus* and to learn which ones are the best values.[5]

We assume that $u(c_t \mid c_t) = \log c_t$; this ensures that in a deterministic and perfectly foreseen environment without self control costs, behavior is the same as with standard logarithmic preferences. To avoid uninteresting approximation issues, we assume that there are a continuum of different kinds of nightclubs available, so that there are many intermediate choices between the beer bar and wine bar.

[5] To fully match the model, this state variable needs to reflect only recent experience: a formerly wealthy wine lover who has been drinking vin de table for many years may take a while to reacquire both a discerning palate and up-to-date knowledge of the wine market.

There are a great many possible functional forms satisfying these properties. Our choice of a specification is guided both by analytic convenience and by evidence (examined below) that short-term risk preferences seem more risk averse than consistent with the logarithmic specification, even when self-control costs are taken into account. This leads us to adopt the functional form

$$u(c_t \mid c_t{*}) = \log c_t{*} - \frac{(c_t / c_t{*})^{1-\rho} - 1}{\rho - 1},$$

where $\rho \geq 1$ corresponds to the short-run self's relative risk aversion over immediate consumption.

With this specification $u(c_t \mid c_t) = \log(c_t)$, and

$$\frac{\partial u(c_t \mid c_t^*)}{\partial c_t^*} = \frac{1}{c_t^*} - \left(\frac{c_t^*}{c_t}\right)^{\rho-2} \frac{1}{c_t}.$$

As a consequence, the first order condition for maximizing $u(c_t \mid c_t{*})$ with respect to $c_t{*}$ implies $c_t{*} = c_t$, and the second order condition is

$$\left.\frac{\partial^2 u(c_t \mid c_t{*})}{\partial c_t^{*2}}\right|_{c_t = c_t^*} = -\frac{1}{c_t^2} - (\rho - 2)\frac{1}{c_t^2} = \frac{1}{c_t^2}(1 - \rho),$$

which is negative when $\rho > 1$.

Durable Consumption: The next step is to specify the agent's preferences for durable versus non-durable consumption. Our goal here is simply to account for the fact that durable consumption exists, and not to explain it, so we adopt a simple Cobb-Douglas-like specification $\tau u(c_t \mid c_t{*}) + (1-\tau)\log c_t^d$; this will lead to a constant share τ of spending on durables. Durable consumption c_t^d can only be adjusted slowly, and seems unlikely to respond at all to the sorts of income shocks received in the lab experiments we study. For this reason we simplify the model by assuming that the time path of c_t^d is chosen for once and for all in the initial time period.[6]

[6] Similarly, we abstract from labor supply, precautionary savings motives, and so forth. However we explicitly introduce durable consumption so that when we calibrate pocket cash the short-run self does not perceive that the rent check and similar expenses are available for short-term amusement.

Cost of Self-Control: The long-run self maximizes the expected discounted present value of the utility of the short-run selves. This is done subject to a cost of self-control. This cost depends on the resources the short-run self perceives as available to himself. These resources determine a *temptation utility* for the short-run self, representing the utility the short-run self perceives as available if allowed unfettered access to those resources, free from the bounds of self-control exerted by the long-run self. Denote this temptation utility by $\overline{u}_t$. The actual realized utility that the long-run self allows the short-run self is u_t, and there may be cognitive load due to other activities, d_t. Then the cost of self-control is $g(d_t + \overline{u}_t - u_t)$ and where the function *g* is continuously differentiable and convex. Until section 7 we suppose that there is no cognitive load from other activities, and set $d_t = 0$. The key idea here is that the cost of self-control depends on the difference between the utility the short-run self is tempted by $\overline{u}_t$ and the utility the short-run self is allowed u_t. In our calibrations of the model, we will take the cost function to be quadratic: $g(v_t) = \gamma v_t + (1/2)\Gamma v_t^2$.

Long-run Self: In the bank no consumption is possible, and so there is no temptation for the short-run self. In the nightclub the short-run self cannot borrow, and wishes to spend all of the available pocket cash x_t on consumption. This pocket cash functions as a commitment device: by spending money on durable consumption or leaving it in the bank, it is not available to the short-run self in the nightclub, so does not represent a temptation that must be overcome by costly self-control.

The problem faced by the long-run self is to choose pocket cash and consumption to maximize the present value using the discount factor δ of short-run self utility net of the cost of self-control. The objective function of the long-run self is

$$U_{RF} = E\sum_{t=1}^{\infty} \delta^{t-1}\left[\tau\left(u(c_t \mid c_t^*) - g(u(x_t \mid c_t^*) - u(c_t \mid c_t^*)\right) + (1-\tau)\log c_t^d)\right] \tag{2.1}$$

which is to be maximized with respect to $c_t \geq 0, c_t^* \geq 0, c_t^d \geq 0, x_t \geq 0$ subject to w_1 given, $w_{t+1} = R(s_t + x_t - c_t)$, $s_t + x_t + c_t^d \leq w_t$ and $w_t \geq 0$.

In this formulation there is a single long-run self with time-consistent preferences. Although the impulsive short-run selves are the source of self-control costs, the equilibrium of the game between the long run self and the sequence of short run selves is equivalent to the optimization of this reduced-form control problem by the single long-run self.

Solution in the Deterministic Case: Suppose that there is no uncertainty, so this is a simple deterministic infinite-horizon maximization problem. Because there is no cost of self-control in the bank, the solution to this problem is to choose $c_t^* = c_t = x_t$. In other words, cash x_t is chosen to equal the optimal consumption for an agent without self-control costs, and c_t^* is the nightclub of the same quality. The agent then spends all pocket cash at the nightclub, and so incurs no self-control cost there. Since $c_t = c_t^* = x_t$, the utility of the short run self is $u(x_t \mid c_t^*) = \log x_t$, and as there is no self-control cost, this boils down to maximizing

$$\sum_{t=1}^{\infty} \delta^{t-1} \left[\tau \log x_t + (1-\tau) \log c_t^d) \right]$$

subject to the budget constraint $w_{t+1} = R(w_t - x_t - c_t^d)$. The solution is easily computed[7] to be $x_t = (1-\delta)\tau w_t$, and $c_t^d = (1-\tau)(1-\delta)w_t$. The corresponding present value utility of the long-run self is

$$U_1(w_1) = \frac{\log(w_1)}{1-\delta} + K \, ,$$

where

$$K = \frac{1}{1-\delta}[\log(1-\delta) + \delta \log(R\delta) + \tau \log \tau + (1-\tau)\log(1-\tau)] .$$

These together give the solution of the simple deterministic budget problem.

Because the time path of durable consumption c_t^d is chosen once and for all in period 1 as a function of initial wealth w_1, the utility from period 2 on is

$$U_2(w_2 \mid w_1) = \frac{\tau \log(w_2 - (1-\tau)R\delta w_1)}{1-\delta} + \frac{(1-\tau)\log((1-\tau)R\delta w_1)}{1-\delta} + K \, ,$$

where K is the same constant as above.

[7] The derivation is standard; an explicit computation in the case where $\tau = 1$ is in Fudenberg and Levine [2006]. Note that equation (1) of that paper contains a typographical error: in place of $(1+\gamma)(\log(1-a) + \log y_0)$ it should read $(1+\gamma)\log(1-a) + \log(y_0)$.

The demand for costly self-commitment in order to reduce the future cost of self-control has many implications. For example, individuals may pay a premium to invest in illiquid assets, as do the quasi-hyperbolic agents in Laibson [1997]. They may also choose to carry less cash than in the absence of self-control costs.

Uncertainty and Unforeseen Choices: The deterministic perfect-foresight savings model is too simple to account for any behavioral phenomena. However, the model does predict that there can be "preference reversals" in evaluating some intertemporal tradeoffs when there is uncertainty that resolves in the nightclub as opposed to as the bank.

Specifically, suppose that unexpectedly the short-run self at the nightclub is offered a choice between an amount z_1 today and an amount θz_1 tomorrow, where $\theta > R$. Let $\bar{w}_2$ denote period 2 wealth in the absence of this unexpected opportunity, and let c_1^* and c_1^d be the quality of the first-period nightclub and first-period durables, which were also chosen before going to the nightclub. Because the choice is unexpected, we have $x_1 = c_1^*$.[8] To analyze this binary choice problem, it is convenient to consider the auxiliary problem where the agent can choose to consume any non-negative portion of z_1 at the nightclub, with the balance "invested" with return θ, so that consuming $c_1 \in [c_1^*, c_1^* + z_1]$ leads to period-2 wealth $\theta(c_1^* + z_1 - c_1) + \bar{w}_2$. In this auxiliary problem, the agent maximizes

$$\begin{aligned}&\tau\left[u(c_1 \mid c_1^*) - g(u(x_1 + z_1 \mid c_1^*) - u(c_1 \mid c_1^*))\right] + (1-\tau)\log c_1^d \\ &+\delta U_2[(\bar{w}_2 + \theta(z_1 - (c_1 - x_1)) \mid w_1)]\end{aligned}$$

over $c_1 \in [c_1^*, c_1^* + z_1]$. The derivative of this objective function with respect to c_1 is

$$\tau\left[1 + g'(u(x_1 + z_1 \mid c_1^*) - u(c_1 \mid c_1^*))\right]\frac{\partial u}{\partial c_1}(c_1 \mid c_1^*) - \delta\theta\frac{\partial U_2}{\partial w_2}[(\bar{w}_2 + \theta(z_1 - (c_1 - x_1)) \mid w_1)],$$

In contrast, the first order condition in the original problem at the bank was

$$\tau\frac{\partial u}{\partial c_1}(c_1 \mid c_1^*) - \delta R\frac{\partial U_2}{dw_2}[(\bar{w}_2 \mid w_1)] = 0.$$

---

[8] The analysis would be essentially the same if this choice was foreseen but had very low ex-ante probability, as in this case it would have a negligible effect on the decisions made at the start of period 1.

Define $\gamma_1 = \lim_{z_1 \downarrow 0} g'(u(x_1 + z_1 \mid c_1^*) - u(c_1^* \mid c_1^*))$ to be the marginal cost of self-control at (c_1^*, c_1^*). The first order condition for the original bank problem implies that in the limit as $z_1 \to 0$, the derivative of the at-the-nightclub objective function evaluated at $c_1 = c_1^*$ is

$$([1 + \gamma_1]\delta R - \delta\theta)\frac{\partial U_2}{dw_2}[(\bar{w}_2 \mid w_1)].$$

This means that if z_1 is small and $\theta < (1 + \gamma_1)R$, the long-run self at the nightclub will strictly prefer to consume now z_1 rather than θz_1 in the future. This is despite the fact that given the choice between z_1 at future date t and θz_1 at date $t + 1$ the long-run self will strictly prefer the later date. The reason is simply that at the nightclub the short-run self was rationed by the available cash x_1. An extra cash payment of z_1 today will cause a temptation to increase spending that is costly for the long-run self to control. By contrast there is no temptation associated with future payoffs, and so there can be "preference reversal" whenever the cost of resisting the short-run temptation is sufficiently high.

