The Effect of Affect on Economic and Strategic Decision Making*

Benjamin E. Hermalin†
University of California

Alice M. Isen‡
Cornell University

First Version: December 1998
This Version: December 5, 1999

Abstract

The standard economic model of decision making assumes a decision maker makes her choices to maximize her utility or happiness. Her *current* emotional state is not explicitly considered. Yet there is a large psychological literature that shows that current emotional state, in particular *positive affect*, has a significant effect on decision making. This paper offers a way to incorporate this insight from psychology into economic modeling. Moreover, this paper shows that this simple insight can parsimoniously explain a wide variety of behaviors.

Keywords: Affect, morale, emotion.
JEL: B41, D99, C70, C73, D81.

1 Introduction

A moment's introspection will convince most people that their decisions are influenced, in part, by their mood. For instance, the decisions we make when happy are not always the same as those we make when unhappy. Nor is this merely an impression: There is a large psychological literature based on experiments that finds a relationship between *affect*—what non-psychologists might call mood, emotions, or feelings—and decision making (see Isen, 1999, for a survey). In particular, this research shows that relatively small changes in positive affect—what a lay person might call happiness and what an economist might call utility—can markedly influence everyday thought processes and that such influence is a common occurrence. Economic modeling of decision making and game playing has, however, essentially ignored the role of affect. The purpose of this paper is to make amends. In particular, we seek to demonstrate that the addition of affect allows us to explain a wide variety of decisions and observed behaviors that are difficult to explain under the standard economic paradigm and to do so within a single, simple framework.

A common reaction by economists to the introduction of psychological insights into economics is that it means abandoning or relaxing the standard assumption of rationality. While it is true that many such attempts have had that flavor (see, e.g., discussions in Lewin, 1996; Rabin, 1998; Elster,

*Financial support from NSF Grant SBR-9616675 and the Willis H. Booth Professorship in Banking & Finance is grateful acknowledged. The authors also thank Matt Rabin, Miguel Villas-Boas, and seminar participants at UC Berkeley and Davis for helpful comments.

†Contact information. Phone: (510) 642-7575. E-mail: hermalin@haas.berkeley.edu. Full address: University of California / Walter A. Haas School of Business / 545 Student Services Bldg. #1900 / Berkeley, CA 94720-1900.

‡Contact information. Phone: (607) 255-4687. E-mail: ami4@cornell.edu. Full address: Johnson Graduate School of Management / 359 Sage Hall / Cornell University / Ithaca, NY 14853-6201.

1998), our approach does *not*. In particular, the actors in our models are completely rational—they make their decisions to maximize the (discounted) value of their utility flow. What distinguishes our approach from the traditional model of decision making is that we assume that current positive affect or utility influences preferences going forward. For instance, consistent with experimental evidence (Isen and Levin, 1972), positive affect tends to increase a person's willingness to aid others; that is, an increase in mood either increases an individual's pleasure from helping or lowers the psychic cost of helping. More generally, the happiness or utility level at the time of decision making affects preferences, which then affects the decision made.

In a one-shot setting, such a change in modeling assumptions would be difficult to distinguish from the more usual assumption of fixed preferences. Moreover, in a one-shot setting, why a person holds certain preferences over others is not, generally, an interesting question in economics. On the other hand, if we consider dynamic settings, then affect becomes much more important: Affect at the beginning of a period influences preferences, which determine decisions, which modify the affective state at the end of the period, which then becomes the relevant affect at the beginning of the next period, and so on. In other words, if $\mathbf{u}_t$ denotes affect (possibly a multi-dimensional variable) at the *end* of period t and $\mathbf{x}_t$ denotes a vector of decisions made in period t, then we have the dynamic:

$$\mathbf{u}_t = U\left(\mathbf{x}_t, \mathbf{u}_{t-1}\right),$$

where U is a function that recognizes that period-t preferences are determined, in part, by affect at the beginning of the period. As we will show, primarily through examples, such dynamics can explain interesting aspects of people's decision making and how they play certain games.

A nice feature of this model is that, although simple, it can encompass a wide range of behaviors. As we will demonstrate, it can, for instance, explain why employers want to hire "happy" workers, workers with "good attitudes," and why they want to take actions that boost morale. It can reconcile rational decision making with the apparent paradox of people eschewing behaviors that correlate with happiness (e.g., socializing, becoming sober, etc.). It offers insights into why moods tend to be persistent and why, for example, pharmacological intervention can be necessary in treating the depressed. It even explains why, in common-interest situations, players try to boost each others' morale and why, in opposing-interest situations, players try to demoralize each other. It also provides a *single* alternative explanation for behavior that has been explained by a wide variety of assumptions: for example, increased incentives from raising *fixed* wages (i.e., a result resembling efficiency wages), seemingly fair or cooperative play in *finitely* repeated games, and cooperative play with*out* punishment strategies in infinitely repeated games.

As Elster (1998) points out, reference to moods and other emotions in economics is rare. When such reference is made, it's usually to make sense of some behavior that seems inconsistent with narrow self interest. For example, honesty in situations where dishonesty would appear to have a larger payoff. If detection is possible, even if not assured, then honesty can be rationalized by assuming that it will be punished with sufficient severity to make the expected utility from being dishonest less than the utility from being honest. But there are many situations in which detection is impossible, or so unlikely, that even the most severe allowable punishment couldn't be a deterrent. In such cases, economists have typically "rationalized" honesty by appealing to the cost of guilt (see, e.g., Becker, 1976; Frank, 1988). Observe, however, that this approach considers only the emotional *consequences* of actions. Decision making is, thus, affected only by the *anticipation* of those consequences. In contrast, our model *also* has the reverse feedback: Having triggered certain emotions, those emotions will affect decision making going forward. That is, for instance, a guilty person will behave differently from a person who doesn't feel guilty (e.g., in search of atonement, the former may donate more to charity than the latter).

Two recent papers in economics (MacLeod, 1996; Kaufman, 1999) have, like us, worried about the effect of emotional state on decision making. They differ from us in that they are interested in modeling the adverse consequences of emotional state on cognitive abilities.[1] In contrast, our actors enjoy normal cognition, and we consider the effects of normal, everyday mild emotional states or feelings. To be sure, we are certainly sympathetic to the view that extreme emotional state can affect cognitive ability,[2] but is worth exploring how emotional state affects behavior without departing from the rational-actor paradigm. Moreover, there is a substantial body of evidence (see Isen, 1999) that at least some affective states (e.g., positive affect) influence behavior with*out* diminishing cognitive ability.[3]

Another strain of the economics literature focuses on "rationalizing" emotions; in particular, to explain why evolutionary forces may have produced them (see, e.g., Frank, 1988; Romer, 1999). Under the supposition, consistent with the fossil record on brain cases (see, e.g., Johanson, 1996), that our homonid predecessors had less cognitive ability than we do, the case can made that there was some advantage to "hardwiring" certain responses. For example, Romer notes that people (like rats) exhibit nausea aversion: If we suffer nausea—for whatever reason—within a short time after eating a particular food, we become averse to that food. For a species with limited cognitive ability, this would seem to be a good way to "learn" what foods are harmful. In contrast, although there is an obvious appeal to such evolutionary theorizing, we do not seek to explain why people have emotions. We take the existence of emotions as given. Our question is what do they influence when it comes to decision making?

The idea that decisions in one period can affect well-being in future periods is a well-known one in economics. The most common formulation of this is in consumption-savings models, where increasing the level of consumption today reduces possible consumption tomorrow. Such "choice-set" effects are absent here—the choice set remains constant over time in our models.[4] In addition, the intertemporal linkage in our model runs solely through affect. In particular, affect, $\mathbf{u}_t$, at time t is a sufficient statistic for predicting future affect levels. Among other implications, this means that there is no *direct* effect of an individual's past behaviors (decisions) on her future behavior: If consumption paths $\{\mathbf{x}_\tau\}_{\tau=1}^t$ and $\{\mathbf{x}'_\tau\}_{\tau=1}^t$ both get the individual to affect level $\mathbf{u}_t$, then behavior thereafter will be the same. Consequently, this paper differs from the habit-formation literature (see, e.g., §4.4 of von Auer, 1998, for a survey), which assumes that the present utility function

[1] MacLeod turns to emotions to justify his model of heuristic problem solving versus the standard optimization techniques that economists typically model decision makers as using. He argues, based on clinical observations of brain-damaged individuals reported in Damasio (1995), that people's heuristic problem-solving abilities are tied to their emotions. MacLeod does not, however, consider how *different* emotional states affect decisions, as we do. Kaufman, building on solid, but preliminary, work in psychology (e.g., Yerkes and Dodson, 1908), suggests that emotional state can enhance or inhibit cognitive function: People who are completely uninterested in a problem or who are panicked over it are less able to solve it (or solve it less efficiently or effectively) than people exhibiting less extreme emotions.

[2] See Ashby et al. (1999) for a hypothesis concerning the role of the neurotransmitter dopamine in tying affect to cognition.

[3] In a related vein, Laibson (1996) has sought to borrow from psychology to understand how preferences and choices can come to be sensitive to contextual variables.

[4] This isn't to suggest, however, that we can't conceive of affective state playing a role in determining the choice set, nor that our model wouldn't apply when the choice set varies over time (for whatever reason). In fact, some work already suggests that positive affect can increase the choice set (Kahn and Isen, 1993). A time-invariant choice set, however, makes more straightforward what the role of the affective state is.

takes past *consumption* as an argument.[5] Moreover, that literature is concerned with "rational addiction" primarily, whereas our approach has broader application.

The rest of the paper proceeds, in the next section, by presenting the basic model and analyzing two example applications. In Section 3, we move from a deterministic model to one with random shocks. In Section 4, we return to a deterministic set-up to explore how affect can affect the play of games. We conclude in Section 5.

2 A Model of Positive Affect & Decision Making

Consider the following model in which positive affect or utility level influences decision making. An individual begins period t with utility $u_{t-1} \in \mathbb{R}$ determined by her past experiences. In period t, she makes decisions $\mathbf{x}_t \in \mathcal{X}$, where $\mathbf{x}_t$ is a vector and $\mathcal{X}$ is the time-invariant feasible set. Let her utility at the end of the period, u_t, be

$$u_t = U\left(\mathbf{x}_t, u_{t-1}\right).$$

The basic behavioral implication of this formulation is captured by the following proposition:

Proposition 1 *Assume that a solution,* $\mathbf{x}^*(u)$*, exists for the program*

$$\max_{\mathbf{x}\in\mathcal{X}} U\left(\mathbf{x}, u\right) \tag{1}$$

for all possible u*. Assume, too, that,*

$$\text{if } u > u', \text{ then } U\left(\mathbf{x}, u\right) > U\left(\mathbf{x}, u'\right) \text{ for all } \mathbf{x} \in \mathcal{X}. \tag{2}$$

Then the solution to

$$\max_{\{\mathbf{x}_t\}_{t=1}^{T}} \sum_{t=1}^{T} \delta^t U\left(\mathbf{x}_t, u_{t-1}\right) \tag{3}$$

(where $\delta > 0$ *and less than one if* $T = \infty$*) is* $\mathbf{x}_t = \mathbf{x}^*\left(u_{t-1}\right)$*; that is, the discounted flow of utility is maximized by making the decisions that maximize each period's utility. Moreover, if* $u > u'$*, then* $U\left(\mathbf{x}^*\left(u\right), u\right) > U\left(\mathbf{x}^*\left(u'\right), u'\right)$*; that is, given rational decision making (i.e., in equilibrium), utility at the end of a period is an increasing function of utility at the beginning of the period.*

Proof. Since future utility is increasing in current utility and current decisions *directly* affect current utility only, maximizing current utility period by period must maximize (3). Hence, $\mathbf{x}^*\left(u_{t-1}\right)$ are the optimal decisions in period t. The last part of the proposition follows from revealed preference and the strict monotonicity of $U\left(\mathbf{x}, \cdot\right)$:

$$U\left(\mathbf{x}^*\left(u\right), u\right) \geq U\left(\mathbf{x}^*\left(u'\right), u\right) > U\left(\mathbf{x}^*\left(u'\right), u'\right).$$

■

[5] Admittedly, in some models current utility is isomorphic to past consumption (e.g., Benhabib and Day, 1981), in which case the approaches are similar—although the motivation is different—but in many contexts there is no isomorphism: Of two equally unhappy people, only one may consume heroin today because only he has consumed it in the past.

In our general analysis, we will maintain the assumptions that the program (1) has a solution for all u and that (2) holds (i.e., *all else equal*, utility at the end of a period is greater, the greater it is at the beginning of the period). In the specific examples, it is readily shown that these assumptions are met.

Observe that the relationship between current utility and past utility is monotonically increasing. This distinguishes our analysis from some related work by Benhabib and Day (1981), where U can be seen as a *non*-monotonic function of u_{t-1}.[6] Although this non-monotonicity yields interesting dynamics—including possibly chaotic dynamics—Benhabib and Day don't offer what we see as a compelling behavioral justification for their utility function.[7]

As a consequence of Proposition 1, utility is defined by the difference equation:

$$u_t = U\left[\mathbf{x}^*\left(u_{t-1}\right), u_{t-1}\right]. \tag{4}$$

Consider, now, a couple of examples.

Example 1 (Creativity & Cooperation): We assume now the decision is one-dimensional. Specifically, $x_t \in \mathbb{R}_+$. Let this choice denote some measure of work effort by an individual. It could, for instance, be some measure of help provided a co-worker; it could be a measure of creativity of thought; or it could just be some measure of effort. There is experimental evidence that positive affect can increase willingness to help (Isen and Levin, 1972); enhance creativity (Isen et al., 1987); and increase intrinsic motivation (Isen and Reeve, 1992). These results can, in turn, be captured by assuming

$$U\left(x_t, u_{t-1}\right) = x_t\sqrt{2\beta} - \frac{x_t^2}{2u_{t-1}} + \bar{u}\left(1-\beta\right), \tag{5}$$

where $\beta \in (0,1)$ determines the marginal benefit of effort and $\bar{u}$ is some constant. To keep the model from being pathological, assume $\bar{u} \geq 0$ and $u_0 > 0$, where u_0 is the individual's time 0 utility. Note that we've chosen to model the effect of positive affect as a reduction in the marginal cost of x; we could, however, equivalently model it as enhancing the marginal benefit of x. In this example, $x^*\left(u_{t-1}\right) = u_{t-1}\sqrt{2\beta}$. Consistent with the experimental evidence, $x^*\left(\cdot\right)$ is an increasing function. Equation (4) becomes

$$u_t = \beta u_{t-1} + \bar{u}\left(1-\beta\right).$$

The solution to this difference equation is

$$u_t = u_0\beta^t + \bar{u}\left(1-\beta^t\right). \tag{6}$$

As $t \to \infty$, $u_t \to \bar{u}$; that is, $\bar{u}$ is the long-run steady-state of this difference equation. Observe that, at any time t, u_t is the weighted average of the steady-state utility and the previous period's utility. Hence, utility is improving if $\bar{u} > u_{t-1}$, falling if $\bar{u} < u_{t-1}$, and unchanging if $\bar{u} = u_{t-1}$. Consequently, $\bar{u}$ is a stable fixed point of this difference equation.

Suppose that the individual in question is employed. Let the per-period benefit to her employer from x be vx, where $v > 0$. Observe that, *ceteris paribus*, the employer prefers to hire a "happier" individual, since the employer's per-period revenue is

$$v\sqrt{2\beta}\left[u_0\beta^{t-1} + \bar{u}\left(1-\beta^{t-1}\right)\right],$$

[6] Benhabib and Day actually assume that current utility equals $x_{1,t}^{g(\mathbf{x}_{t-1})}x_{2,t}^{1-g(\mathbf{x}_{t-1})}$, where $g\left(\cdot\right)$ is an increasing function of x_1. One could, however, make $g\left(\cdot\right)$ a function of previous utility, which would make their analysis more similar to ours.

[7] They suggest that an individual is choosing the amount of leisure to enjoy each period and that the greater the level of past leisure, the more leisure the individual desires today (e.g., the less the individual worked last period, the more vacation she desires today; conversely, the harder she worked last period, the less vacation she desires today).

which is clearly increasing in u_0. This corresponds to the well-known adage that "happy workers make good workers." If we also supposed that $\bar{u}$ denoted a steady-state utility of the individual—e.g., something akin to "attitude"—then we see, in part, why employers seek employees with good attitudes or other attributes associated with long-run positive affect. Relatedly, to the extent that the employer can undertake activities to raise u_0 or $\bar{u}$ or both (e.g., pay a signing bonus or ensure a pleasant working environment), he will have incentive to do so, since this yields him greater benefit.

Worker output can also be a function of incentives. Suppose, for instance, that $\bar{u}$ is $\gamma(w)$, where $\gamma' > 0$ and $\gamma'' < 0$ is a function relating the per-period wage, w, to utility. For convenience, assume infinite employment. Assume a constant discount factor of $\delta \in (0,1)$; i.e., assume an interest rate of $(1-\delta)/\delta$. Then the wage will be set to maximize

$$\sum_{t=0}^{\infty} \delta^t \left[v\sqrt{2\beta}\left[u_0\beta^t + \gamma(w)\left(1-\beta^t\right)\right] - w \right].$$

The solution is defined by

$$\gamma'(w^*) = \frac{1-\beta\delta}{\delta(1-\beta)\,v\sqrt{2\beta}}. \tag{7}$$

Note first that, although the employer is paying a fixed wage, it nevertheless has important incentive effects. In some ways, it is like an efficiency-wage story (see, e.g., Akerlof, 1982; Shapiro and Stiglitz, 1984), but differs in so far as it is not explicitly dependent on the threat of unemployment or the existence of an alternative employment sector. The two models could be more closely linked by assuming that $\gamma(w)$ is also a function of the unemployment rate and relative wages, similar to what Akerlof (1982) does.

Inverting the right-hand side of (7), we see the derivative of the inverse with respect to β is

$$\frac{1}{2}\delta v\sqrt{2}\frac{1-3\beta+\beta\delta+\beta^2\delta}{(1-\beta\delta)^2\sqrt{\beta}}.$$

This is positive for β less than

$$\frac{1}{2\delta}3-\delta-\sqrt{\left(9-10\delta+\delta^2\right)}$$

and negative for β greater than that, which means w^* is increasing in β for β less than that and is decreasing in β for β greater than that ($\gamma(\cdot)$, recall, is concave). In words, the wage is, at first, increasing in the worker's intrinsic motivation, $\sqrt{2\beta}$, and, then decreasing in it. This occurs because, when intrinsic motivation is low, utility is mostly a function of the wage. Hence, raising the wage has a big impact on utility. The impact of this, however, depends on intrinsic motivation. The greater the intrinsic motivation, the greater the return from inducing positive affect. Consequently, as intrinsic motivation rises, the marginal return to the employer from raising the wage is increasing, so he increases the wage. When, however, intrinsic motivation is high, utility is relatively insensitive to the wage. Consequently, there is less to be gained from a high wage, so the wage rate begins to fall with intrinsic motivation.

If x can be measured directly, then we could also consider paying a piece rate, s, as an incentive. Changing the model somewhat, suppose

$$U(x_t, u_{t-1}) = sx_t - \frac{x_t^2}{2u_{t-1}} + \tilde{u}.$$

Then $x^*(u_{t-1}) = su_{t-1}$ and the difference equation becomes

$$u_t = \frac{s^2 u_{t-1}}{2} + \tilde{u}.$$

Its solution is

$$\begin{aligned} u_t &= \left(\frac{s^2}{2}\right)^t u_0 + \left[1 - \left(\frac{s^2}{2}\right)^t\right] \frac{\tilde{u}}{1 - \left(\frac{s^2}{2}\right)} \\ &= \beta(s)^t u_0 + \left[1 - \beta(s)^t\right] \frac{\tilde{u}}{1 - \beta(s)} \end{aligned}$$

hence,

$$x_t = s\left(\beta(s)^{t-1} u_0 + \left[1 - \beta(s)^{t-1}\right] \frac{\tilde{u}}{1 - \beta(s)}\right)$$

The employer's time-0 problem is, thus,

$$\begin{aligned} & \max_s \sum_{t=0}^{\infty} \delta^t (v - s) s \left(\beta(s)^t u_0 + \left[1 - \beta(s)^t\right] \frac{\tilde{u}}{1 - \beta(s)}\right) \\ = & \max_s \frac{s(v-s)}{1-\delta} \left[\frac{(1-\delta)u_0 + \delta\tilde{u}}{1 - \beta(s)\delta}\right] \\ = & \max_s \frac{s(v-s)}{1-\delta} H(s) . \end{aligned}$$

We see that $H(s) > 0$ and

$$H'(s) = \frac{(1-\delta)u_0 + \delta\tilde{u}}{(1 - \beta(s)\delta)^2} \beta'(s) > 0.$$

Suppose $\delta = 0$—that is, only one period mattered or, equivalently, $U(x_t, u_{t-1})$ was standard and didn't depend on u_{t-1}—then $s^* = v/2$.[8] For $\delta > 0$ (i.e., in our model), $s^* > v/2$. Hence, in our model we should see stronger piece rates than the standard model: It pays to invest in a "happy" worker, which means raising the piece rate. Observe, as well, that s^* is independent of u_0;[9] that is, the piece rate does *not* depend on the level of initial affect—yet affect is still important for determining the piece rate through its dynamic effect.