Now suppose that the choice instead of between certain rewards the rewards $z_1, \theta z_1$ have a common probability p less than one of being received (so with complementary probability $1 - p$ the reward is 0.) Because the reward is less likely, the costs and benefits of resisting temptation are lower. If the cost of self-control is linear, these effects exactly offset, but if marginal cost is increasing ($\Gamma > 0$), then when the rewards are less likely, the marginal cost of self-control is smaller as compared to its benefit, so preference reversal will only occur for a smaller range of returns θ. In particular, for some values of θ the reversal will occur when the reward is certain, but not when there is a sufficiently small chance of receiving the reward. The following data from Keren and Roelsofsma [1995] shows that this is exactly what happens.[9]

[9] This experimental result is confirmed by Weber and Chapman [2005], and discussed in Halevy [2008], who proposed an objective function that is consistent with these choices. Note that the experiment was in Dutch Florins. We converted from Dutch Florins to U.S. Dollars using an exchange rate typical of the early 1990s of 1.75 Florin per Dollar.

Table 1 – "Hyperbolic" Discounting with Random Reward

		Probability of reward	
		1.0	0.5
A	$175 now	0.82	0.39
	$192 4 weeks	0.18	0.61
B	$172 26 weeks	0.37	0.33
	$192 30 weeks	0.63	0.67

Note that this dependence of the choices on the probability of reward is not consistent with quasi-hyperbolic preferences (as in Laibson [1997]) or with the version of the independence axiom (for choices over menus) imposed as an axiom by Gul and Pesendorfer [2001] and Dekel, Lipman, and Rustichini [2008].

Mental Accounting: A crucial aspect of the model is the "pocket cash" x_t that serves to ration consumption and so reduce the temptation to the sort-run self. In the simple perfect-foresight version of the model, it is optimal to give the short-run self exactly the amount to be spent at the nightclub, and so avoid temptation and self-control cost entirely. In effect the long-run self hands the pocket cash to the short-run self to take to the nightclub and says "here…go crazy….spend it all." The actual decision about how much pocket cash to allocate to the short-run self is taken at a location – "the bank" – where there are no tempting consumption possibilities.

In Fudenberg and Levine [2006] the notion of a bank and pocket cash were taken literally. In practice there are many strategies that individuals use to reduce the temptation for impulsive expenditures. The view we take here is that pocket cash is determined by mental accounting of the type discussed by Thaler [1980], and not necessarily by physically isolating money in a bank. In other words, x_t should not be viewed as the literal amount of money the short-run self has in their wallet or the amount available including cash cards and so forth, but should be viewed as the amount of resources that the short-run self feels entitled to use. The strategies individuals use for this type of commitment can be varied. For example some people may choose to carry only a limited amount of cash and no credit cards. Others may allow the short-run self to spend money only in the "right pocket." Yet others may engage in more direct mental accounting of the form "you may spend $100 at the nightclub, but no more."

In this formulation we do not imagine that x_t is directly observable, for example by surveying individuals about how much money they have in their wallets. However, because of the intertemporal optimization problem, we can calculate x_t (along with c_t^* which is also not directly observable) from knowledge of the underlying parameters of preferences δ, ρ.

3. Risky Drinking: Nightclubs and Lotteries

Suppose in period 1 (only) that when the agent arrives at the nightclub of her choice, she has the choice between two lotteries, A and B with returns $\tilde{z}_1^A, \tilde{z}_1^B$. Initially we will assume that this choice is completely unanticipated – that is, it has prior probability zero. This means that $c_t^* = x_t = (1-\delta)\tau w_t$. What is the optimal choice of lottery given x_1, c_1^*? For simplicity, we assume throughout that the agent does not expect to have lottery opportunities at nightclubs after period 1; as noted earlier, the overall savings and utility decision will not change significantly provided that the probability of getting future lottery opportunities is small.

The lotteries $\tilde{z}_1^A, \tilde{z}_1^B$ may involve gains or losses, but we suppose that the largest possible loss is less than the agent's pocket cash. There are number of different ways that the dual-self model can be applied to this setting, depending on the timing and "temptingness" of the choice of lottery and spending of its proceeds. In this paper we assume that the short-run self in the nightclub simultaneously decides which lottery to pick and how to spend for each possible realization of the lottery.

Since the highest possible short-run utility comes from consuming the entire proceed of the lottery, the temptation utility is $\max\{Eu(x_1+\tilde{z}_1^A, c_1^*), Eu(x_1+\tilde{z}_1^B, c_1^*)\}$ where $\tilde{z}_1^j$ is the realization of lottery $j = A, B$. This temptation must be compared to the expected short-run utility from the chosen lottery. If we let $\tilde{c}_1^j(z_1^j)$ be the consumption chosen contingent on the realization of lottery *j,* the self-control cost is

$$g\Big(\max\{Eu(x_1+\tilde{z}_1^A, c_1^*), Eu(x_1+\tilde{z}_1^B, c_1^*)\} - Eu(\tilde{c}_1^j, c_1^*)\Big).$$

Let γ_1 denote the marginal cost of self-control in the first period. We let $\hat{c}_1^j(\gamma_1)(z_1^j)$ be the solution to the first-order condition for a maximum for a given marginal cost of self control, that is, the unique solution to

$$(c_1^j)^\rho = \left(c_1^*\right)^{\rho-1} \frac{(1-\delta)(1+\gamma_1)}{\delta}\left(\tau w_1 + z_1^j - c_1^j\right),$$

and find the corresponding marginal cost of self control

$$\hat{\gamma}_1^j(\gamma_1) =$$
$$g'(\max\{Eu(x_1 + \tilde{z}_1^A, c_1^*), Eu(x_1 + \tilde{z}_1^B, c_1^*)\} - Eu(\min\{\hat{c}_1^j(\gamma_1)(\tilde{z}_1^j), x_1 + \tilde{z}_1^j\}, c_1^*))$$

We show in the Appendix that we can characterize the optimum as follows:

Theorem 1*: For given* (x_1, c_1^*) *and each* $j \in \{A, B\}$ *there is a unique solution to*

$$\gamma_1^j = \hat{\gamma}_1^j(\gamma_1^j)$$

and this solution together with $\tilde{c}_1^j = \min(\hat{c}_1^j(\gamma_1^j)(z_1^j), x_1 + z_1^j]$ *and the choice of* j *that maximizes long-run utility is necessary and sufficient for an optimal solution to the consumer's choice between lotteries A and B.*

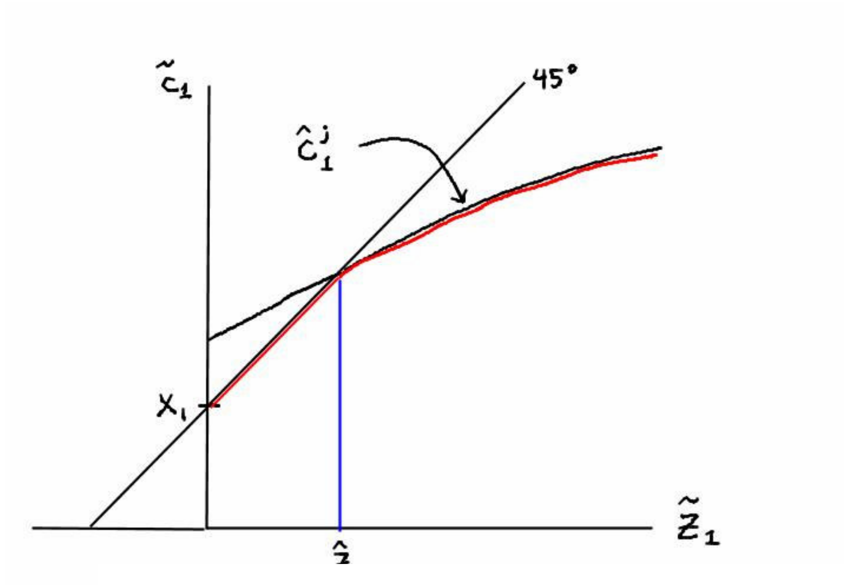


Figure 1 - The "Consumption" Function

The "consumption function" is $\tilde{c}_1^j = \min(\hat{c}_1^j(\gamma_1^j)(z_1^j), x_1 + z_1^j\}$. Let $\hat{z}$ such that all the winnings are spent, that is, $\hat{c}_1^j(\gamma_1)(\hat{z}) = x_1 + \hat{z}$. From the first order condition this can be computed equal to

$$\hat{z}_1 = \left(c_1^*\right)^{\frac{\rho-1}{\rho}} \left[\frac{1-\delta}{\delta}(1+\gamma_1)[\tau w_1 - x_1]\right]^{1/\rho} - x_1$$

Note that for arbitrary x_1, c_1^* we may have $\hat{z}_1$ negative. This is sketched in Diagram 1. For $z_1^j < \hat{z}_1$ no self-control is used, and all winnings are spent. Above this level self control is used, with only a fraction of winnings consumed, and the rest going to savings. When the time period is short, $\hat{c}_1^j$ is very flat, so that only a tiny fraction of the winnings are consumed immediately when receipts exceed the critical level. Thus when the agent is patient he is almost risk neutral with respect to large gambles. However the agent is still risk averse to small gambles, as these will not be smoothed but will lead to a one for one change in current consumption.

Increasing Marginal Cost of Self-Control: So far we have focused on the first-order effect of self-control costs. We showed that when an unexpected gamble arises at the nightclub the problem of self-control that was not present at the bank leads to a wedge in marginal utilities and consequently a high marginal propensity to consume out of small gains. Our main goal is to understand how changes in the chances of winning a prize, as in the Allais paradox or the Keren and Roelsofsma [1995] data, can lead to preference reversals. As we shall see, changing the chances of winning a prize has complicated effects, depending both on temptation, how the overall prize money is spent, and so forth. However, we can gain some intuition about these effects by considering the simple conceptual experiment of changing the temptation holding fixed the constraints and other parameters of the decision problem.