As a second example,

Example 2 (Socializing & Sobriety): Again assume the decision is $x_t \in \mathbb{R}_+$. Let x_t denote energy or effort expended on some task. For instance, x_t could be effort at socializing with others; or energy spent keeping to a diet or staying sober; or some other effort similar to that considered in the previous example. Suppose that

$$U(x_t, u_{t-1}) = \phi(u_{t-1}) x_t - \frac{x_t^2}{2};$$

that is, here, utility at the beginning of a period modifies the marginal benefit of the action, x_t. Assume that $\phi(\cdot)$ is at least twice continuously differentiable, that $\phi(\cdot) \geq 0$, and that $\phi'(\cdot) > 0$. These assumptions reflect the idea that socializing is more pleasurable the happier one is, that a positive mood makes it more rewarding to keep to a diet or stay sober, or the other behavioral evidence cited in Example 1. Observe that $x^*(u_{t-1}) = \phi(u_{t-1})$. Hence, equation (4) becomes

$$u_t = \frac{\phi(u_{t-1})^2}{2}. \tag{8}$$

[8] To ensure that $\beta(s) < 1$, it's necessary that $v < 2\sqrt{2}$.

[9] Note that we need to assume that $u_0 > 0$ and $4 > 2\delta v^2$ for the model to make sense.

Unless $\phi(u) \propto \sqrt{u}$, this is a nonlinear difference equation. The second derivative of the right-hand side function is

$$\left[\phi'(u_{t-1})\right]^2 + \phi(u_{t-1})\,\phi''(u_{t-1}).$$

From this it follows that if $\phi(\cdot)$ is strictly concave—improving initial utility has a bigger impact on future utility when initial utility is small than when it's large—then this difference equation can be convex for low values of u_{t-1} and concave for high values of u_{t-1}. In turn, this means it is possible that the right-hand side function crosses the 45°-line three times (see Figure 1). This would be true, for instance, if

$$\phi(u) = \frac{\beta u}{u+1}, \tag{9}$$

where $\beta > \sqrt{8}$. The three points of crossing would then be 0, $\frac{1}{4}\beta^2 - 1 - \frac{1}{4}\beta\sqrt{\beta^2 - 8}$, and $\frac{1}{4}\beta^2 - 1 + \frac{1}{4}\beta\sqrt{\beta^2 - 8}$ (e.g., if $\beta = 3$, the points would be 0, $\frac{1}{2}$, and 2). Returning to the general case in which (8) crosses the 45°-line three times, each point of crossing is a fixed point. Only the first, $\hat{u}_1$, and third, $\hat{u}_3$, however, are stable: To the left of the second, $\hat{u}_2$, the process converges toward $\hat{u}_1$ and to the right of $\hat{u}_2$, the process converges toward $\hat{u}_3$. Some points about this model:

- Small differences in initial utility (level of positive affect) can lead to large differences in future utility: Consider two individuals, one with initial utility $\hat{u}_2 - \varepsilon$ and one with initial utility $\hat{u}_2 + \varepsilon$ (assume $\hat{u}_3 - \hat{u}_2 > \varepsilon > \hat{u}_2 - \hat{u}_1$). The former's utility will be constantly decreasing, while the latter's will be constantly increasing. For example, if $\phi(\cdot)$ is given by (9) with $\beta = 3$, then starting the two individuals at .499 and .501 respectively will lead to them having utilities of .206 and .802 respectively by $t = 20$. By $t = 43$, both will be within ε of their stable fixed points (0 and 2, respectively). Correspondingly, there will be increasing differences in behavior: Initially, they will expend .999 and 1.001 units of effort or energy, but, by $t = 20$, it will be .642 and 1.266, respectively (nearly a two-to-one margin).
- As indicated by the previous point, the population will tend to divide between the "very happy" and the "less happy." Moreover, this difference in affective state will tend to be persistent all else being equal.[10]
- There will be a strong correlation between behavior and affect (e.g., happy people socialize more), but the conventional causal inference will be wrong: People are *not* so much unhappy because they don't socialize or fail to keep to a diet or drink too much, rather they behave in these ways because they aren't happy. That is, although behavior affects affect, affect affects behavior and it is not, therefore, always possible to modify affect in a desired way through behavior. It is even possible that unhappy people themselves confuse correlation for causation: Mistakenly declaring that they would be happier if they socialized more, kept to their diets, stayed off the bottle, etc.[11]
- Recall that our decision makers are behaving optimally, there is *no* way for them to modify their behavior to achieve a better utility time path.[12] Hence, it would be wrong to blame the unhappiness of the "recluse" or failed dieter on his or her lack of effort or will power; and it would seem wrong, as well, to blame it on irrational behavior.

[10] This could, however, be interrupted by actions of others. For instance, harmful acts, even neglect, by others could be a shock to the dynamic system. We consider such shocks in the next section and strategic interactions with others in Section 4.

[11] The idea that people might not understand why they're unhappy (or even what would make them happy) is not implausible: If people were expert at understanding their own psychology, why would there be any market for psycho-therapists?

[12] Since we're considering a single-dimensional choice set of actions, we're abstracting from the possibility of actions

In terms of policy, this suggests that in an employment situation (i.e., one in which x is a measure of effort), the firm wants to identify workers with initial utility (happiness, attitude, etc.) greater than $\hat{u}_2$ or induce such a utility initially (e.g., by giving a signing bonus) and arrange conditions so that positive affect is not dispelled. A policy prescription for the recluse or the failed dieter might be to directly try to improve utility rather than focus on the deficient behavior. For instance, pharmacological or other intervention might be beneficial by directly enhancing mood (e.g., by affecting the amount of a neuro-transmitter like dopamine or serotonin). In extreme cases physicians may prescribe a mood elevating drug, and once u_t gets above $\hat{u}_2$, the pharmacological intervention could be discontinued.[13]

Returning to the general formulation, define $U^*(u) = U[\mathbf{x}^*(u), u]$. Assume that $u_0 \in \mathcal{U}$ and $U^* : \mathcal{U} \to \mathcal{U}$, where $\mathcal{U}$ is an interval in $\mathbb{R}$. Let $I_{\mathcal{U}}$ be the greatest lower bound on $\mathcal{U}$ and $S_{\mathcal{U}}$ be the least upper bound on $\mathcal{U}$; that is,

$$I_{\mathcal{U}} = \inf \mathcal{U} \text{ and } S_{\mathcal{U}} = \sup \mathcal{U}.$$

If $U^*(\cdot)$ has *no* stable fixed point, then, given that $U^*(\cdot)$ is increasing, u_t must trend, but never reach, either $I_{\mathcal{U}}$ or $S_{\mathcal{U}}$. Particularly if that limit is $+\infty$ or $-\infty$, such a dynamic would seem unrealistic: One's affective state is, ultimately, a function of physical processes in the brain and all physical processes are bounded. Assuming that $U^*(\cdot)$ is continuous and differentiable,[14] the following proposition establishes conditions under which this dynamic process must have a stable fixed point.

Proposition 2 *Assume that $U^* : \mathcal{U} \to \mathcal{U}$ is continuous and differentiable,[15] that $u_0 \in \mathcal{U}$, and that $\mathcal{U}$ is an interval in $\mathbb{R}$. Assume, in addition, that there is* no *u such that*

$$u - U^*(u) = 1 - U^{*\prime}(u) = 0. \tag{10}$$

Then the dynamic process that defines affect, $u_t = U^*(u_{t-1})$, possesses at least one stable fixed point in $\mathcal{U}$ if at least one assumption in column "A" holds and if at least one assumption in column "B" holds.

A	B
(i) $\lim_{u \downarrow I_{\mathcal{U}}} U^*(u) - u > \varepsilon_I > 0$	*(i)* $\lim_{u \uparrow S_{\mathcal{U}}} u - U^*(u) > \varepsilon_S > 0$
(ii) $I_{\mathcal{U}} \in \mathcal{U}$	*(ii)* $S_{\mathcal{U}} \in \mathcal{U}$

Where ε_I and ε_S are arbitrary positive constants.

on other dimensions (e.g., going to an enjoyable film or giving oneself a treat to self-induce positive affect). But the point carries over to a vector of activities. That is, "happy" people could tend to choose the vector $\mathbf{x}$, while "unhappy" people tend to choose the vector $\hat{\mathbf{x}}$, $\hat{\mathbf{x}} \neq \mathbf{x}$. Yet this difference is not the *cause* of happiness or unhappiness, but merely a correlate. In particular, the "therapy" of behaving like happy people would still be inappropriate.

[13]This is not inconsistent with actual medical practice. Informal discussions with physicians indicate that "accepted practice" for first-time treatment with selective serotonin reuptake inhibitors (SSRIs) is to put someone on them for 6 to 12 months and, then, wean him or her off them. For many patients this is sufficient (i.e., u_t is now greater than $\hat{u}_2$), and future medication is not necessary. Other patients cycle on and off them, suggesting that their brain chemistry or life experience is such that they are periodically and randomly thrown well below $\hat{u}_2$, necessitating intervention to escape. Of course, intervention by others (e.g., taking the person to an enjoyable film or giving her a treat) can also be effective in many cases.

[14]Note continuity is a sufficient, but not necessary, condition for a fixed point to exist. Almost all our analysis would carry through if $U^*(\cdot)$ were not continuous everywhere (although, since it is monotonic, it must be continuous almost everywhere). Indeed, there could be important threshold (discontinuous) phenomena. These, however, lie outside the scope of this paper.

[15]Since $U^*(\cdot)$ is strictly increasing by Proposition 1, it is continuous and differentiable almost everywhere. Whereas continuity *everywhere* is necessary for what follows, we could carry out the same analysis with it merely being differentiable almost everywhere. Assuming it's differentiable everywhere, however, simplifies the proof.

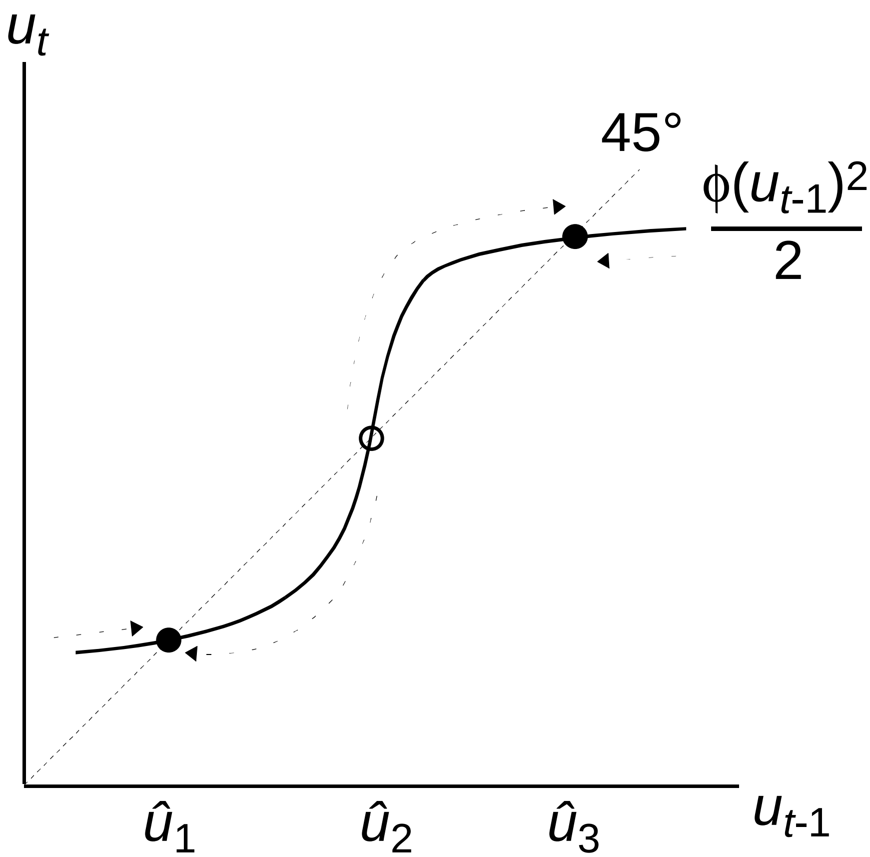

Figure 1: Possible relationship between u_{t-1} and u_t in Example 2

Proof: Please see Appendix.

Remark 1 *The menu eliminates the non-existence that would occur if*

$$[u - U^*(u)] \times [u' - U^*(u')] > 0$$

for all u, $u' \in \mathcal{U}$. *Condition (10) rules out the non-existence that could occur if every fixed point lay in an interval of fixed points (i.e., if* $U^*(\cdot)$ *coincided with the* $45°$ *line for an interval of u's). Alternatively, we could rule out this second type of non-existence by assuming that, for any* u *satisfying the equations in (10),* $U^{*\prime\prime}(u) = 0$ *and* $U^{*\prime\prime\prime}(u) < 0$ *(see Theorem 1.14 of Elaydi, 1996). Even without these assumptions, we would still be assured of a stable* bounded *range: For any* T, *there would exist a* bounded *set* $\mathcal{U}^*(T) \subset \mathcal{U}$ *such that* $u_t \in \mathcal{U}^*(T)$ *for all* $t > T$ *if at least one of the assumptions in column A held and at least one of the assumptions in column B held.*

3 Random Effects

In the model considered so far, utility follows a deterministic path. Moreover, this path is always monotonic: Utility either improves steadily over time or it declines steadily (although the pace of change slows as it approaches a stable fixed point). Formally,

Proposition 3 *Assume that* $U^* : \mathcal{U} \to \mathcal{U}$ *is continuous, that* $u_0 \in \mathcal{U}$, *and that* $\mathcal{U}$ *is an interval in* $\mathbb{R}$. *Then, in equilibrium, utility changes monotonically over time. Specifically, it is either non-decreasing or non-increasing (i.e.,* $(u_{t+1} - u_t)(u_t - u_{t-1}) \geq 0$ *for all* t*).*

Proof. If $u_t - u_{t-1} = 0$, then we're at a fixed point, so $u_{t+1} - u_t = 0$, which establishes the result. So assume, instead, that $|u_t - u_{t-1}| > 0$. We'll consider only the case $u_t - u_{t-1} > 0$. The case $u_t - u_{t-1} < 0$ is proved similarly. Since $U^*(\cdot)$ is continuous, there exists an open interval[16] in $\mathcal{U}$ containing u_{t-1} such that $U^*(u) > u$ for all u in that interval. Let (I, S) be the largest such interval. Now I is either a fixed point or it's $I_{\mathcal{U}}$. Similarly, S is either a fixed point or it's $S_{\mathcal{U}}$. Either way, since $U^*(\cdot)$ is increasing, $I \leq U^*(I) < U^*(S) \leq S$. Hence, $U^*[(I, S)] \subseteq (I, S)$. Consequently, u_t is also in this interval, implying $u_{t+1} = U^*(u_t) > u_t$. The result follows. ■

Taken literally, these monotonicity results (Proposition 3) are somewhat unrealistic. People's moods do not move monotonically: We *can* have good days, followed by bad days, followed by good days, and so forth. A more realistic model can be achieved by assuming that mood is subject to random events outside the decision maker's control (e.g., for many people, sunny days boost mood, while grey days lower it). Consequently, there is a stochastic aspect:

$$\tilde{u}_t = \hat{U}(\mathbf{x}_t, \tilde{u}_{t-1}, \zeta_{t-1}),$$

where ζ_{t-1} is a random shock. The $t - 1$ index reflects our interest in the impact of affect on decision making; that is, we assume that ζ is realized before $\mathbf{x}_t$ is chosen. Since we can always transform the random shock as necessary, we are thus free to write

$$\hat{U}(\mathbf{x}_t, \tilde{u}_{t-1}, \zeta_{t-1}) = U(\mathbf{x}_t, \tilde{u}_{t-1} + \zeta_{t-1}).$$

[16] Half open if $u_{t-1} = I_{\mathcal{U}}$ or $= S_{\mathcal{U}}$.

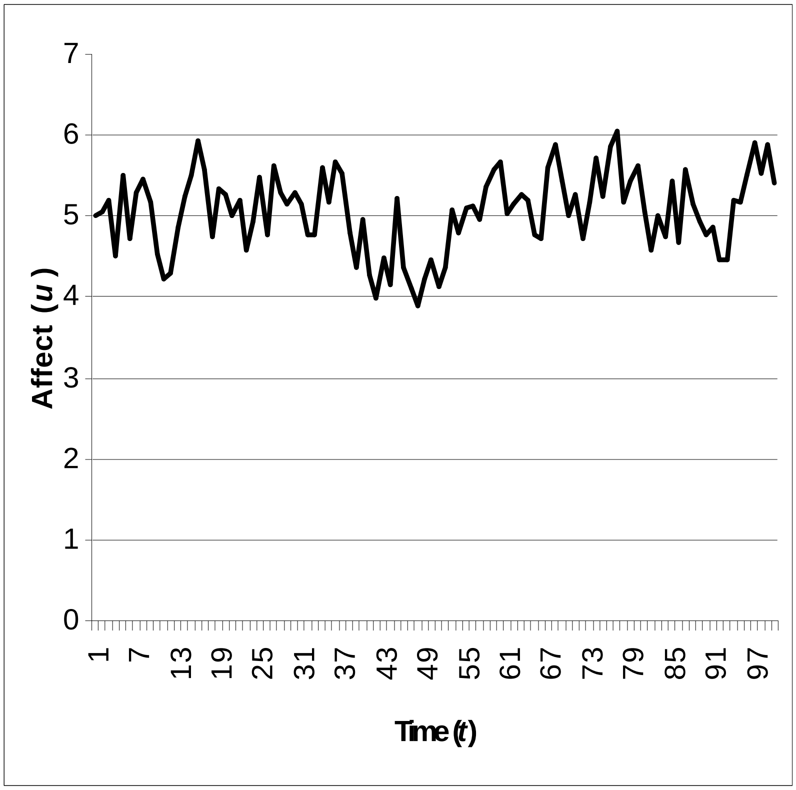

Figure 2: Random path of utility for Example 1 ($u_0 = 5$, $\beta = .5$, $\bar{u} = 5$, $\zeta \sim \mathrm{U}\left[-1.5, 1.5\right]$).

Note the dynamics are precisely the same if we write

$$u_t = U\left(\mathbf{x}_t, u_{t-1}\right) + \zeta_t,$$

where $u_t = \tilde{u}_t + \zeta_t$. Since this last formulation is the most convenient, it's the one we'll use in what follows. Provided ζ can be both positive and negative, then utility won't, in general, move monotonically (see, e.g., Figure 2, which is based on Example 1, where ζ_t is an i.i.d. draw from a uniform distribution on $\left[-\frac{3}{2}, \frac{3}{2}\right]$).

If we assume that $U\left(\cdot\right)$ is defined by expression (5), then

$$u_t = \beta u_{t-1} + \left(1 - \beta\right)\bar{u} + \zeta_t;$$

hence,

$$u_t = \beta^t u_0 + \left(1 - \beta^t\right)\bar{u} + \sum_{\tau=0}^{t-1} \beta^{t-\tau-1}\zeta_\tau.$$

Note that the direct impact of any random shock, ζ_τ, is diminishing over time: $d\beta^{t-\tau-1}/dt = \beta^{t-\tau-1}\ln\beta < 0$ (since $\beta < 1$). This makes sense: Mood-affecting random events last year likely have little impact on your mood today, whereas this morning's mood-affecting events probably do.

Since if ζ had a non-zero mean, we could build it into $\bar{u}$, there's no loss of generality in assuming that $\mathbb{E}\{\zeta_t\} = 0$. It follows then that

$$\mathbb{E}\{u_t\} = \beta^t u_0 + \left(1 - \beta^t\right)\bar{u};$$

that is, the non-stochastic model considered in Example 1 represents an unbiased predictor of the dynamics of the stochastic model. Finally, note that if $\zeta \in [\zeta_L, \zeta_H]$, then

$$\beta^t u_0 + \left(1 - \beta^t\right)\bar{u} + \zeta_L \frac{1-\beta^t}{1-\beta} \leq u_t \leq \beta^t u_0 + \left(1 - \beta^t\right)\bar{u} + \zeta_H \frac{1-\beta^t}{1-\beta};$$

that is, at any time there is a bounded neighborhood of $\bar{u}$, the non-stochastic stable fixed point, in which u_t must lie. Moreover, as time passes, that neighborhood depends less and less on initial utility.