Specifically, the overall objective function we are interested in has the form

$$U_1(\overline{u}) =$$
$$\tau\Big(u(c_1 \mid c_1^*) - g(\overline{u} - u(c_1 \mid c_1^*))\Big) + (1-\tau)\log c_1^d + \delta U_2(\overline{w}_2 + \theta(z_1 - (c_1 - x_1)))$$

where $\overline{u}$ represents the temptation. This weights the utility of the current short-run self $\tau\Big(u(c_1 \mid c_1^*) - g(\overline{u} - u(c_1 \mid c_1^*)\Big) + (1-\tau)\log c_1^d)$ against the utility of future short-run selves $\delta U_2(\overline{w}_2 + \theta(z_1 - (c_1 - x_1)))$. What happens to the objective function with quadratic cost of self-control when we change the (unforeseen) temptation $\overline{u}$ to $\overline{u}\,'$? Then

$$\begin{aligned}(U_1(\overline{u}\,') - U_1(\overline{u}))/\tau &= -\gamma(\overline{u}\,' - \overline{u}) - (\Gamma/2)((\overline{u}\,' - u(c_1 \mid c_1^*))^2 - (\overline{u} - u(c_1 \mid c_1^*))^2) \\ &= -\left[\gamma(\overline{u}\,' - \overline{u}) + (\Gamma/2)(\overline{u}\,'^2 - \overline{u}^2)\right] + \Gamma(\overline{u}\,' - \overline{u})u(c_1 \mid c_1^*)\end{aligned}$$

The term in square brackets just involves $\overline{u}\,', \overline{u}$ which are constants that do not matter for the decision. The second term shows that if $\Gamma > 0$, the weight $\Gamma(\overline{u}\,' - \overline{u})$ on the utility $u(c_1 \mid c_1^*)$ of the current short-run self increases as the temptation increases. In other words, more temptation implies greater weight on the utility of the current short-run self in the optimal decision. If there is great temptation then the current short-run self calls the shots; if there is little temptation, then the agent's maximization problem is more like the usual exponential discounting case. It is this idea that generates reversals such as those in the Allais paradox and the Keren and Roelsofsma [1995] data.

Note well the implications of this analysis for risk aversion. The preferences of the current short-run self $u(c_1 \mid c_1^*)$ are CES with the base level of "wealth" corresponding to the endogenous pocket cash x_1. On the other hand, the preference of the future short-run selves, as measured by $U_2(\overline{w}_2 + \theta(z_1 - (c_1 - x_1)))$, are logarithmic with the base level of wealth equal to lifetime wealth. In other words, the current-short self is very risk averse, and the future long-run self is nearly risk neutral. So as the temptation increases and more weight is placed on the preferences of the current short-run self, the individual will behave in a more risk averse fashion.

4. Basic Calibration

The first step in our calibration of the model is to pin down as many parameters as possible using estimates from external sources of data. We will subsequently use data from laboratory experiments to calibrate risk aversion parameters and to determine the cost of self control.

To measure the subjective interest rate r we ordinarily think of taking the difference between the real rate of return and the growth rate of per capital consumption. However, we must contend with the equity premium puzzle. From Shiller [1989], we see that over a more than 100 year period the average growth rate of per capita consumption has been 1.8%, the average real rate of returns on bonds 1.9%, and the real rate of return on equity 7.5%. Fortunately if the consumption lock-in once a nightclub is chosen lasts for six quarters,[10] the problem of allocating a portfolio between stocks and bonds is essentially the same as that studied by Gabaix and Laibson [2001], which is a simplified version of Grossman and Laroque [1990].[11] Their calibrations support a subjective interest rate of 1%. This rate, and any rate in the range 1-7%, can explain the data on the Allais paradox.[12] To explain the Benjamin, Brown, and Shapiro data on Chilean high school students requires a higher interest rate of 7%, at the high end of what can be supported by Shiller's data.

From the Department of Commerce Bureau of Economic Analysis, real per capita disposable personal income in December 2005 was $27,640. To consider a range of income classes, we will use three levels of income $14,000, $28,000, and $56,000. To

---

[10] We have implicitly assumed it lasts for only a day, but the length of lock-in plays no role in the analysis, no result or calculation changes if the lock-in is six quarters.

[11] They assume that once the nightclub is chosen, no other level of consumption is possible. We allow deviations from the nightclub level of consumption – but with very sharp curvature, so in practice consumers are "nearly locked in" to their choice of nightclub. Chetty and Szeidl [2006] show that these models of sticky consumption lead to the same observational results as the habit formation models used by Constantinides [1990] and Boldrin, Christiano and Fisher [2001].

[12] Subjective interest rates outside this range are also consistent with the Allais data, we did not explore this as we focus on the range that has some prior support from macroeconomic data.

infer consumption from the data we do not use current savings rates, as these are badly mis-measured due to the exclusion of capital gains from the national income accounts. We instead use the historical long-term savings rate of 8% (see FSRB [2002]) measured when capital gains were not so important. This enables us to determine wealth and consumption from income.

We estimate wealth as annual consumption divided by our estimate of the subjective interest rate r: $w_1 = 0.92y_1 / r$, where y_1 denotes steady state income. In determining pocket cash, we need take account of consumption c_t^d that is not subject to temptation: housing, consumer durables, and medical expenses. At the nightclub, the rent or mortgage was already paid for at the bank, and it is not generally feasible to sell one's car or refrigerator to pay for one's impulsive consumption. As noted by Grossman and Laroque [1990], such consumption commitments increase risk aversion for cash gambles.[13] For consumption data, we use the National Income and Product Accounts from the fourth quarter of 2005. In billions of current dollars, personal consumption expenditure was \$8,927.8. Of this \$1,019.6 was spent on durables, \$1,326.6 on housing, and \$1,534.0 on medical care, which are the non-tempting categories. This means that the share of income subject to temptation $\tau = 0.57$.

Finally, we must determine the time horizon Δ of the short-run self. This is hard to pin down accurately, in part because it seems to vary both within and across subjects, but the most plausible period seems to be about a day. For the purposes of robustness we checked that none of our results are sensitive to assuming a time horizon of a week: details can be found in the earlier working paper version available on line.

Putting together all these cases, we estimate pocket cash to be
$x_1 = 0.57 \times 0.92 \times y_1 / 365 = .00144 \times y_1$.

[13] Chetty and Szeidl [2006] extend Grossman and Laroque to allow for varying sizes of gambles and costly revision of the commitment consumption. Postelwaite, Samuelson and Silberman [2006] investigate the implications of consumption commitments for optimal incentive contracts.

Table 2 – Calibrated Parameter Summary

Percent interest r		$y_1 = 14$K		$y_1 = 28$K		$y_1 = 56$K	
annual	daily	w_1	x_1	w_1	x_1	w_1	x_1
1	.003	1.3M	20	2.6M	40	5.2M	80
3	.008	.43M		.86M		1.7M	
5	.014	.30M		.61M		1.2M	

To determine a reasonable range of self control costs, we need to find how the marginal propensity to consume "tempting" goods changes with unanticipated income. The easiest way to parameterize this is with the "self-control threshold," which is the level of consumption at which self-control kicks in. The consumption cutoff corresponding to $\hat{z}_1$ is given by

$$\hat{c}_1 \equiv x_1 + \hat{z}_1 = (x_1)^{\frac{\rho-1}{\rho}} \left[\frac{\tau(1-\delta)}{\delta}(1+\gamma)[w_2] \right]^{1/\rho}$$
$$\approx x_1 (1+\gamma)^{1/\rho}$$

where we use the facts that $w_2 \approx w_1 = x_1 / \tau(1-\delta)$, and that $\delta \approx 1$. Define $\mu_1(\gamma_1) = (1+\gamma_1)^{1/\rho} \approx \hat{c}_1 / x_1$. Because γ_1 is measured in units of utility, its numerical value is hard to interpret. For this reason we will report $\mu_1(\gamma_1)$ rather than γ_1.

We can also relate μ_1 to consumption data. Abdel-Ghany et al [1983] examined the marginal propensity to consume semi- and non-durables out of windfalls in 1972-3 CES data.[14] In the CES, the relevant category is defined as "inheritances and occasional large gifts of money from persons outside the family...and net receipts from the settlement of fire and accident policies," which they argue are unanticipated. For windfalls that are less than 10% of total income, they find a marginal propensity to consume out of income of 0.94. For windfalls that are more than 10% of total income they find a marginal propensity to consume out of income of 0.02. Since the reason for

[14] The Imbens, Rubin and Sacerdote [2001] study of consumption response to unanticipated lottery winnings shows that big winners earn less after they win, which is useful for evaluating the impact of winnings on labor supply. Their data is hard to use for assessing μ_1, because lottery winnings are paid as an annuity and are not lump sum, so that winning reduces the need to hold other financial assets. It also appears as though the lottery winners are drawn from a different pool than the non-winners since winners earn a lot less than non-winners before the lottery.

the 10% cutoff is not clear from the paper, we will view 10% as a general indication of the cutoff. [15] As we have figured the ratio of income to pocket cash to be $y_1 / x_1 = 696$, the value of μ_1 corresponding to 10% of annual income is 69.6.

5. Small Stakes Risk Aversion

To demonstrate how the model works and calibrate the basic underlying model of risk preference, we start with the "Rabin Paradox": the small-stakes risk aversion observed in experiments implies implausibly large risk aversion for large gambles.[16] The central issue is the case of small gambles. Following Rabin's proposal, let option A be $(.5 : -100, .5 : 105)$, while option B is to get nothing. We expect that as Rabin predicts many people will choose B. Since the combination of pocket cash and the maximum winning is well below our estimates of $\hat{c}_1$, all income is spent, and the consumer simply behaves as a risk-averse individual with wealth equal to pocket cash and a coefficient of relative risk aversion of ρ. Let us treat pocket cash as an unknown for the moment, and ask how large could pocket cash be given that a logarithmic consumer is willing to reject such a gamble. That is, we solve $.5\log(x_1 - 100) + .5(x_1 + 105) = \ln(x_1)$ for pocket cash; for larger values of x_1 the consumer will accept the gamble, and for smaller ones he will reject. In this sense, as we argued in Fudenberg and Levine [2006], short-run logarithmic preferences are consistent with the Rabin paradox.[17]

The problem with this analysis is that the gamble $(.5 : -100, .5 : 105)$ has comparatively large stakes. Laboratory evidence shows that subjects will reject considerably smaller gambles, which is harder to explain with short-run logarithmic preferences. We use data from Holt and Laury [2002], who did a careful laboratory study of risk aversion. Their subjects were given a list of ten choices between an A and a B lottery. The specific lotteries are shown below, where the first four columns show the

---

[15] Landsburger [1966] with both CES data and with data on reparation payments by Germany to Israeli citizens reaches much the same conclusion.

[16] Rabin thus expands on an earlier observation of Samuelson [1963].