Figure 3 illustrates a stochastic affect path for a variation on the model in Example 2. Here,

$$\phi(u) = \begin{cases} \frac{3u}{1+u}, \text{ if } u \geq 0 \\ 0, \text{ if } u < 0 \end{cases}.$$

In this model, there are three fixed points, 0, $\frac{1}{2}$, and 2, of which only the first and last are stable. Starting at the unstable fixed point, $\frac{1}{2}$, we see, in this example, that utility "bounces" around the low fixed point for awhile, then, due to a positive affect shock, jumps to hang around the high fixed point.

Figure 3 raises the question of whether stochastic utility can always "escape" oscillating around a low fixed point (e.g., 0) to reach a higher fixed point and, conversely, whether oscillations around a high fixed point can jump down towards a low fixed point. Here we will explore only the possibility rather than the likelihood of such escapes (the latter can be explored using the techniques in Freĭlin and Wentzell, 1984, for instance). Assume that the largest possible ζ is $z > 0$ and the smallest is $-z$.[17] Clearly, if z is large, then escape must be possible.[18] But what if z isn't large? To answer, consider the *deterministic* processes: $y_t = U^*(y_{t-1}) + z$ and $w_t = U^*(w_{t-1}) - z$ (where, again, $U^*(u) = U[\mathbf{x}^*(u), u]$). It follows that, at any time,

$$U^*(u_{t-1}) - z \leq u_t \leq U^*(u_{t-1}) + z.$$

Consequently, by examining the dynamic paths w_t and y_t, we can determine whether escape is possible. Figures 4a and 4b illustrate two possibilities based on the dynamics of Example 2. In both figures, u_t must lie between the bounds represented by the two dashed curves (the upper one corresponds to $U^*(u_{t-1}) + z$ and the lower one to $U^*(u_{t-1}) - z$). In Figure 4a, the lower bound crosses the 45° line once (from above). This means that it is *possible* for u_t to fall towards the lower stable fixed point of $U^*(\cdot)$. The absolute lower bound on u_t is defined by where $U^*(u_{t-1}) - z$ crosses the 45°. Note that the upper bound crosses the 45° three times (like $U^*(\cdot)$ itself). Since $\hat{u}_1^+$ is a stable fixed point of $U^*(u_{t-1}) + z$, it follows that u_t can never get above $\hat{u}_1^+$ if $u_\tau \leq \hat{u}_1^+$ for any $\tau < t$. In other words, although u_t can escape the higher stable fixed point in Figure 4a, it can*not* escape the lower stable fixed point (in *contrast* to the path shown in Figure 3). Moreover, u_t is "doomed" to eventually fall below $\hat{u}_1^+$ if $u_\tau \leq \hat{u}_2^+$ at some time τ. Figure 4b is essentially

[17] Symmetry is not necessary, but it simplifies the analysis.

[18] It can be shown that the z in Figure 3 is a large enough for u_t to escape any neighborhood around a stable fixed point of the deterministic process.

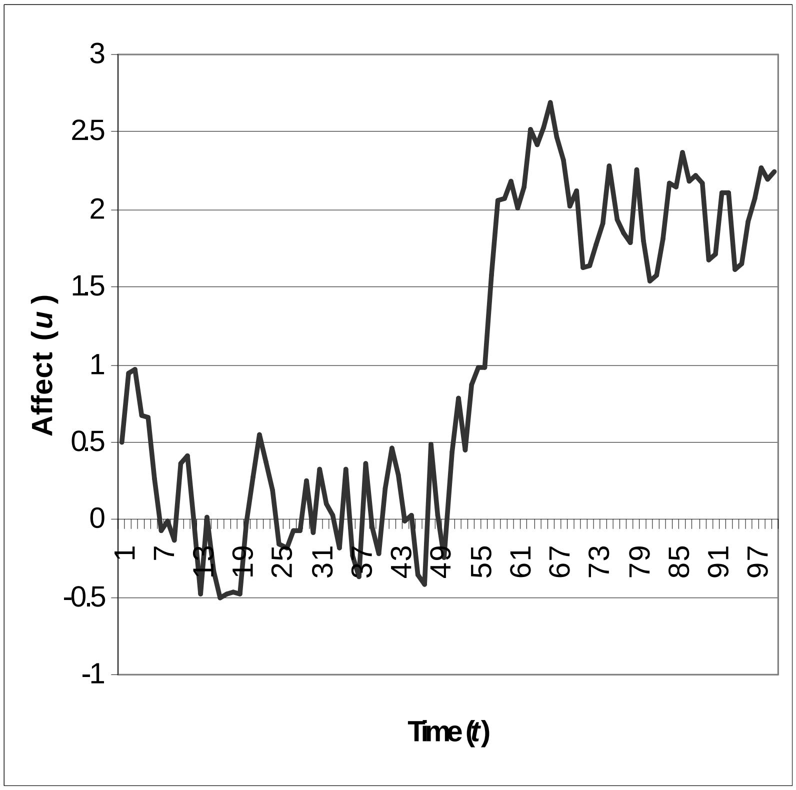

Figure 3: Utility (positive affect) path when U has three fixed points (0 and 2 are stable, .5 is unstable). The error, ζ, is distributed uniformly on $[-.5, .5]$.

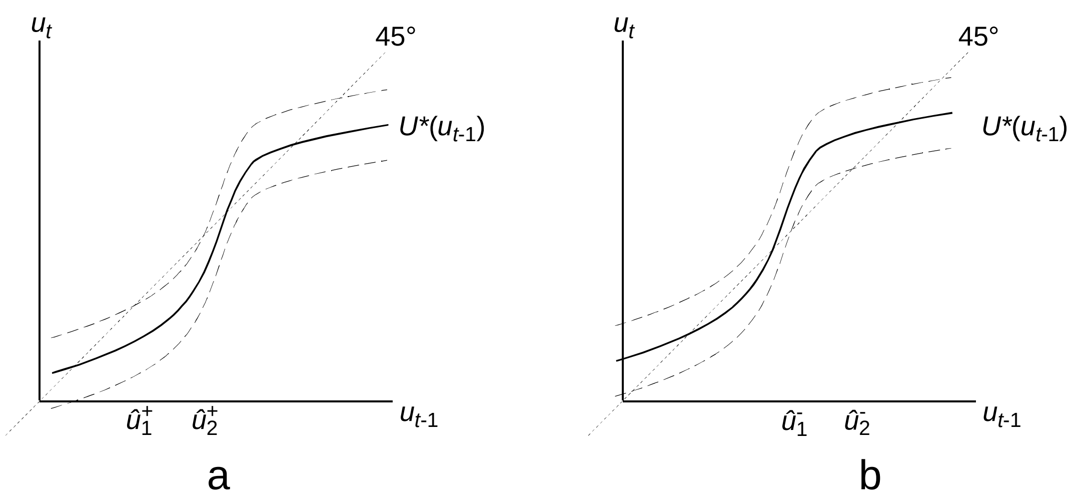

Figure 4: The stochastic utility process is bounded between the dashed curves.

the reverse scenario. Now u_t can escape the neighborhood of the lower stable fixed point of $U^*(\cdot)$, but it can't escape a neighborhood of the higher stable fixed point: Specifically, if $u_\tau \geq \hat{u}_2^-$, then $u_t \geq \hat{u}_2^-$ for all $t \geq \tau$. Moreover, u_t is "destined" to be high if $u_\tau \geq \hat{u}_1^-$ at some time τ. In both figures, if we shrunk z, the size of the bound, then there would exist neighborhoods around each stable fixed point of $U^*(\cdot)$ such that u_t could never escape that neighborhood. In other words, if the perturbations to utility are small, then conclusions similar to those reached in Example 2 will continue to hold.

It is worth noting that nothing in this analysis requires that ζ be an i.i.d. random variable. In particular, experience suggests that people tend to recall positive events more or with greater frequency than negative events. Hence, a $\zeta > 0$ leads to positive serial correlation in "shocks," as the recall of past positive shocks boosts positive affect. In contrast, a $\zeta < 0$ could have a much less serial correlation going forward. Indeed, there is no reason that u_t and ζ_t couldn't be described by complicated lag structures. We leave this issue, however, to future research.

To this point, we've assumed that the random effects are outside the control of the decision maker (e.g., they're due to weather, traffic, finding money, etc.). We could also consider decision making over gambles. In particular, there is evidence that individuals in whom positive affect has been induced behave in a more risk-averse fashion than a control group; i.e., those in a "neutral" affective state (see Isen et al., 1988, for evidence). This is not surprising if the dynamic process resembles the one in Figure 1: An individual above the unstable fixed point $\hat{u}_2$ faces a dire downside risk—she could get switched from trending up toward $\hat{u}_3$ to trending down toward $\hat{u}_1$—versus a modest upside potential. In contrast, an individual below $\hat{u}_2$ faces a sizeable upside potential—switching from trending down toward $\hat{u}_1$ to trending up toward $\hat{u}_3$—versus a modest downside risk. We are not, however, claiming that *all* the dynamic processes considered here will exhibit a positive correlation between utility and risk aversion—if the dynamic process is described, for instance, by equation (6), then attitudes toward risk will be independent of utility. Rather our point is that a model of decision making that is sensitive to the impact of affect can provide new insights into decision making under uncertainty and explain experimental results about such decision making.

4 The Impact of Positive Affect on Game Playing

We've so far considered only individual decision makers. In this section we extend our analysis to situations in which our decision makers interact (i.e., games). Although there are many potential models to explore, we will consider only one: There are two decision makers (players), indexed by *super*scripts. A given player's utility at the end of the tth period is assumed to be

$$u_t^i = x_t^i + \alpha^i x_t^j - \frac{\left(x_t^i\right)^2}{2u_{t-1}^i};$$

that is, each player i's utility is the same as in Example 1 (with $\beta = \frac{1}{2}$ and $\bar{u} = 0$) but with the addition of the impact of the *other* player's (j's) action on her utility. For simplicity, j's externality on i is assumed to enter linearly. If $\alpha^i > 0$, then it's a positive externality. If $\alpha^i < 0$, then it's negative. It seems reasonable—at least in many contexts—to imagine that the externality is not too large relative to the direct effect. Consequently, we limit attention to models in which $\left|\alpha^i\right| \leq 1$ for $i = 1, 2$.

To begin, consider the following *sequential* play game: Player 1 chooses her x in period 1, Player 2 chooses his x in period 2, and then the game ends. Given the finite time horizon, we may ignore discounting. Hence, the sums of the two periods' utilities are

$$U^1 = \alpha^1 x_2^2 + x_1^1 - \frac{\left(x_1^1\right)^2}{2u_0^1} \text{ and}$$
$$U^2 = x_2^2 - \frac{\left(x_2^2\right)^2}{2\alpha^2 x_1^1} + \alpha^2 x_1^1,$$

respectively. So that Player 2 has a well defined strategy, assume $\alpha^2 > 0$; Player 1 provides a positive externality for Player 2. Solving for the subgame perfect equilibrium, we see that Player 2's optimal play is

$$x_2^{2*}\left(x_1^1\right) = \alpha^2 x_1^1.$$

Hence, Player 1 maximizes

$$\alpha^1 \alpha^2 x_1^1 + x_1^1 - \frac{\left(x_1^1\right)^2}{2u_0^1}.$$

Solving yields

$$x_1^{1*}\left(u_0^1\right) = \left(1 + \alpha^1\alpha^2\right) u_0^1.$$

Some remarks: Observe that Player 1's behavior when there's a Player 2 is different than when there is no Player 2. Specifically, she does more x when $\alpha^1 > 0$ (Player 2's action provides a positive externality for her) and she does less when $\alpha^1 < 0$ (Player 2's action imposes a negative externality on her). When she does more (i.e., when $\alpha^1 > 0$), it might appear that she is "internalizing" the benefit she provides Player 2. This, however, would be an incorrect inference: She is playing selfishly, but she understands that doing more of something Player 2 likes benefits her because the happier is Player 2, the more Player 2 does of an action she likes. If Player 2 couldn't act or, equivalently, if $\alpha^1 = 0$, then Player 1 would cease to seem so generous. Consequently, although it can yield similar behavior, our model is different than an altruism model in which Player 2's utility would be an argument of Player 1's utility function. Note, too, that $x_2^{2*}(\cdot)$ is an increasing function. Particularly when $\alpha^1 > 0$, the resulting behavior of more x^1 leading to more x^2 could be viewed as Player 2 reciprocating Player 1's "generosity" or Player 2's behavior being governed by a fairness norm. Again, such a view would be, here, inaccurate: Player 2 doesn't produce more x^2 because he has a preference for reciprocating or being fair—in fact, he doesn't possess such preferences at all—rather he produces more because Player 1's action has changed his preferences such that he finds that he enjoys more of the action.

We are not arguing, however, that our approach is superior to assuming altruism or norms of fairness and reciprocity. Rather we are pointing out that behavior consistent with such motives need not be due to those motives (even in one-shot or finite-horizon games). Of course, these other approaches could be complementary to our approach: We could, for instance, assume that Player 2's utility is boosted from the "warm glow" that comes from knowing that Player 1 has done something extra for him (e.g., we could assume

$$\tilde{U}^2 = x_2^2 - \frac{\left(x_2^2\right)^2}{\kappa_0 + \kappa_1\left(x_1^1 - \bar{x}^1\right)} + \alpha^2 x_1^1,$$

where κ_0 and κ_1 are positive constants and $\bar{x}^1$ is the optimal level of x were Player 1 *not* playing with Player 2—here, u_0^1).

Consider now a two-period game with *simultaneous* moves. Now,

$$\begin{aligned} u_2^i &= \alpha^i x_2^j + x_2^i - \frac{\left(x_2^i\right)^2}{2u_1^i} \text{ and} \\ u_1^i &= \alpha^i x_1^j + x_1^i - \frac{\left(x_1^i\right)^2}{2u_0^i}. \end{aligned}$$

Solving for the subgame perfect equilibrium, we see that

$$x_2^{i*}\left(u_1^i\right) = u_1^i.$$

Hence, player i chooses x_1^i to maximize

$$\overbrace{\alpha^i \underbrace{\left(\alpha^j x_1^i + x_1^j - \frac{\left(x_1^j\right)^2}{2u_0^j}\right)}_{x_2^{j*}} + \frac{1}{2}\left(\alpha^i x_1^j + x_1^i - \frac{\left(x_1^i\right)^2}{2u_0^i}\right)}^{u_2^i} + \overbrace{\alpha^i x_1^j + x_1^i - \frac{\left(x_1^i\right)^2}{2u_0^i}}^{u_1^i}. \tag{11}$$

Thus,

$$x_1^{i*}\left(u_0^i\right) = \left(1 + \frac{2}{3}\alpha^i \alpha^j\right) u_0^i.$$

For future reference, note that x_t^{i*} is a linear function of i's $t-1$ utility only; in particular, the other state variable, u_{t-1}^j, is not relevant to her decision.

Define three scenarios:

Friendly game: $\alpha^1 > 0$ and $\alpha^2 > 0$;

Antagonistic game: $\alpha^1 < 0$ and $\alpha^2 < 0$; and

Mixed game: $\alpha^i \alpha^j \leq 0$, $i \neq j$.

Observe that, relative to a situation in which player i is alone, her choice of x_1^i is greater in a friendly or antagonistic game (since $\alpha^i \alpha^j > 0$), but smaller in a mixed game (since $\alpha^i \alpha^j < 0$). The intuition is straightforward: When $\alpha^i > 0$, player i wants to make player j happier, since that will yield a greater externality in the next period. Conversely, when $\alpha^i < 0$, she wants to reduce player j's happiness, since that will yield less of the externality in the next period. Player i boosts (reduces) player j's happiness by *increasing* x_1^1 when her action has a positive (negative) externality on j. She boosts (reduces) his happiness by *decreasing* x_1^1 when her action has a negative (positive) externality on him. This makes sense: Consider, for example, that, in many sporting events (examples of antagonistic games), the players seem to expend more energy or play with greater intensity in the first half than the second.[19] Similarly, as an example of a friendly game,

[19] Admittedly, physical fatigue also plays a role in explaining this pattern.

the home team tries to get its fans "into the game" early on and the fans tend to cheer a lot at the beginning (e.g., during player introductions). Note that if just *one* player is unaffected by the other's action, then *neither* player deviates from what he or she would have done if playing alone. That is, the apparent concern about the other player in the first period is strategic and not inherent (an insight also borne out by the fact that the last-period action is equal to the playing-alone action for both players regardless of scenario).

Substituting the definitions of x_1^{i*} and x_1^{j*} into player i's "lifetime" utility, expression (11), yields

$$U^i = \left[\alpha^i\alpha^j\left(1+\frac{2}{3}\alpha^i\alpha^j\right)+\frac{3}{2}\theta\left(1+\frac{2}{3}\alpha^i\alpha^j\right)\right]u_0^i+\alpha^i\left(1+\frac{2}{3}\alpha^i\alpha^j\right)\left(2-\frac{1}{3}\alpha^i\alpha^j\right)u_0^j,$$

where $\theta(q) = q - \frac{1}{2}q^2$. The sign of the last term equals the sign of α^i, since $\min 1 + \frac{2}{3}\alpha^i\alpha^j = \frac{1}{3}$ and $\min 2 - \frac{1}{3}\alpha^i\alpha^j = \frac{2}{3}$. Hence, when j's action provides a positive externality for i, she prefers that j have high initial utility. Conversely, when j's action imposes a negative externality on her, she prefers that j have low initial utility. This is consistent, for example, with the observation from sports that teams do better against demoralized opponents. It could also explain norms of sportsmanship and even why there are rules against directly trying to demoralize opponents (e.g., against excessive celebration after a touchdown).[20]

Finally, consider an infinite-horizon version of this simultaneous-move game. Let δ be the common discount factor. Once we consider an infinite horizon, a myriad of equilibria arise because of the players' abilities to reward cooperative play and punish uncooperative play. Since such equilibria would also emerge in a game with externalities but without moods affecting behavior, we won't explore such equilibria here. Instead, we will focus on a Markov equilibrium of the game. In a Markov equilibrium, play at period t can be conditioned only on the state variables, in this case u_{t-1}^1 and u_{t-1}^2. Moreover, if $\left(u_{t-1}^1, u_{t-1}^2\right) = \left(u_{\tau-1}^1, u_{\tau-1}^2\right)$ for any t and τ, then the equilibrium strategies of the players going forward from either t or τ must be the same. Note that Markov equilibria are also subgame perfect. In the Markov equilibrium of the game in which past utility didn't matter (e.g., one in which $u_t^i = \alpha^i x_t^j + x_t^i - \left(x_t^i\right)^2/2$), a player would play the same x each period and that x would maximize his utility with*out* regard for his opponent's utility. This will *not* be true in a Markov equilibrium of the game where past utility *does* matter.

Proposition 4 *There exists a Markov equilibrium of the infinite-horizon game in which*

$$x_t^{i*} = \mu^* u_{t-1}^i,$$

where

$$\begin{aligned}\mu^* &= 1-\frac{2^{\frac{1}{3}}Y}{3\delta Z}+\frac{Z}{3\delta\times 2^{\frac{1}{3}}},\\ Y &= 6\delta-3\delta^2-6\alpha^i\alpha^j\delta^2,\ \textit{and}\\ Z &= \left(54\alpha^i\alpha^j\delta^3+\sqrt{4Y^3+2916\left(\alpha^i\alpha^j\right)^2\delta^6}\right)^{\frac{1}{3}}.\end{aligned}$$

[20] Under current NCAA rules for American football, a team guilty of excessive celebrating is cited for unsportsmanlike play and penalized 15 yards on the next play. Note that the motivation for this penalty is unlikely to be (solely) the fact that celebrating delays play: There already exists a delay-of-game penalty (a five-yard penalty).

Proof: Please see Appendix.

Observe that if $\alpha^i = 0$ or $\alpha^j = 0$, then the equilibrium response constant, μ^*, equals one—the same value it would take if each player were playing in isolation (see Example 1). Consequently, as in the two-period game, if just one player is unaffected by the other's action, then *neither* player deviates from what he or she would have done if playing alone. It follows that what distinguishes this equilibrium from isolated play is $\alpha^i\alpha^j \neq 0$. The next proposition addresses how:

Proposition 5 *The equilibrium response constant, μ^*, is increasing in $\alpha^i\alpha^j$. Hence, if the game is friendly or antagonistic ($\alpha^i\alpha^j > 0$), then both players do more of the action conditional on their utility than they would if they played alone. Conversely, if the game is mixed ($\alpha^i\alpha^j < 0$), then both players do less of the action conditional on their utility than they would if they played alone.*

Proof. Observe that the right-hand side of equation (14) equals 1 if $\mu^* = 0$. Hence the right-hand side of (14) must cross the 45° line from above. Hence, since the right-hand side is increasing in $\alpha^i\alpha^j$, it follows that increasing $\alpha^i\alpha^j$ must increase the μ^* at which the right-hand side crosses the 45° line. ■

As noted, we would observe larger responses in both friendly and in antagonistic games. For the observer of a friendly game, a natural interpretation would be that the players are exploiting infinite repetition to sustain a cooperative outcome (i.e., that promotes the positive externality) or otherwise playing in some reciprocal fashion. In this case, that interpretation would, however, be wrong. Here cooperation is not a consequence of infinite repetition nor any other direct motive to reciprocate. The players appear to cooperate only because each player understands it's better to have a happy opponent than an unhappy opponent. For the observer of an antagonistic game, a natural interpretation—at least at the start of the game—is that the players are punishing each other for *not* cooperating (not doing less of the action in recognition of the negative externality). Again, this interpretation would be incorrect in this context. The players are not punishing so much as attempting to "demoralize" their opponents, since, now, it's better to have an *un*happy opponent than a happy opponent.