[17] Note that this theory predicts that if payoffs are delayed sufficiently, risk aversion will be much lower. Experiments reported in Barberis, Huang and Thaler [2003] suggest that there is appreciable risk aversion for gambles where the resolution of the uncertainty is delayed as well as the payoffs themselves. However, delayed gambles are subject to exactly the same self-control problem as regular ones, so this is consistent with our theory. In fact the number of subjects accepting the risky choice in the delayed gamble was in fact considerably higher than the non-delayed gamble, rising from 10% to 22%.

probabilities of the rewards, and the first four rows, which are irrelevant to our analysis are omitted.

Table 3 – Laboratory Preferences Towards Risk

Option A		Option B		Fraction of Subjects Choosing A			
\$2.00	\$1.60	\$3.85	\$0.10	1X	20X	50X	90X
0.5	0.5	0.5	0.5	.70	.85	1.0	.90
0.6	0.4	0.6	0.4	.45	.65	.85	.85
0.7	0.3	0.7	0.3	.20	.40	.60	.65
0.8	0.2	0.8	0.2	.05	.20	.25	.45
0.9	0.1	0.9	0.1	.02	.05	.15	.40
1.0	0.0	1.0	0.0	.00	.00	.00	.00

Initially subjects were told that one of the ten rows would be picked at random and they would be paid the amount shown. Then they were given the option of renouncing their payment and participating in a high stakes lottery, for either 20X, 50X or 90X of the original stakes, depending on the treatment. The high-stakes lottery was otherwise the same as the original: a choice was made for each of the ten rows, and one picked at random for the actual payment. Everyone in fact renounced their winnings from the first round to participate in the second. The choices made by subjects are shown in the table above.

In the table we have highlighted (in yellow and turquoise respectively) the decision problems where roughly half and 15-20% of the subjects chose A. We will take these as characterizing median and high risk aversion respectively. The bottom 15th percentile exhibits little risk aversion, suggesting that perhaps they do not face much in the way of a self-control problem.

Since the stakes plus pocket cash remain well below our estimate of $\hat{c}$, we can fit a CES utility function with respect to our pocket cash estimates of \$21, \$42, \$84, \$155, \$310 and \$620, in each case estimating the value of ρ that would leave a consumer indifferent to the given gamble, assuming the chosen nightclub is equal to pocket cash.

To do this, we use the CES functional form measured in units of marginal utility of income, so functional form being

$$-x_1 \frac{(c_1 / x_1)^{1-\rho} - 1}{\rho - 1}.$$

We can then compute the utility gain from option A for each of the highlighted gambles for each value of the unobserved parameter x_1. The theory says this should be zero. We estimate the ρ's corresponding to the median and 85th percentile choices by minimizing the squared sum of these utility gains pooled across all of the gambles in the relevant cells. The results are shown in Table 3.

Table 4 – Estimated Relative Risk Aversion

	Pocket Cash x_1		
	\$20	\$40	\$80
ρ median	1.06	1.3	1.8
ρ 85th	2.1	2.8	4.3

Because the estimated ρ's are greater than 1, these preferences are not logarithmic. Notice that the data does not let us separately identify pocket cash and risk aversion; various combinations of these two are observationally equivalent.

As a check on the estimation procedure, we can compute for each of the estimated risk preferences and each scale of gamble (1X, 20X, 50X, 90X) the unique probability of reward that makes the individual indifferent between option A and option B. Tables 4 and 5 report these indifference points along with the indifference points in the data. That is, the actual indifference points reported in Table 4 are those from Table 2 at which about half the population chose A, and the actual indifference points in Table 5 are those from Table 2 at which about 85% of the agents chose A. The theoretical indifference points are the probabilities which would make an individual with the given CES preference indifferent.

Table 5: Indifference Probabilities for Median Estimated Value of ρ

		Pocket cash and corresponding median estimated ρ		
		\$20,1.06	\$40,1.3	\$80,1.8
	Actual indifference point	Theoretical indifference points		
1X	.60	.47	.46	.46
20X	.70	.65	.62	.60
50X	.70	.72	.71	.71
90X	.80	.79	.79	.81

Table 6: Indifference Probabilities for 85th Percentile Estimated Value of ρ

		Pocket cash and corresponding median estimated ρ		
	Actual Indifference point	Theoretical indifference points		
		\$20, 2.1	\$40, 2.8	\$80, 4.3
1X	.70	.50	.48	.47
20X	.80	.81	.79	.78
50X	.90	.90	.91	.94

If the CES model fit the data perfectly then in each row the theoretical probabilities corresponding to different levels of pocket cash would be identical to the actual probabilities. For the 20X and above treatments the fit is quite good. However, the 1X treatments fit less well, suggesting that for very small gambles risk aversion is even greater than in the CES specification.[18]

[18] It is possible that the size of the choices might have been confounded with the order in which the choices were given. Harrison, Johnson, McInnes and Rutstrom [2005] find that corrected for order the impact of the size of the gamble is somewhat less than Holt and Laurie found, a point which Holt and Laurie [2005] concede is correct. The follow-on studies which focus on the order effects do not contain sufficient data for us to get the risk aversion estimates we need.

6. The Allais Paradox

We proceed next to examine the Allais paradox in the calibrated model. We assume that the choice in this (thought) experiment is completely unanticipated. In this case the solution is simple: there is no self-control problem at the bank, so the choices is $c_1^* = x_1$ and spend all the pocket cash in the nightclub of choice. Given this, the problem is purely logarithmic, so the solution is to choose $x_1 = (1-\delta)w_1$.

In the Kahneman and Tversky [1979] version of the Allais Paradox the two options in the first scenario are A_1 given by $(.01:0,.66:2400,.33:2500)$, while B_1 is 2400 for certain. Many people choose option B_1. In scenario two the pair of choices are $A_2 = (.33:0,.34:2400,.33:2500)$ and $B_2 = (.32:0,.68:2400)$.[19] Here many people choose A_2. Expected utility theory requires the same option A or B be chosen in both scenarios.

To describe the procedure we will use for reporting calibrations concerning choices between pairs of gambles, let us examine in some detail the choice between A_1 and B_1 in the base case where the annual interest rate $r = 1\%$, annual income is \$28,000, wealth is \$860,000, so pocket cash and the chosen nightclub are $x_1 = c_1^* = 40$. Recall the cost of self-control $g(v_1) = \gamma v_1 + (1/2)\Gamma {v_1}^2$. Consider first the case $\Gamma = 0$ of linear cost of self-control. Here we have an expected utility model, so the optimal choice is independent of the scenario, and we can solve for the numerically unique value ${\gamma_1}^*$

[19] These were thought experiments. We are unaware of data from real experiments where subjects are paid over \$2000, though experiments with similar "real stakes" are sometimes conducted in poor countries. There is experimental data on the Allais paradox with real, but much smaller, stakes, most notably Battalio, Kagel and Jiranyakul [1990]. Even for these very small stakes, subjects did exhibit the Allais paradox, and even the reverse Allais paradox. The theory here cannot explain the Allais paradox over such small amounts, as to exhibit the paradox, the prizes must be in the region of the threshold $\mu_1(\gamma_1)$, while the prizes in these experiments ranged from \$0.12 to \$18.00, far out of this range. However, indifference or near indifference may be a key factor in the reported results. In set 1 and set 2 the two lotteries have exactly the same expected value, and the difference between the large and small prize is at most \$8.00, and there was only one chance in fifteen that the decision would actually be implemented. So it is easy to imagine that subjects did not invest too much time and effort into these decisions. By way of contrast Harrison [1994] found that with various small stakes the Allais paradox was sensitive to using real rather than hypothetical payoffs, and found in the real payoff case only 15% of the population exhibited the paradox. Although Colin Camerer pointed out the drop from 35% when payoffs were hypothetical was not statistically significant, a follow study by Burke, Carter, Gominiak and Ohl [1996] found a statistically significant drop from 36% to 8%. Conlisk [1989] also finds little evidence of an Allais paradox when the stakes are small. He examines payoffs on the order of \$10, much less than our threshold values of $\mu_1(\gamma_1)$ of roughly 1% of annual income. These studies suggest that when played for small real stakes there is no Allais paradox, as our theory predicts.

($\mu_1(\gamma_1{}^*) = 9.60$) such that there is indifference between the two gambles A and *B*.[20] A numerical computation shows that $Eu(\tilde{c}_1^B(\gamma_1{}^*)) > Eu(\tilde{c}_1^A(\gamma_1{}^*))$, so that when γ_1^* is chosen so that the long-run self is indifferent, the short-run self prefers the sure outcome *B*. On the other hand, when there is no cost of self-control, it is easy to compute that the long-run self prefers the risky outcome A.

Next, suppose that $\Gamma > 0$. Suppose we have solved the optimization problem as described by Theorem 1. As before, let $\overline{u}_1$ be the temptation utility. This is calculated by letting the short-run self choose the preferred lottery B and spend the entire proceeds.

We compute two possible values of the marginal cost of self-control separately depending on whether option A or option B is chosen by the long-run self.

$$\begin{aligned} \gamma_1{}^A &= \gamma + \Gamma\left(\overline{u}_1 - Eu(\tilde{c}_1^A(\gamma_1{}^A))\right) \\ \gamma_1{}^B &= \gamma + \Gamma\left(\overline{u}_1 - Eu(\tilde{c}_1^B(\gamma_1{}^B))\right) \end{aligned} \tag{6.1}$$

Suppose we start with γ slightly smaller than γ_1^* and $\Gamma = 0$. Then in both scenarios the risky option is strictly preferred. If we increase Γ slightly then in the high temptation scenario 1, γ_1 will rise more than in the low temptation scenario 2, creating the possibility that we will get a reversal in the high temptation scenario without creating a reversal in the low temptation scenario. This is exactly the Allais paradox.

To verify that this construction works, we computed γ_1^* for each of our cases, then constructed valued of γ, Γ with γ close to $\gamma_1{}^*$ and solved (6.1) iteratively to find in scenario 1 $\gamma_1^A[1], \gamma_1^B[1]$ and in scenario 2 $\gamma_1^A[2], \gamma_1^B[2]$. We then verified that in fact *B* is preferred in scenario 1 and *A* is preferred in scenario 2. The parameters used are reported in Table 7.[21]

---

[20] We omit the subscript on *A* and *B* since here the independence axiom is satisfied, so it does not matter which pair of choices we consider. Subsequently when we add some curvature the indifference will be broken, and, as we shall see, in opposite ways for the first and second pair of choices.

[21] The programs used for the computations were in Octave, a free equivalent of Matlab. They can be found at www.dklevine.com/papers/allais.zip.