The equilibrium dynamics are

$$\begin{aligned} u_t^i &= \theta(\mu^*)\, u_{t-1}^i + \alpha^i \mu^* u_{t-1}^j \text{ and} \\ u_t^j &= \alpha^j \mu^* u_{t-1}^i + \theta(\mu^*)\, u_{t-1}^j. \end{aligned}$$

Define

$$\mathbf{u}_t = \begin{pmatrix} u_t^i \\ u_t^j \end{pmatrix} \text{ and } \mathbf{M} = \begin{pmatrix} \theta(\mu^*) & \alpha^i\mu^* \\ \alpha^j\mu^* & \theta(\mu^*) \end{pmatrix},$$

so

$$\mathbf{u}_t = \mathbf{M}\mathbf{u}_{t-1}.$$

It can be shown (see, e.g., Elaydi, 1996, §3.1) that

$$\mathbf{u}_t = \begin{pmatrix} \frac{1}{2}\left(\lambda_1^t + \lambda_2^t\right) & \frac{1}{2}\frac{\lambda_1^t - \lambda_2^t}{\sqrt{\alpha^i\alpha^j}}\alpha^i \\ \frac{1}{2}\frac{\lambda_1^t - \lambda_2^t}{\sqrt{\alpha^i\alpha^j}}\alpha^j & \frac{1}{2}\left(\lambda_1^t + \lambda_2^t\right) \end{pmatrix}\mathbf{u}_0,$$

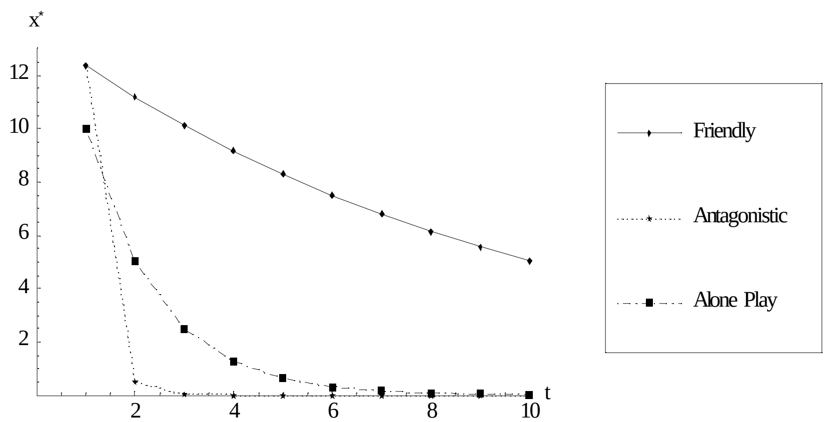

Figure 5: Example path for x_t^{i*} for friendly and antagonistic games plus for playing alone.

where $\lambda_h = \theta(\mu^*) + (-1)^{h-1} \mu^* \sqrt{\alpha^i \alpha^j}$ are the eigenvalues of $\mathbf{M}$.[21]

Restricting attention to the friendly and antagonistic cases, we see that

$$u_t^i > \frac{1}{2} \left(\lambda_1^t + \lambda_2^t \right) u_0^i$$

in the friendly case (since $\alpha^i > 0$), but

$$\frac{1}{2} \left(\lambda_1^t + \lambda_2^t \right) u_0^i > u_t^i$$

in the antagonistic case (since $\alpha^i < 0$). Hence, the players' utilities are always greater in a friendly game than in an antagonistic game. More importantly, consider the friendly game with α^i and α^j and the "polar" opposite antagonistic game $\tilde{\alpha}^i$ and $\tilde{\alpha}^j$, where $\tilde{\alpha} = -\alpha$. In these two games, μ^* would be the same. Yet, since $x_t = \mu^* u_{t-1}$, the actions would be less in every period (except the first) in the antagonistic game than in the friendly game. An observer might be tempted to interpret this as the players in the antagonistic game internalizing the negative externality (perhaps because of repeated play). This, however, wouldn't be correct: The players do less in the antagonistic game because they want to do less having been "demoralized" by their opponents. Figure 5 illustrates an example time path of x_t^{i*} (here, $\delta = .9$, $\alpha^i = \alpha^j = .35$, $\tilde{\alpha}^i = \tilde{\alpha}^j = -.35$, and $u_0^i = u_0^j = 10$). Observe that, in Figure 5, the level of action is initially higher in an antagonistic game than in isolated play, but then falls below it. Again, the tempting interpretation is that, after an initial "mistake," the players cooperate by doing less of the action because of the negative externality. And again, that's an incorrect interpretation in this model: The players do less, because the initial high level has succeeded in demoralizing them.

Finally, in a game, a given player's utility no longer needs to follow a monotonic path—see Figures 6. Hence, consistent with earlier discussion, strategic interactions can introduce non-monotonicity into an individual's utility time path.

[21] It might seem this wouldn't apply in the mixed case, where $\alpha^i \alpha^j < 0$. But working through with the resulting complex numbers, it can be shown that it also applies in the mixed case. We will, however, restrict attention to the cases in which $\alpha^i \alpha^j > 0$. In any case, note the model is only valid as long as u_t^i is not driven below zero. In the examples considered, this property is satisfied.

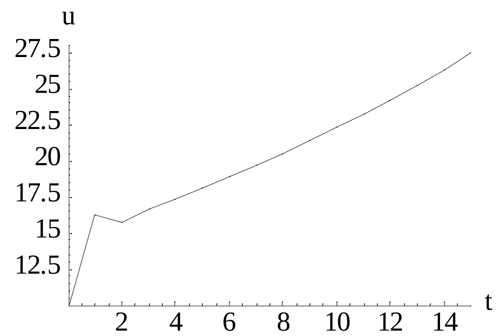

Figure 6: u_t^i in a friendly game where $\alpha^i = .86$, $\alpha^j = .23$, $\delta = .9$, and $u_0^i = u_0^j = 10$.

Admittedly, we've analyzed only a limited number of games in which affect could be relevant. Yet, our analysis gives some sense of the issues that arise. In particular, behavior that can be explained by altruism, fairness norms, reciprocity norms, or the exploitation of repeated play can also be explained by affect. In many ways, particularly for *finite* games, affect is more consistent with conventional models of rationality and more parsimonious: It doesn't require players obeying norms that aren't in their immediate self interest or taking someone else's utility as an argument in their own utility functions. Rather, using affect, players simply want to do what's best for them, but they recognize that the affect they induce in their opponents will, through interaction, feedback on them (see Batson, 1991, for empirical evidence that striving to increase others' moods can be motivated by self-interest). Moreover, affect models in games build directly on behavior in single-player decision problems in a way that these other approaches don't. Finally, affect models offer an explanation within the rational-actor paradigm for such behaviors in strategic situations of trying to demoralize your rivals or cheering on your allies, phenomena for which these other approaches don't account.

Clearly, our approach can be applied to a much larger set of games. Moreover, as we've suggested above, it could serve to complement other approaches. People, for instance, may strive to be fair or reciprocate in cooperative situations because they've found, as a rule of thumb, that the positive affect it induces in others, and consequently the level of others' actions, ultimately pays off for them.

5 Conclusions

In this paper, we have shown that incorporating the psychological finding that affect influences decision making can greatly enrich rational-actor models of decision making and strategic interaction. Although a modest change in our standard assumptions—yet possessing strong empirical backing (e.g., Isen, 1999)—it nevertheless gives insights into a number of behavioral phenomena:

- the persistence of mood, especially a happy mood;
- increased incentives from increases in a *fixed* wage (similar to an efficiency-wage story);
- the setting of piece rates;

- the paying of signing bonuses and other *non*-contingent rewards that serve to boost worker morale;

- the apparent paradox of people not pursuing behaviors correlated with well-being;

- pharmacological treatment strategies for depression;

- decision making under uncertainty;

- apparently cooperative play in finite-period games; and

- attempts to demoralize opponents and to build the morale of friends.

Moreover, we suspect that we've only scratched the surface with respect to the economic applications. We can, for instance, foresee applications to issues of morale building within organizations, promotion of products, building of customer loyalty, relationship marketing, and policy issues and social welfare, among others.

Perhaps more importantly, the methodology outlined here can be extended to other emotions (we've focused on positive affect—happiness—because the experimental evidence gave us a clear guide as to the nature of the relationship between affect and behavior in that context). For instance, we might imagine that guilt affects behavior; perhaps according to a dynamic similar to

$$u_t = -g_{t-1}e^{-x_t} \text{ and } g_t = g_{t-1}e^{x_t}$$

where g_t is the level of guilt at time t and x_t is a guilt-inducing action ($x_t < 0$ is a guilt-reducing action).[22] For example, x_t could be the amount of money the individual spends on herself in period t (a negative x_t would, then, indicate spending on others). Although simple, this model does capture the ideas that guilt reduces utility and engaging in a guilt-inducing activity increases guilt going forward, but is pleasurable today. Other emotions could be similarly incorporated into models.

Both psychological theory, including evolution-based theorizing (e.g., Johnston, 1999), and empirical work (e.g., Isen, 1999; LeDoux, 1998)—to say nothing of common sense—make it clear that behavior is affected by emotions. In addition, increasingly work in neuro-science (e.g., Damasio, 1995; Ashby et al., 1999; LeDoux, 1998) is working out the links and underpinnings between feelings and behavior. Further, clinical evidence from those who've suffered certain brain injuries demonstrates a rather suggestive set of correlations between behavior and emotion. Given all this, it seems that economic modeling of behavior should pay attention to emotions. Otherwise, our models will be better suited to Mr. Spock and his fellow Vulcans than to *homo sapiens*. On the other hand, one of the amazing things about our species is the ability to employ rational thought, including planning. Consequently, we've sought to develop a model that integrates these elements. Building on over 20 years of psychological research on positive affect, we've shown it's possible to build a model that reflects what we know about its role in decision making while maintaining the assumption of rationality. Moreover, we believe, that we've shown this combination offers real explanatory power with regard to real-life behavior.