Table 7 Explaining the Allais Paradox with $r = 1\%$

income	$x_1 = c_1{}^*$	ρ	$\mu_1(\gamma_1{}^*)$	γ	Γ	$\mu_1(\gamma_1^A[1])$	$\mu_1(\gamma_1^B[1])$	$\mu_1(\gamma_1^A[2])$	$\mu_1(\gamma_1^B[2])$
14000	20	1.06	19.41	21.08	1	19.69	19.67	19.33	19.31
14000	20	2.10	7.20	59.08	32	7.22	7.20	7.16	7.15
28000	40	1.30	9.61	17.04	1.5	9.68	9.68	9.55	9.54
28000	40	2.80	4.04	46.37	64	4.06	4.05	4.03	4.02
56000	80	1.80	4.80	15.02	4	4.85	4.84	4.80	4.79
56000	80	4.20	2.45	39.96	128	2.46	2.45	2.45	2.44

Notice that for each case there is a wide range of parameters that will generate paradoxes. The basic limitation is that if the curvature Γ is very large, then it will be impossible to generate values of γ_1^A, γ_1^B that are sufficiently close to γ_1^* to give

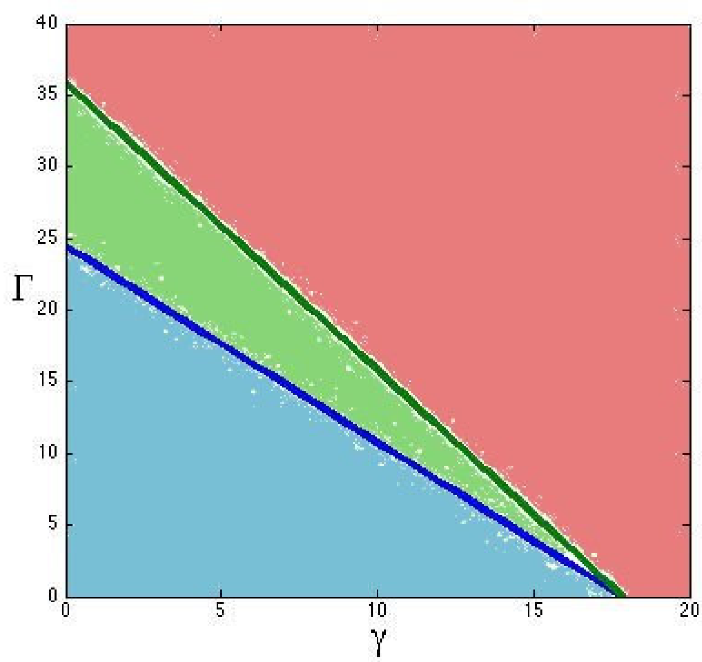


Figure 2 - Allais Paradoxes

paradoxes. This is shown in Figure 2, where the parameter values generating Allais paradoxes are computed for the base case of $28,000 annual income and risk aversion

$\rho = 1.3$. In the blue shaded blue region with low costs of self-control, the long-run optimum *A* is the best choice in both scenarios. In the red shaded region with high costs of self-control, the short-run optimum *B,* is the best choice in both scenarios. In the green shaded region in between an Allais reversal occurs, as the optimal choice is *B* in the high temptation scenario and *A* in the low temptation scenario.

The comparative statics are driven by how the marginal cost of self-control must change to maintain indifference between the two gambles in the face of changes in the marginal utilities of current and future consumption. Recall that the short-run self prefers the safe gamble *B,* while, starting in period 2 the long-run self prefers the risky gamble *A*. If the marginal utility of current consumption increases relative to future consumption, this will break the tie in favor of the earlier period, that is, the short-run self. However, lowering the marginal cost of self-control effectively lowers the weight on the short-run self's preferences, and restores the tie. The short version: more weight on the present implies the marginal cost of self-control must fall be lower in order to maintain indifference.

The first comparative static we consider is to change the interest rate. If we increase r to 3% or 5%, which increases the weight on the present, the marginal cost of self-control must fall to accommodate the paradox. However, in the calibration when we change r we also change wealth correspondingly. Changing r from 1% to 5% raises the weight on the present period by a little more than a factor of 5, but is lowers wealth by a factor of 5, and since second period value is logarithmic, raises the marginal utility of second period wealth by a factor of 5. The net effect is a very small increase in the weight on the present, and when we did the calculation, the values of $\mu_1(\gamma_1{}^*)$ change only in the third significant digit.

In contrast, raising risk aversion holding everything else fixed makes the gamble less attractive to the short-run self, increasing the temptation. This effectively increases the weight on the first period, so must lead to a reduced marginal cost of self-control, as happens in Table 7. Increasing income has a different effect: it has little effect on the decision problem, since that is formulated in relative terms. That means that the cutoff in dollars cannot change much, and so the cutoff relative to pocket cash, which has increased, must go down.

The values of the self-control parameter $\mu_1(\gamma_1{*})$ range from 2.44-19.4, which is considerably smaller than the 69.6 figure Abdel-Ghany et al [1983] found in CES data that we discussed above. However, the windfall income in the CES is considerably larger than these Allais gambles, so poses a greater temptation, and with increasing marginal cost of self-control should generate higher marginal self-control costs.

Notice that we are able to explain the Allais paradox with exactly the same risk aversion parameters that we used to explain the Rabin paradox. The theory here give a consistent explanation of both paradoxes, and it does so with a decision model that is consistent with long-run savings behavior being logarithmic as in growth and macroeconomic models.

Now we examine whether these parameters are consistent with the Keren and Roelsofsma [1995] data on hyperbolic discounting reported above. The non-linearity here is not sufficient to generate a reversal in the Keren and Roelsofsma experiment: Computation shows that even with the 50% chance of a prize, an individual with any of the parameters in the table above strictly prefers to take the money now. The reason is that our estimates of the linear coefficient γ are too large to support a reversal. If we use the same value of the curvature coefficient Γ as before, but a lower intercept, then a reversal is generated.[22] However, this lower value of the linear coefficient cannot explain the Allais paradox, because it reduces the temptation of the Allais gambles so much that the riskier gamble *A* is preferred in both scenarios. It might be possible to accommodate both the Allais choices and Keren and Roelsofsma data by departing from the quadratic specification of control costs, but we have not explored this possibility.

Original Allais Paradox: The original Allais paradox involved substantially higher stakes, so it would be difficult to implement other than as a thought experiment: option A_1 was $(.01:0,.89:1{,}000{,}000,.1:5{,}000{,}000)$ and B_1 was 1,000,000 for certain,; the second scenario was $A_2 = (.90:0,.10:5{,}000{,}000)$ and $B_2 = (.89:0,.11:1{,}000{,}000)$. Here the paradoxical choices were B_1 and A_2. In our base case of median income the results for the original Allais paradox are reported in Table 8.

[22] For example, with low income, high risk aversion, and an annual subjective interest rate of 3%, if the intercept is taken to be 4.4 rather than 26.4, and we use the estimated curvature of 23.9, a reversal is generated for the Keren and Roelsofsma [1995] data.

Table 8: Explaining the Original Allais Paradox with $r = 1\%$

income	$x_1 = c_1{*}$	ρ	$\mu_1(\gamma_1{*})$	γ	Γ	$\mu_1(\gamma_1^A[1])$	$\mu_1(\gamma_1^B[1])$	$\mu_1(\gamma_1^A[2])$	$\mu_1(\gamma_1^B[2])$
28000	40	1.3	8631.76	124351.02	131072	8768.31	8595.35	8512.30	8331.91
28000	40	2.8	115.87	571255.84	10747904	117.93	113.83	117.88	113.77

Notice that the values of $\mu_1(\gamma_1{*})$ are considerably larger here than they are for the Kahneman-Tverksy version of the paradox. This is as it should be: $\mu_1(\gamma_1{*})$ is endogenous and determined by temptation. The original Allais paradox involves larger stakes and thus larger temptations than the Kahneman-Tverksy version, so the theory predicts that the marginal cost of self-control should be larger. The value corresponding to low risk aversion, however, is implausible. It implies that the cutoff in dollars is about \$400,000, meaning that if the outcome is favorable (winning either \$1,000,000 or \$5,000,000) the long-run self intends to allow the short-run self to spend this amount on the first day. The value corresponding to high risk aversion is still large but more sensible: in case of success the long-run self will allow the short run self to spend \$5200 immediately. Since our model was calibrated on data with real payoffs, and that original paradox involves very large amounts that subjects may find difficult to evaluate, the discrepancy does not seem like a major concern.

7. Cognitive Load

The theory predicts that increasing cognitive load should increase the marginal cost of self-control and lead to reversals similar to those in the Allais paradox. Relatively few experiments have been conducted on the effect of cognitive load on decisions involving risk. One recent one is an experiment conducted with Chilean high school juniors by Benjamin, Brown and Shapiro [2006]. We analyze their data to show that their subjects have Allais-like reversals brought about by cognitive load as predicted by the theory.

In the experiment students made choices about uncertain outcomes both under normal circumstances and under the cognitive load of having to remember a seven-digit number while responding. In scenario 1 the choice was between a 50-50 gamble between 650 pesos and nothing versus a sure option of 250 pesos. In scenario 2 the sure option was replaced by a 50-50 gamble between 300 and 200 pesos.[23] The table below summarizes the fraction of the population taking the risky choice, with the number in parentheses following the treatment indicating the number of subjects.

Table 9 - Students Taking the Low Risk Option

650/0 versus 250		650/0 versus 300/200	
No load (13)	Cognitive Load (21)	No Load (15)	Cognitive Load (22)
70%	24%	73%	68%

These were real, and not hypothetical choices, the subjects were paid in cash at the end of the session. To provide some reference for these numbers, 1 \$US= 625 pesos and the subjects average weekly allowance was around 10,000 pesos from which they had to buy themselves lunch twice a week. [24]

The key fact in the table is that introducing cognitive load when the alternative is safe induces many subjects to switch to the safe alternative, while there is no such reversal when the "safe" alternative is the 300/200 gamble. This is as the theory predicts. If the short-run self prefers the safe alternative to the risky one we should see the first reversal. However, the 300/200 gamble is less tempting than the sure alternative of 250, so a cognitive load that will lead to a reversal in the first scenario need not do so in the second.

To calibrate the model, we take pocket cash to be the average weekly allowance of 10,000 pesos divided by 7 that amount in the daily case, or about \$2.29. We then work out wealth and income indirectly using the utility-function parameters that we calibrated in the Allais experiments.[25] To explain a preference reversal, the parameters must lead the two choices to have sufficiently similar levels of utility that self-control matters. Within the range of parameters in our calibration, the only set of parameters for which

[23] We thank the authors for providing us with this data. There is data on a risky alternative involving four other size prizes that are not relevant for our purposes. There is one anomaly in the data that we cannot explain: the fraction of people choosing the risky option against the sure alternative under cognitive load actually decreases as the size of the prize is increased. This may be due to sampling error.

[24] Many of them buy lunch at McDonald's for 2000 pesos twice a week, leaving an apparent disposable income of 6000 pesos per week.

[25] It is unclear that we should use the same value of τ but the results are not terribly sensitive to this.

this is true is when the annual interest rate is 7% and risk aversion is at the lower median level. For these parameters, we can calculate the values of $\mu_1(\gamma_1{*})$ that leads to indifference in the first and second scenario respectively.

Table 10 - Parameters for Indifference for the Chilean Gambles

r	income	w_1	$x_1 = c_1^*$	ρ	$\mu_1(\gamma_1{*})$ 1	$\mu_1(\gamma_1{*})$ 2
7%	1.6K	21K	2.29	1.06	23.91	23.95

Note that the values of $\mu_1(\gamma_1{*})$ of 24.66-24.71 needed to create indifference for the Chilean gambles are close to the value $\mu_1(\gamma_1{*}) = 19.2$ from the Allais paradox for the 5% calibration. This makes sense, because the temptations are of the same order of magnitude,.

In both scenarios, the risky option has the greater temptation, meaning that it will be chosen only for low marginal cost of self-control or equivalently, low values of $\gamma_1{*}$. The risky option, however, is preferred in the absence of cost of self-control. Recall that in our model the marginal cost of self-control is $\gamma + \Gamma(d_1 + \overline{u}_1 - u_1)$ where d measures the cognitive load. Suppose that $\mu_1(\gamma_1) < 23.90$ and that Γ is not too large. Then when cognitive load $d_1 = 0$, marginal cost of self-control is low enough in both scenarios that the risky alternative will be chosen. On the other hand, when cognitive load is high so $d_1 = \overline{d}_1 > 0$, for an appropriate value of $\overline{d}_1$, there will be a greater marginal cost of self-control $23.90 \leq \mu_1(\gamma_1) \leq 23.95$. That means that in scenario 1 the marginal cost of self-control is high enough that the safe alternative will be chosen, while in scenario 2 the marginal cost of self-control is low enough so that the risky alternative will continue to be chosen.

Note that the cognitive load calibration is more sensitive to the interest rate than the Allais calibration, as with cognitive load we require that r be at least 5%. In both cases, increasing the interest rate slightly increases the weight on the present relative to the future, due to the offsetting effect of changing wealth in the calibration. The intuition for this is that in the Allais case it is the levels of the marginal cost of self-control that matters, while in the cognitive load case what matters is the difference between two different marginal costs of self-control. The latter is much smaller than the absolute level, and so smallish increases in the level can result in largish increases (proportionally) in the difference.

In both cases, increasing the interest rate increases by a small (due to the offsetting effect of changing wealth in the calibration) amount the weight on the present relative to the future. In both cases this has the effect of slightly reducing the cost of self-control that leads to indifference. In the cognitive load calibration, there is less temptation in scenario 2 than scenario 1, meaning that for indifference $\mu_1(\gamma_1{}^*)$ is larger in scenario 2 than scenario 1, as required to explain the data. However, for $r = 1\%, 3\%$, the weight on the first period is so small that the algorithm is unable to find indifference. As we increase the weight on the early period the amount by which we must adjust self-control to maintain indifference for a given drop in temptation increases. At $r = 5\%$ the computer can find it, and the gap expands considerably as we increase the interest rate further. For interest rates higher than 5% – not implausible for high school students – we find that the range expands farther, making it more likely that cognitive load could push the marginal cost of self-control into the intermediate range needed to explain the data.

8. Making the Evening's Plans: Pocket Cash and Choice of Club

Our base model supposes that that the choice between A and B is completely unanticipated. How does the optimal choice of nightclub c_1^* and pocket cash change if the decision maker realizes that she will face a gamble? Specifically, let π^G denote the probability of getting the gambles. Our assumption has been that $\pi^G = 0$. In this case the solution is simple: there is no self-control problem at the bank, so the choices is $c_1^* = x_1$ and spend all the pocket cash in the nightclub of choice. Given this, the problem is purely logarithmic, so the solution is to choose $x_1 = (1-\delta)w_1$.

To examine the robustness of our results, consider then the polar opposite case in which $\pi^G = 1$, that is, the agent knows for certain she will be offered the choice between A and B. Since we will derive qualitative results only, we will simplify to the case $\tau = 1$. Given the choice of venue and pocket cash c_1^*, x_1 the choice of which lottery to choose at the nightclub and how much to spend are the same regardless of the beliefs that led to the choice of c_1^*, x_1. To keep things simple, we will assume that $\tilde{z}_1^k \ll \hat{z}_1$ so that all the proceeds of the gamble will be spent at the nightclub and x_1 will be chosen to be strictly positive.

In the Appendix we show that

Proposition 2*: First order conditions necessary for an optimum are*

$$(1-\delta)x_1 + \delta \frac{E(x_1+\tilde{z}_1^k)^{1-\rho}}{E(x_1+\tilde{z}_1^k)^{-\rho}} = (1-\delta)w_1 \tag{8.1}$$

$$c_1^* = \left(E\left(x_1+\tilde{z}_1^k\right)^{\rho-1}\right)^{1/(\rho-1)} \tag{8.2}$$

where k is the chosen alternative.

To understand (8.1), suppose that $\tilde{z}_1^k$ is constant, not random. Then (8.1) reduces to $x_1 + \delta E\tilde{z}_1^k = (1-\delta)w_1$. Here $E\tilde{z}_1^k$ does not substitute for pocket cash x_1 on a 1-1 basis, as it has a miniscule effect on life-time wealth, but as δ is nearly one, as we would expect it nearly does so. The second condition (8.2) then implies that c_1^* is chosen equal to the certain expenditure at the nightclub.

When the variance of $\tilde{z}_1^k$ is positive, observe that by assumption $\rho-1$ is positive so that $(\bullet)^{1/(\rho-1)}$ is increasing, and $(\bullet)^{\rho-1}$ is concave or convex as $\rho<2, \rho>2$. This implies

$$E(x_1+\tilde{z}_1^k)^{\rho-1} < \left[E(x_1+\tilde{z}_1^k)\right]^{\rho-1}$$

if $\rho<2$ with the inequality reversed if $\rho>2$. This in turn implies that $c_1^* < (>) E\left(x_1+\tilde{z}_1^k\right)$ as $\rho < (>) 2$. An individual with low risk aversion chooses a less attractive venue in the face of risk, and individual with high risk aversion chooses a more attractive venue.

9. Conclusion

We have argued that a simple self-control model with quadratic cost of self-control and logarithmic preferences can account quantitatively for both the Rabin and Allais paradoxes. We have argued also that the same model can account for risky decision making of Chilean high school students faced with differing cognitive loads. Ranges of income from half to double the median income; subjective interest rates in the range of 1-7%; short-run risk aversion in the range from 1-4; and a self-control cost switchpoint $\mu_1(\gamma_1{*})$ in the range 15-30 cover all of the cases. Except for the Chilean data, these results are quite robust. The Chilean students' behavior, however, require high subjective interest rates of 7%.

We find it remarkable that the behavior of Chilean high school students can be explained with essentially the same parameters that explain the Allais paradox. This finding is not trivial, as their possible observations that are not consistent with the theory. For example, cognitive load in the Chilean experiment could have caused subjects to switch in the reverse, "anti-Allais," direction, which we would not be able to explain. Second, there is enormous heterogeneity in the data; only a fraction of subject populations exhibit reversals, and the populations in the various experiments are very different, so there is no reason to believe that there is a single set of individual parameter values that will explain all of the data. However, while we have allowed the parameters to vary somewhat across experiments, it is important that all the parameters we use fall within a plausible range.

The main problem we have in calibrating the model is with respect to the degree of self-control. The model predicts that there should be a threshold level of unanticipated income, with marginal propensity to consume of 100% below the threshold and a very low marginal propensity above it. As we indicated, the permanent consumption data analyzed by Abdel-Ghany et al [1983] indicates that this may be true, and that the threshold is about 10% of annual income or 69.6 times pocket cash. We find, however, that to explain the data we consider, the threshold must be in the range of 4.04-22.3, which is considerably smaller than the threshold found in household consumption surveys. If we make the plausible assumption that windfall income measured in household consumption surveys is much less tempting than cash paid on the spot then this lower threshold makes more sense.

The existing model most widely used to explain a variety of paradoxes, including the Allais paradox, is prospect theory, which involves an endogenous reference point that is not explained within the theory.[26] In a sense, the dual-self theory here is similar to prospect theory in that it has a reference point, although in our theory the reference point is a particular value, pocket cash. They key aspect of pocket cash is that it is not arbitrary,

---

[26] See Kozegi and Rabin [2006] for one way to make the reference point endogenous, and Gul and Pesendorfer [2007] for a critique.

but is endogenous and depends in a specific way on the underlying preference parameters of the individual. The theories are also quite different in a number of respects. Prospect theory makes relatively *ad hoc* departures from the axioms of expected utility, while our departure is explained by underlying decision costs. Our theory violates the independence of irrelevant alternatives, with choices dependent on the menu from which choices are made, while prospect theory satisfies independence of irrelevant alternatives. Our theory can address issues such as the role of cognitive load and explains intertemporal paradoxes such as the hyperbolic discounting phenomenon and the Rabin paradox about which prospect theory is silent. Finally, a primary goal of our theory is to have a self-contained theory of intertemporal decision making; by way of contrast, it is not transparent how to embed prospect theory into an intertemporal model.[27]

In the other direction, prospect theory allows for individuals who are simultaneously risk averse in the gain domain and risk loving over losses. This is done in part through the use of different value functions in the gain and loss domains, and in part through its use of a probability weighting function, which can allow individuals to overweight rare events.[28] Most work on prospect theory has estimated a representative-agent model; Bruhin, Fehr-Duda, and Epper [2007] refined this approach by classifying individuals as expected utility maximizing or as prospect theory types,[29] and find that most individuals are prospect theory types. It is interesting to note that given the functional forms they estimate, individuals with expected utility preferences are assumed to be risk averse throughout the gains domain, while in their data individuals are risk loving for small probabilities of winning, while for higher probability of success they are risk averse. This can be explained within the expected utility paradigm by means of a Savage-style S-shaped utility function that is risk loving for small increases in income and risk averse for larger increases.[30]

[27] Kozegi and Rabin [2007] develop but do not calibrate a dynamic model of reference dependent choice.
[28] See Prelec [1998] for an axiomatic characterization of several probability weighting functions, and a discussion of their properties and implications.
[29] Their estimation procedure tests for and rejects the presence of additional types.
[30] Notice that it is possible to embed such short-run player preferences in our model although we have focused on the risk averse case. Indeed, such preferences are consistent even with long-run risk aversion: the envelope of S-shaped short-term utility functions can be concave provided that there is a kink between gains and losses, with strictly higher marginal utility in the loss domain. There is evidence that this is the case.