[22] The idea that guilt is, for some, a persistent emotion is borne out by numerous anecdotes of people who devote large portions of their lives seeking to atone for their misdeeds or the misdeeds of their family or people.

Appendix

Proof of Proposition 2: We'll first prove a related proposition: If there exist I and S, both *in* $\mathcal{U}$, such that

$$\begin{aligned} U^*(I) - I &\geq 0 \geq U^*(S) - S \text{ and} \quad (12) \\ I &\leq u_0 \leq S, \quad (13) \end{aligned}$$

then the dynamic process $u_t = U^*(u_{t-1})$ has at least one stable fixed point. To see this, observe first, since $U^*(\cdot)$ is continuous and increasing, that $U^*([I, S]) = [U^*(I), U^*(S)] \subset [I, S]$. From the Brouwer Fixed-Point Theorem, it follows, then, that $U^*(\cdot)$ has at least one fixed point. Suppose, first, that $I < U^*(I) < U^*(S) < S$. Then, since $U^*(\cdot)$ starts out above the 45° line and ends below it, it must *cross* it at least once from above. Let u^* be such a point. Since $U^*(\cdot)$ crosses from above, $U^{*\prime}(u^*) \leq 1$. Condition (10) rules out equality; that is, $U^{*\prime}(u^*) < 1$. It follows from Theorem 1.12 of Elaydi (1996) that u^* is a stable fixed point. If $I = U^*(I)$ or $U^*(S) = S$ or both, then I or S is a fixed point, respectively, or they both are. Assume $I = U^*(I)$ (the case for S is similar). If $u_0 = I$, we're done—the process will stay there. Assume, then, that $u_0 > I$. If $U^{*\prime}(I) > 1$, then there exists an $I_+ \in (I, u_0)$ that satisfies the conditions (12) and (13) with $U^*(I_+) > I_+$, so the above argument can be applied. Otherwise, $U^{*\prime}(I) < 1$; so I is, itself, a stable fixed point. This establishes our related proposition.

We now show that the assumptions in the "menu" imply conditions (12) and (13). If A(ii) holds, then, since $U^* : \mathcal{U} \to \mathcal{U}$, $U^*(I_{\mathcal{U}}) \geq I_{\mathcal{U}}$, and we can set $I = I_{\mathcal{U}}$ in conditions (12) and (13). A similar argument works for B(ii). If $I_{\mathcal{U}} \notin \mathcal{U}$, then the interval $(I_{\mathcal{U}}, u_0) \subset \mathcal{U}$ is non-empty. By A(i), there exists an $\bar{I} \in \mathcal{U}$ such that $U^*(u) - u > \varepsilon_I > 0$ for $u \in (I_{\mathcal{U}}, \bar{I}) \subset \mathcal{U}$. Now pick I from the intersection of $(I_{\mathcal{U}}, u_0)$ and $(I_{\mathcal{U}}, \bar{I})$. Clearly, this I satisfies (12) and (13). A similar argument works for B(i). ■

Proof of Proposition 4: Suppose that player j is playing this strategy. We need to show that it is a best response for player i to do the same. As before, define $\theta(q) = q - q^2/2$. Then, for player j,

$$u_t^j = \alpha^j x_t^i + \theta(\mu^*) u_{t-1}^j.$$

Solving this difference equation yields

$$u_t^j = \alpha^j \sum_{\tau=1}^{t} \theta(\mu^*)^{t-\tau} x_\tau^i + \theta(\mu^*)^t u_0^j$$

(note we're employing the convention that $\sum_{\tau=1}^{0} q(\tau) = 0$). Player i's utility is

$$\begin{aligned} u_t^i &= x_t^i - \frac{(x_t^i)^2}{2u_{t-1}^i} + \alpha^i \mu^* u_{t-1}^j \\ &= x_t^i - \frac{(x_t^i)^2}{2u_{t-1}^i} + \alpha^i \mu^* \alpha^j \sum_{\tau=1}^{t-1} \theta(\mu^*)^{t-1-\tau} x_\tau^i + \alpha^i \mu^* \theta(\mu^*)^{t-1} u_0^j. \end{aligned}$$

If i is playing a best response to j, then her strategy, $\{x_t^i\}_{t=1}^{\infty}$, maximizes

$$\sum_{t=1}^{\infty} \delta^t u_t^i = \sum_{t=1}^{\infty} \delta^t \left(x_t^i - \frac{(x_t^i)^2}{2u_{t-1}^i} + \mu^* \alpha^i \alpha^j \sum_{\tau=1}^{t-1} \theta(\mu^*)^{t-1-\tau} x_\tau^i + \alpha^i \mu^* \theta(\mu^*)^{t-1} u_0^j \right).$$

The "principle of optimality" (see, e.g., Stokey and Lucas, 1989, §4.1) tells us that we can solve this maximization problem by ensuring that each x_t^i maximizes the discounted sum of all the utilities it *directly* affects. This yields the first-order condition for x_t^i:

$$\delta^t \left(1 - \frac{x_t^i}{u_{t-1}^i} + \delta\mu^* \alpha^i \alpha^j \sum_{\tau=0}^{\infty} \delta^\tau \theta \left(\mu^*\right)^\tau \right) = 0.$$

Hence,

$$x_t^i = \left(1 + \frac{\delta\mu^* \alpha^i \alpha^j}{1 - \delta\theta\left(\mu^*\right)}\right) u_{t-1}^i.$$

Observe that the expression in the large parentheses is time-*in*variant. The proof is, therefore, complete if

$$\mu^* = 1 + \frac{\delta\mu^* \alpha^i \alpha^j}{1 - \delta\theta\left(\mu^*\right)}. \tag{14}$$

Tedious algebra reveals that is precisely what it equals.[23] ■

References

Akerlof, George A., "Labor Contracts as Partial Gift Exchange," *Quarterly Journal of Economics*, November 1982, *97* (4), 543–569.

Ashby, F. Gregory, Alice M. Isen, and And U. Turken, "A Neuropsychological Theory of Positive Affect and its Influence on Cognition," *Psychological Review*, 1999. Forthcoming.

Batson, C. Daniel, *The Altruism Question: Toward a Social-psychological Answer*, Hillsdale, NJ: Erlbaum, 1991.

Becker, Gary S., *The Economic Approach to Human Behavior*, Chicago: University of Chicago Press, 1976.

Benhabib, Jess and Richard H. Day, "Rational Choice and Erratic Behaviour," *Review of Economic Studies*, July 1981, *48* (3), 459–471.

Damasio, Antonio R., *Descartes' Error*, New York: Avon Books, 1995.

Elaydi, Saber N., *An Introduction to Difference Equations*, Berlin: Springer-Verlag, 1996.

Elster, Jon, "Emotions and Economic Theory," *Journal of Economic Literature*, March 1998, *36* (1), 47–74.

Frank, Robert H., *Passions within Reason*, New York: Norton, 1988.

Freĭlin, Mark I. and Alexander D. Wentzell, *Random Perturbations of Dynamical Systems*, Berlin: Springer-Verlag, 1984.

[23] It can be shown that the equation (14) has only one real solution (the other solutions are imaginary).

Isen, Alice M., "Positive Affect and Decision Making," in Michael Lewis and Jeannette M. Haviland, eds., *Handbook of Emotions*, 2nd ed., New York: The Guilford Press, 1999. In press.

_ **and John Marshall Reeve**, "The Influence of Positive Affect on Intrinsic Motivation," 1992. Working paper, Cornell University.

_ **and Paula F. Levin**, "Effect of Feeling Good on Helping: Cookies and Kindness," *Journal of Personality and Social Psychology*, March 1972, *21* (3), 384–388.

_ **, Kimberly A. Daubman, and Gary P. Nowicki**, "Positive Affect Facilitates Creative Problem Solving," *Journal of Personality and Social Psychology*, June 1987, *52* (6), 1122–1131.

_ **, Thomas E. Nygren, and F. Gregory Ashby**, "Influence of Positive Affect on the Subjective Utility of Gains and Losses; It is Just not Worth the Risk," *Journal of Personality and Social Psychology*, 1988, *55* (5), 710–717.

Johanson, Donald, *From Lucy to Language*, New York: Simon & Schuster, 1996. With contributions by Blake Edgar.

Johnston, Victor S., *Why We Feel*, Reading, MA: Perseus Books, 1999.

Kahn, Barbara and Alice M. Isen, "The Influence of Positive Affect on Variety Seeking among Safe, Enjoyable Products," *Journal of Consumer Research*, 1993, *20*, 257–270.

Kaufman, Bruce E., "Emotional Arousal as a Source of Bounded Rationality," *Journal of Economic Behavior & Organization*, 1999, *38*, 135–144.

Laibson, David I., "A Cue-Theory of Consumption," 1996. Working paper, Department of Economics, Harvard University.

LeDoux, Joseph, *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*, New York: Simon and Schuster, 1998.

Lewin, Shira B., "Economics and Psychology: Lessons for Our Own Day from the Early Twentieth Century," *Journal of Economic Literature*, September 1996, *34* (3), 1293–1323.

MacLeod, W. Bentley, "Decision, Contract, and Emotion: Some Economics for a Complex and Confusing World," 1996. Working paper, C.R.D.E., Université de Montréal.

Rabin, Matthew, "Psychology and Economics," *Journal of Economic Literature*, March 1998, *36*, 11–46.

Romer, Paul M., "Thinking and Feeling," 1999. Working paper, Graduate School of Business, Stanford University.

Shapiro, Carl and Joseph E. Stiglitz, "Equilibrium Unemployment as a Worker Discipline Device," *American Economic Review*, June 1984, *74* (3), 433–444.

Stokey, Nancy L. and Robert E. Lucas Jr., *Recursive Methods in Economic Dynamics*, Cambridge, MA: Harvard University Press, 1989. With Edward C. Prescott.

von Auer, Ludwig, *Dynamic Preferences, Choice Mechanisms, and Welfare*, Berlin: Springer-Verlag, 1998.

Yerkes, R.M. and J.D. Dodson, "The Relation of Strength of Stimulus to Rapidity of Habit-Formation," *Journal of Comparative Neurology and Psychology*, 1908, *18*, 459–482.