While S-shaped utility can explain risk seeking for small chances of gain and risk aversion for larger chances, it does not explain the Allais paradox, while prospect theory can potentially do so. But it appears that the parameters needed to explain individuals who are simultaneously risk averse and risk loving cannot at the same time explain the Allais paradox. Neilson and Stowe [2002] conducted a systematic examination of the parameters needed to fit prospect theory to various empirical facts, and concluded that

> parameterizations based on experimental results tend to be too extreme in their implications. The preference function estimated by Tversky and Kahneman (1992) implies an acceptable amount of risk seeking over unlikely gains and risk aversion over unlikely losses, but can accommodate neither the strongest choice patterns from Battalio, Kagel, and Jiranyakul (1990) nor the Allais paradox, and implies some rather large risk premia. The preference functions estimated by Camerer and Ho (1994) and Wu and Gonzalez (1996) imply virtually no risk seeking over unlikely gains and virtually no risk aversion over unlikely losses, so that individuals will purchase neither lottery tickets nor insurance…. We show that there are no parameter combinations that allow for both the desired gambling/insurance behavior and a series of choices made by a strong majority of subjects and reasonable risk premia. So, while the proposed functional forms might fit the experimental data well, they have poor out-of-sample performance.

The survey examined the original Allais paradox holding relative risk aversion constant, which as we have already noted is quite difficult because with expected utility individuals are not near indifference with reasonable degrees of risk aversion. However, if we use the Bruhin, Fehr-Duda, and Epper [2007] estimates from the Zurich 03 gains-domain treatment, the prospect theory types have preferences give by

$$U = \sum_i \frac{.846 p_i^{.414}}{846 p_i^{.414} + (1 - p_i)^{.414}} x_i^{1.056}$$

where p_i is the probability of winning the prize x_i.[31] In the Kahnemann and Tversky version of the Allais paradox, A_1 is $(.01:0,.66:2400,.33:2500)$, and B_1 is 2400 for

[31] Bruhin, Fehr-Duda, and Epper [2007] specified a utility function only for two-outcome gambles, this seems the natural extension to the three or more outcomes demanded to explain the Allais paradox. Note also that this utility function has the highly unlikely global property that if we fix the probabilities of the outcomes and vary the size of the rewards it exhibits strict risk loving behavior.

certain. This gives $U(A_1) = 3874.58$ and $U(B_1) = 3711$. In other words, an individual with these preferences would prefer A_1 to B_1 and so would not exhibit an Allais paradox.

Our overall summary, then, is that the dual-self model explains choices over lotteries about as well as prospect theory, while explaining phenomena such as commitment and cognitive load that prospect theory cannot. Moreover, the dual-self model is a fully dynamic model of intertemporal choice that is consistent with both traditional models of savings (long-run logarithmic preferences) and with the equity premium puzzle.[32]

In conclusion, there is no reason to think that the dual-self model has yet arrived at its best form, but its success in providing a unified explanation for a wide range of phenomena suggests that it should be viewed as a natural starting point for attempts to explain other sorts of departures from the predictions of the standard model of consumer choice. One possible next step would be try to more explicitly account for the evident heterogeneity of the population, and estimate distributions of self-control parameters as opposed to simply fitting the median or some other fractile as we have done here.

[32] The "behavioral life cycle model" of Shefrin and Thaler [1988] can also explain many qualitative features of observed savings behavior, and pocket cash in our model plays a role similar to that of "mental accounts" in theirs. The behavioral life cycle model takes the accounts as completely exogenous, and does not provide an explanation for preferences over lotteries. It does seem plausible to us that some forms of mental accounting do occur as a way of simplifying choice problems. In our view this ought to be derived from a model that combines the long-run/sort-run foundations of the dual-self model with a model of short-run player cognition.

References

Abdel-Ghany, M., G. E. Bivens, J. P. Keeler, and W. L. James, [1983], "Windfall Income and the Permanent Income Hypothesis: New Evidence," *Journal of Consumer Affairs* 17**:** 262–276.

Barberis, Huang and Thaler [2003] "Individual Preferences, Monetary Gambles, and the Equity Premium," NBER WP 9997.

Battalio, R. C., J. H. Kagel, and K. Jiranyakul [1990], "Testing Between Alternative Models of Choice and Uncertainty: Some Initial Results," *Journal of Risk and Uncertainty*, 3: 25-90.

Benjamin, D., S. Brown, and J. Shapiro [2006] "Who is Behavioral: Cognitive Ability and Anomalous Preferences, " mimeo.

Benhabib, J.and A. Bisin, [2005] "Modeling Internal Commitment Mechanisms and Self-Control: A Neuroeconomics Approach to Consumption-Saving Decisions." *Games and Economic Behavior* 52:460-492.

Bernheim, B. Douglas and A. Rangel [2004] "Addiction and Cue-Triggered Decision Processes." *American Economic Review*, 94*:*1558-1590.

Boldrin, M., L. Christiano and J. Fisher [2001], "Habit Persistence, Asset Returns, and the Business Cycle," *American Economic Review* 91:149-166.

Brocas, I. and J. Carrillo [2005] "The Brain as a Hierarchical Organization," mimeo.

Bruhin, A., H. Fehr-Duda, and T. Epper [2007] "Risk and Rationality: Uncovering Heterogeneity in Probability Distortion," mimeo.

Burke, M. S., J. R. Carter, R. D. Gominiak and, D. F. Ohl [2006], "An experimental note on the allais paradox and monetary incentives," *Empirical Economics*.

Chetty, R. and A Szeidl [2007] "Consumption Commitments and Risk Preferences," *Quarterly Journal of Economics* 122:831-877.

Conlisk, J. [1989], "Three Variants on the Allais Example," *American Economic Review*, 79: 393-407

Constantinides, G. [1990] "Habit Formation: A Resolution of the Equity Premium Puzzle," *Journal of Political Economy* 98:519-543.

Cox, J., V. Sadiraj, B. Vogt, and U. Dasgupta [2007], "Is There a Plausible Theory for Risky Decisions?" Georgia State University

Dekel, E. , Lipman, B., and Rustichini, Aldo. [2008] "Temptation-Driven Preferences." forthcoming in the *Review of Economic Studies.*

Frederick, S. [2005] "Cognitive Ability and Decision Making," *Journal of Economuc Perspectives,* 19:24-42.

FRBSF [2002], "What's Behind the Low U.S. Personal Saving Rate?" *FSRB Economic Letter*, 2002-09; March 29, 2002.

Fudenberg, D. and D.K. Levine [2006] "A Dual Self Model of Impulse Control," *American Economic Review* 96: 1449-1476.

Gabaix, X. and D. Laibson [2001], "The 6D Bias and the Equity Premium Puzzle", *NBER Macroeconomics Annual.*

Gailliot, M., R. F. Baumeister, C. N. DeWall, J. K. Maner, E. A. Plant, D. M. Tice, L. E. Brewer, [2008] "Self-Control Relies on Glucose as a Limited Energy Source: Willpower Is More Than a Metaphor," *Journal of Personality and Social Psychology*, forthcoming

Grossman, S. and G. Laroque [1990], "Asset Pricing and Optimal Portfolio Choice in the Presence of Illiquid Durable Consumption Goods," *Econometrica* 58: 25-52

Gul, F. and W. Pesendorfer [2001] "Temptation and Self Control." *Econometrica*, *69*:1403-1436.

Gul, F. and W. Pesendorfer [2005] "The Case for Mindless Economics," mimeo.

Halevy, Y. [2008], "Strotz Meets Allais: Diminishing Impatience and the Certainty Effect," *American Economic Review*, **98** 1145-1162.

Harrison, G. W. [1994], "Expected Utility Theory and the Experimentalists," *Empirical Economics*, 19: 223-243

Harrison, G. W., E. Johnson, M. M. McInnes, E. E. Rutström [2005], "Risk Aversion and Incentive Effects: Comment," *American Economic Review* 95:897-901.

Holt, C. and S. Laury [2002] "Risk Aversion and Incentive Effects," *American Economic Review* 92:1644-1655.

Holt, C. and S. Laury [2005], "Risk Aversion and Incentive Effects: New Data Without Order Effects" *American Economic Review* 95:902-904.

Imbens, G., D. Rubin, B. Sacerdote [2001] "Estimating the Effect of Unearned Income on Labor Earnings, Savings, and Consumption: Evidence from a Survey of Lottery Players," *American Economic Review* 91:778-794.

Kahneman, D. and A. Tversky [1979] "Prospect Theory: An Analysis of Decisions under Risk," *Econometrica* 47, 313-327.

Keren, G. and P. Roelsofsma [1995], "Immediacy and Certainty in Intertemporal Choice," *Organizational Behavior and Human Decision Making*, **63** 297-297.

Koszegi. B. and M. Rabin [2006] "A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics* 121:1133-1165.

Koszegi, B. and M. Rabin [2007] "Reference-Dependent Consumption Plans," mimeo.

Laibson, D. [1997] "Golden Eggs and Hyperbolic Discounting." *Quarterly Journal of Economics,* 112: 443-477.

Landsberger, M. [1966]: "Windfall Consumption: Comment", *American Economic Review*, 56: 534-540

Loewenstein, G. and T. O'Donoghue [2004]"Animal Spirits: Affective and Deliberative Processes in Economic Behavior," mimeo.

Muraven, M. and R. Baumeister [2000] "Self-Regulation and Depletion of Limited Resources: Does Self-Control Resemble a Muscle?" *Psychological Bulletin, 126* 247-259.

Muraven, M. M. D. Tice and R. Baumeister, [1998] "Self-Control as a Limited Resource." *Journal of Personality and Social Psychology, 74* . 774-789.

Neilson, W. and J. Stowe [2002] "A Further Examnation of Cumulative Prospect Theory Parameterizations," *Journal of Risk and Uncertainty* 24:31-46.

Ozdenoren, E., S. W. Salant and D. Silverman [2006], "Willpower and the Optimal Control of Visceral Urges" NBER W12278.

Postelwaite, A., L. Samuelson, and D. Silberman [2004] "Consumption Commitments and Preferences for Risk," mimeo.

Prelec, D. [1998] "The Probability Weighting Function," *Econometrica* 66:497-527.

Rabin, M. [2000] "Risk Aversion and Expected-Utility Theory: A Calibration Theorem." *Econometrica*, *68*:1281-1292.

Samuelson, P. [1963] "Risk and Uncertainty: A Fallacy of Large Numbers," *Scientia* 98:108-113.

Shefrin, H. and R. Thaler [1988] "The Behavioral Life-Cycle Hypothesis," *Economic Inquiry,* 26:609-643.

Shiv, B. and A. Fedorikhin [1999] "Heart and Mind in Conflict: The Interplay of Affect and Cognition in Consumer Decision Making." *Journal of Consumer Research*, *26:* 278-292.

Shiller, R. [1989]: *Market Volatility*, MIT Press: Cambridge.

Thaler, R. H. [1980]: "Towards a positive theory of consumer choice," *Journal of Economic Behavior and Organization*, 1:39-60

Ward, A. and T. Mann, Traci [2000] "Don't Mind If I Do: Disinhibited Eating Under Cognitive Load." *Journal of Personality and Social Psychology, 78*: 753-763.

Weber, B. and G. Chapman [2005], "The Combined Effects of Risk and Time on Choice: Does Uncertainty Eliminate the Immediacy Effect? Does Delay Eliminate the Certainty Effect?" *Organizational Behavior and Human Decision Processes*, **96**, 104-118

Appendix 1

Theorem 1*:* (a) *For given* (x_1, c_1^*) *and each* $j \in \{A, B\}$ *there is a unique solution to*

$$\gamma_1{}^j = \hat{\gamma}_1{}^j(\gamma_1{}^j).$$

This solution together with $\tilde{c}_1^j = \min(\hat{c}_1^j(\gamma_1{}^j)(z_1^J), x_1 + z_1^j\}$ *and the choice of* j *that maximizes long-run utility is necessary and sufficient for an optimal solution.*

Proof: Consider random unanticipated income $\tilde{z}_1^j$ at the nightclub. If z_1 is the realized income, the short-run self is constrained to consume $c_1 \leq x_1 + z_1$. Period 2 wealth is given by

$$w_2 = R(s_1 + x_1 + z_1 - c_1 - c_1^d) = R(w_1 + z_1 - c_1 - c_1^d) \ .$$

The utility of the long-run self starting in period 2 is given by the solution of the problem without self control, that is:

$$U_2(w_2) = \frac{\log(w_2)}{1-\delta} + K.$$

Let $\tilde{c}_1$ be the optimal response to the unanticipated income $\tilde{z}_1$. This is a random variable measurable with respect to $\tilde{z}_1$. The overall objective of the long-run self is to maximize

$$\tau\Big(Eu(\tilde{c}_1^j, c_1^*) - \overline{g}(x_1, \tilde{c}_1^j, c_1^*)\Big) + \frac{\delta}{(1-\delta)} E \log(w_1 + \tilde{z}_1^j - \tilde{c}_1^j - c_1^d) + K \ . \quad \text{(A.1)}$$

Let $\overline{u}_1(x_1, c_1^*) = \max\{Eu(x_1 + \tilde{z}_1^A, c_1^*), Eu(x_1 + \tilde{z}_1^B, c_1^*)\}$ denote the maximum possible utility given c_1^* and the pair of lotteries A,B. We then have that

$$\begin{aligned} &Eu(\tilde{c}_1^j, c_1^*) - \overline{g}(x_1, \tilde{c}_1^j, c_1^*) = \\ &Eu(\tilde{c}_1^j, c_1^*) - g\Big(\overline{u}_1(x_1, c_1^*) - Eu(\tilde{c}_1^j, c_1^*)\Big)' \end{aligned}$$

and since $\overline{u}_1$ does not depend on $\tilde{c}_1^j$, the optimal level of consumption can be determined for each lottery realization by pointwise maximization of (A.1) with respect to $c_1 = c_1^j(z_1^j)$. The marginal cost of self-control is given by

$$\gamma_1 = g'\Big(\overline{u}_1(x_1, c_1^*) - Eu(\tilde{c}_1^j, c_1^*)\Big) = g'\Big(\overline{u}_1(x_1, c_1^*) - \sum_{z_1^j} \Pr(z_1^j) u(z_1^j, c_1^*)\Big). \quad \text{(A.2)}$$

First we show that the first order conditions corresponding to optimal consumption for a given choice *j* have a unique solution. Observe that

$$\frac{d\gamma_1}{dc_1^j(z_1^j)} = -\Pr(z_1^j)u'(c_1^j(z_1^j), c_1^*)g''\Big(\overline{u}_1(x_1, c_1^*) - Eu(\tilde{c}_1^j, c_1^*)\Big) \leq 0 .$$

Because

$$\frac{\partial u(c_1^j, c_1^*)}{\partial c_1^j} = (c_1^j)^{-\rho}\left(c_1^*\right)^{\rho-1},$$

the derivative of (A.1) with respect to $c_1^j = c_1^j(z_1^j)$ evaluated at z_1^j as

$$\tau(1+\gamma_1)(c_1^*)^{\rho-1}(c_1^j)^{-\rho} - \frac{\delta}{(1-\delta)}\frac{1}{w_1 + z_1^j - c_1^j - c_1^d}$$

From this we can compute the second derivative

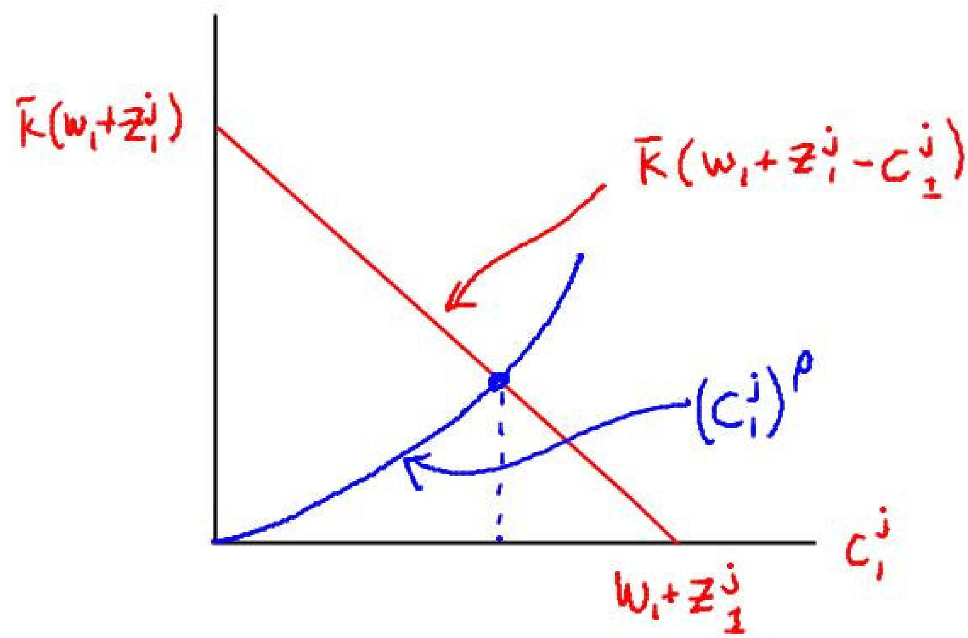
$$\tau(c_1^*)^{\rho-1}(c_1^j)^{-\rho}\frac{d\gamma_1}{dc_1^j} - \tau\rho(1+\gamma_1)\left(c_1^*\right)^{\rho-1}(c_1^j)^{-\rho-1} - \frac{\delta}{(1-\delta)}\frac{1}{\left(w_1 + z_1^j - c_1^j - c_1^d\right)^2} < 0$$

implying that the function is globally concave.

Because the objective function is globally concave with respect to c_1^j, it follows that the unique maximum is given by the solution to the first order condition, which may be written as

$$\begin{aligned}(c_1^j)^{\rho} &= \left(c_1^*\right)^{\rho-1}\frac{\tau(1-\delta)(1+\gamma_1)}{\delta}\left(w_1 + z_1^j - c_1^j - c_1^d\right) \\ &= \overline{K}\left(w_1 + z_1^j - c_1^j - c_1^d\right)\end{aligned}.$$

From the diagram below we can see both the uniqueness of the solution of the first order condition, and also see that the solution is increasing in $\bar{K}$, that is, decreasing in δ and



increasing in γ_1, and that the solution is increasing in $w_1 + z_1^j$.

We now show that the conditions in the Theorem are necessary and sufficient for an optimum. Examine necessity first. Suppose that an optimum exists. Once we know the choice j, for any given consumption plan in j the marginal cost of self control γ_1 is defined by A.2, and the optimal consumption plan must satisfy the first order condition with respect to that γ_1 because our conditions preclude a boundary solution. That is, ${\gamma_1}^j = {\hat{\gamma}_1}^j({\gamma_1}^j)$ must hold.

Next we show sufficiency. Suppose we have $j, {\gamma_1}^j$ satisfy the conditions of the theorem and that this is not the optimum. Since the problem is one of maximizing a continuous function over a compact space, an optimum exists. That optimum must yield more utility in (3.1) than choosing $-j$ and any consumption plan in $-j$, so the unique consumption plan that comes from solving ${\gamma_1}^{-j} = {\hat{\gamma}_1}^{-j}({\gamma_1}^{-j})$. Given that j is chosen, the optimal consumption is the unique solution of the first order condition. On the other hand, if $-j$ was chosen, we could do no better than the consumption plan defined by $\gamma_1{}^{-j} = \hat{\gamma}_1{}^{-j}(\gamma_1{}^{-j})$, and by assumption this is not as good as choosing j.

☑

Proposition 2: *When* $\tilde{z}_1^j \ll \hat{z}_1, j = A, B$ *the first order conditions necessary for an optimum are*

$$(1-\delta)x_1 + \delta \frac{E(x_1 + \tilde{z}_1^k)^{1-\rho}}{E(x_1 + \tilde{z}_1^k)^{-\rho}} = (1-\delta)w_1 \tag{8.1}$$

$$c_1^* = \left(E\left(x_1 + \tilde{z}_1^k \right)^{\rho-1} \right)^{1/(\rho-1)} \tag{8.2}$$

where k *is the chosen alternative.*

Proof: When $\tilde{z}_1^j \ll \hat{z}_1, j = A, B$ all income is spent at the nightclub so $\tilde{c}_1^j = x_1 + \tilde{z}_1^j$, and consequently the gamble most preferred by the short-run self is chosen. There is no self-control cost, so the objective function, given the choice of gamble of gamble j, is

$$Eu(\tilde{c}_1^j, c_1^*) + \frac{\delta}{(1-\delta)} E \log(w_1 - x_1) + K_2 \,. \tag{A.3}$$

The first order condition with respect to c_1^* is

$$\frac{1}{c_1^*} - (c_1^*)^{\rho-2} E\left(\tilde{c}_1^j \right)^{1-\rho} = 0$$

which solves to give (8.2).

The first order condition with respect to x_1 is

$$\left(c_1^* \right)^{\rho-1} E(x_1 + \tilde{z}_1^k)^{-\rho} - \frac{\delta}{1-\delta} \frac{1}{w_1 - x_1} = 0\,.$$

Substituting (8.2) and rearranging gives (8.1).

☑